Footprinting Titration Studies on the Binding of Echinomycin to DNA Incapable of Forming Hoogsteen Base Pairs†

Eric W. Sayers‡ and Michael J. Waring*

Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, England

Received March 1, 1993; Revised Manuscript Received June 7, 1993

ABSTRACT: In order to investigate the possible importance of Hoogsteen base pairing to the DNA-binding ability of echinomycin, quantitative DNase I footprinting has been performed. The substrate was the *tyrT* DNA restriction fragment, either "native" or substituted with one of the purine analogs 2′-deoxy-7-deazaadenosine and 2′-deoxy-7-deazaguanosine in both strands. The modified DNA species were prepared by PCR and selectively labeled at the 5′ terminus of one strand (usually the upper "Watson" strand) with [^{32}P]ATP and polynucleotide kinase. Proper incorporation of the analog nucleotides was verified by Maxam–Gilbert G- and C-sequencing reactions as well as exposure to osmium tetroxide and diethyl pyrocarbonate. OsO_4 was found to react strongly with the 7-deaza nucleotides, providing a good check of faithful incorporation. The previously observed echinomycin-induced hyperreactivity of purines toward diethyl pyrocarbonate was eliminated by incorporating the appropriate 7-deazapurine. The DNase I footprinting titration studies greatly refined the existing knowledge of the DNA-binding characteristics of echinomycin, as they revealed five general types of concentration-dependent behavior at single-bond resolution. Estimates of microscopic binding constants at individual DNA binding sites were obtained by measuring the antibiotic concentration which produced a half-maximal effect on the concentration of a given DNase I cleavage product. All binding sites contained one or more CpG steps, and all CpG steps analyzed formed part of a binding site for echinomycin. No consistent differences in the estimated binding constants for these sites were observed by comparing normal and modified DNAs, indicating that the abolition of formal Hoogsteen pairs did not significantly alter the thermodynamics of echinomycin–DNA interaction. The lack of any detectable decrease in binding constants for critical sites in the 7-deazapurine-substituted DNAs argues against any *anti–syn* conformational transition of purine nucleosides occurring in association with the bis-intercalative complex formation.

Echinomycin is a member of the family of quinoxaline antibiotics, a group of drugs which contain two quinoxaline chromophores attached to a cyclic depsipeptide backbone (Figure 1). First isolated from *Streptomyces echinatus* (Corbaz et al., 1957), it was shown to be active against Gram-positive bacteria, viruses, and several tumors (Katagiri et al., 1975). More recently, echinomycin has been brought into phase I and phase II clinical trials against various forms of cancer (Foster et al., 1986; Muss et al., 1990a,b). Early investigations revealed that the antibiotic binds to DNA, and it was suggested that this binding, and the concomitant disabling of DNA transcription and replication, was the basis for the antitumor activity of the antibiotic (Ward et al., 1965; Sato et al., 1967). Later studies revealed that echinomycin binds by bis-intercalating into DNA (Waring & Wakelin, 1974; Wakelin & Waring, 1976), thereby sandwiching two base pairs between its two chromophores (Waring, 1979, 1990). This model was further refined by DNase I and methidium propyl-ETDA/Fe(II) footprinting studies, both of which found that the dinucleotide step CpG is present in the preferred binding sites of echinomycin (Low et al., 1984; Van Dyke & Dervan, 1984). It should, however, be noted that echinomycin can bind to poly(dA-dT) as well as to poly(dC-dG) (Wakelin & Waring, 1976), indicating that the CpG step is not the only possible binding site. Indeed, it has recently been demonstrated that echinomycin can bind weakly to CpA steps (Phillips et al., 1990).

At the same time as the initial footprinting experiments were being performed, quinoxaline–DNA complexes were being examined by X-ray crystallography. The first complex whose structure was solved was that between triostin A, a close relative of echinomycin, and the self-complementary hexamer d(CGTACG) (Wang et al., 1984). The structures of complexes of echinomycin with the same hexamer (Ughetto et al., 1985) and of triostin A with the octamer d(GCGTACGC) (Quigley et al., 1986; Wang et al., 1986) soon followed. In all of these crystal structures, the quinoxaline antibiotic was found to bis-intercalate around the CpG steps in the oligonucleotide as expected. The complexes were stabilized by three hydrogen bonds between L-alanine residues of the antibiotic and guanosines in the oligonucleotide, together with a host of van der Waals interactions. These structures thus confirmed the expected ability of echinomycin (and triostin A) to bis-intercalate into DNA.

One unexpected finding from these structural studies was the observation of non-Watson–Crick base pairing in the antibiotic–oligonucleotide complexes. The internal AT pairs of all three complexes, as well as the terminal GC pairs of the octamer complex, were found to have shifted from Watson–Crick pairs to Hoogsteen pairs, an unusual pairing arrangement where the purines adopt a *syn* conformation (Figure 2). As shown, the Hoogsteen GC pair requires the protonation of N3 of cytidine. The apparent advantage of Hoogsteen pairing arises from the observation that the sugar–phosphate backbones of the oligonucleotides are brought almost 2 Å closer together than they would be with Watson–Crick pairs, allowing them to fold more closely around the antibiotic (Quigley et

† This work was supported by grants from the Cancer Research Campaign and the Royal Society. E.W.S. acknowledges the award of a Churchill scholarship funded in part by Glaxo Ltd.

* To whom correspondence should be addressed.

‡ Present address: Department of Pharmacology, Yale University, New Haven, CT 06510.

0006-2960/93/0432-9094$04.00/0

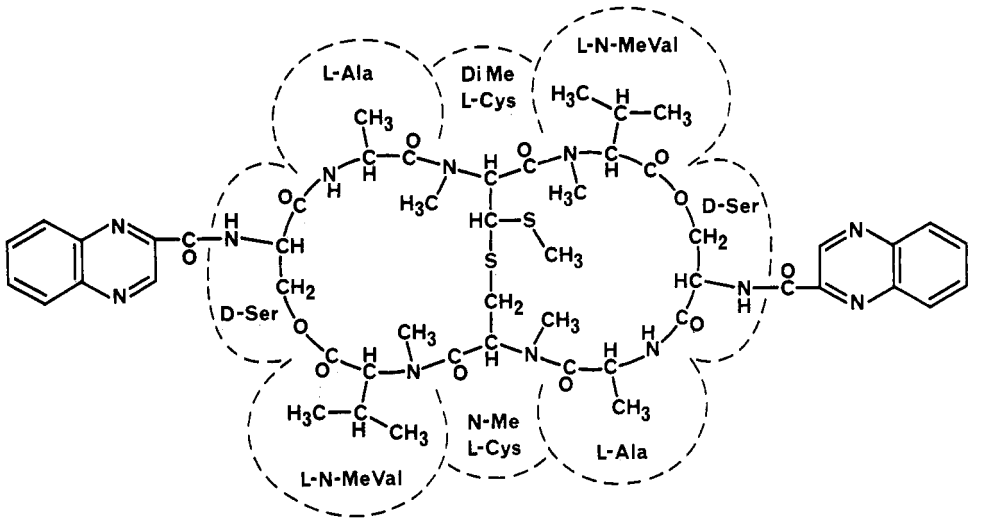

FIGURE 1: Chemical structure of echinomycin.

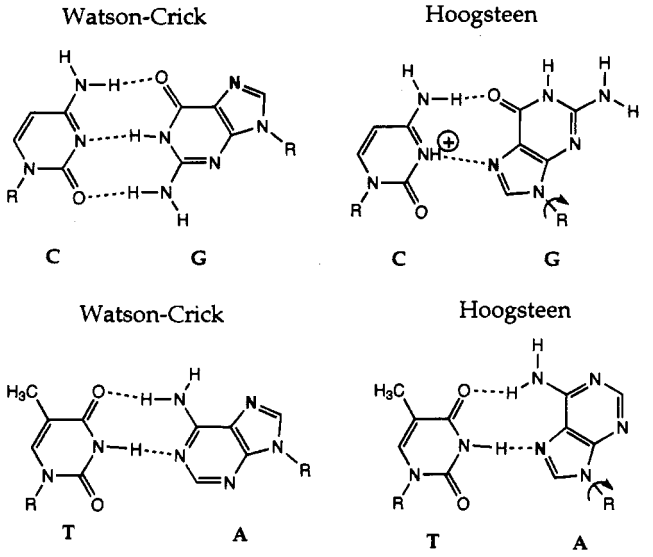

FIGURE 2: Structures of Watson–Crick and Hoogsteen base pairs. The N7 of each purine is in bold type, and dashed lines indicate hydrogen bonding. Arrows indicate the rotation around the glycosidic bond whereby the purines shift from the *anti* conformation of Watson–Crick pairs to the *syn* conformation of Hoogsteen pairs. Hoogsteen pairs are stabilized by the ability of the N7 to hydrogen bond to the 3-NH group of the pyrimidines. Notice that in the case of cytosine, this requires the protonation of N3. In all cases, R = 2′-deoxyribofuranoside.

al., 1986). Later NMR studies confirmed the presence of Hoogsteen pairs, although their formation was found to be highly dependent on both sequence and temperature. The first complexes thus studied were those between echinomycin and the tetramers d(ACGT) and d(TCGA). The two AT pairs in the d(ACGT) complex were found to be Hoogsteen paired, while the AT pairs in the d(TCGA) complex were Watson–Crick paired (Gao & Patel, 1988). Interestingly, NMR studies of the complex between echinomycin and d(AAACGTTT) revealed that all AT pairs were Watson–Crick paired (Gao & Patel, 1989). Similar studies of the complexes between echinomycin and the oligomers d(GCGC) and d(CCGG) revealed that the terminal GC pairs in the d(GCGC) complex were Watson–Crick paired at neutral pH, but that they shifted to Hoogsteen pairing as the pH was lowered, with a pK_a of 5.1 for the transition. However, no Hoogsteen pairs were observed in the d(CCGG) complex at any pH (Gao & Patel, 1989). Further studies involving a complex between echinomycin and the octamer d(ACGTACGT) revealed that all four AT pairs adopt stable Hoogsteen pairs at 1 °C. But as the temperature was raised to 45 °C, the two internal AT pairs appeared to be exchanging between Hoogsteen pairing and either an open or a Watson–Crick conformation. The two terminal AT pairs remained stably Hoogsteen paired throughout the temperature range (Gilbert et al., 1989). Similar studies on the echinomycin complex with the octamer d(TCGATCGA) revealed that all AT pairs adopt the Watson–Crick conformation throughout the temperature range 1–45 °C (Gilbert & Feigon, 1991). In summary, stable Hoogsteen pairs were observed only in the terminal base pairs of complexes having purines 5′ and pyrimidines 3′ to the CpG (i.e., ACGT, GCGC), and only at low pH in the case of GC pairs. That the formation of Hoogsteen GC pairs should be pH dependent is not surprising, since cytidine must be protonated in order for a Hoogsteen pair to form. Significantly, internal AT pairs (which are more constrained by the helix than terminal ones) appeared to be unable to adopt a stable Hoogsteen pair (and possibly any base pairing at all) at physiological temperature (Gilbert et al., 1989). Moreover, Hoogsteen pairing at terminal AT pairs was eliminated simply by adding two additional terminal AT pairs. Clearly, echinomycin appears to cause dramatic changes upon binding to oligonucleotides. The question remains as to whether similar effects occur in full-length DNA and specifically whether or not they include the formation of Hoogsteen pairs.

Several groups have sought for a specific means to recognize Hoogsteen pairs in DNA restriction fragments. The first probe tried was DEPC, which displayed hyperreactivity toward purines both proximal and distal to echinomycin binding sites (Mendel & Dervan, 1987), but the hyperreactivity was eventually traced to echinomycin-induced unwinding of the DNA helix rather than Hoogsteen base pair formation (Portugal et al., 1988; McLean & Waring, 1988; Jeppesen & Nielsen, 1988). Further studies using osmium tetroxide (OsO_4), potassium tetrachloropalladate, and dimethyl sulfate produced no evidence for the existence of Hoogsteen pairs (McLean & Waring, 1988). Most recently, the OsO_4 modification reaction and DNase I footprinting were used to investigate the binding of echinomycin to DNA containing 2′-deoxy-7-deazaadenosines in place of adenosines in one strand (McLean et al., 1989). Because such 7-deazapurines lack the N7 which is required for Hoogsteen pair formation,

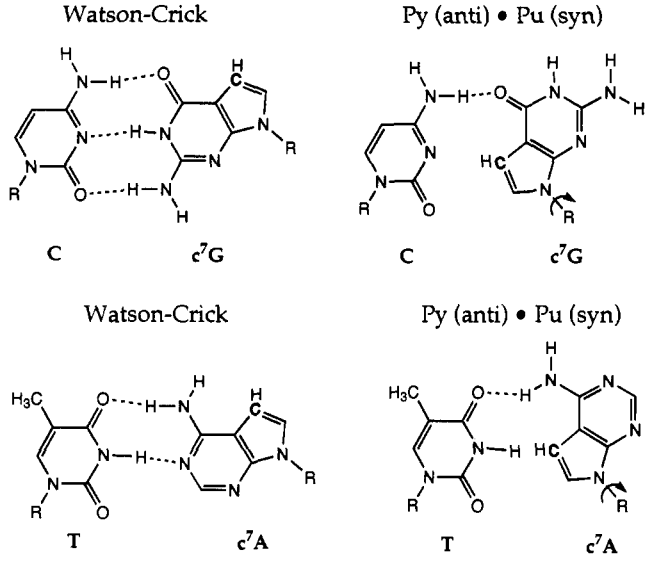

FIGURE 3: Structures of Watson–Crick and Py(*anti*)·Pu(*syn*) base pairs between normal pyrimidines and 7-deazapurines. The C7 of each purine analog is in bold type, and dashed lines indicate hydrogen bonding. Rotation around the glycosidic bond of the 7-deazapurine nucleosides (arrowed) allows only one hydrogen bond to be formed because of the absence of the N7. Consequently, these purine analogs are incapable of forming Hoogsteen pairs. In all cases, R = 2′-deoxyribofuranoside, c^7A = 2′-deoxy-7-deazaadenosine, and c^7G = 2′-deoxy-7-deazaguanosine.

DNA containing them is incapable of forming stable Hoogsteen pairs (Figure 3). However, it should be noted that it remains possible for 7-deazapurines to adopt a *syn* conformation, which would result in a weak Hoogsteen-like pair stabilized by only one hydrogen bond. Thus, DNA containing these analogs affords a means of determining the importance, if any, of stable Hoogsteen base pair formation as a determinant of the ability of echinomycin to bind to DNA. When DNA containing c^7A in one strand was subjected to DNase I footprinting and OsO_4 modification, no differences were observed between this modified DNA and normal DNA (McLean et al., 1989). Thus, preventing Hoogsteen A-T base pair formation did not prevent echinomycin from binding to DNA.

In this paper we report the results of extended experiments employing 7-deazapurines to study the importance of Hoogsteen pairs to the DNA-binding ability of echinomycin. Using the technique of PCR, c^7A and c^7G were incorporated into both strands of the *tyrT* fragment. Conveniently, both c^7-dATP and c^7dGTP are acceptable substrates for DNA polymerases. This was demonstrated for c^7dATP by the studies mentioned above (McLean et al., 1989). c^7dGTP has long been known to be a substrate for DNA polymerases and is now routinely used to reduce "G-compression" in dideoxy sequencing protocols (Mizusawa et al., 1986). Moreover, there are already reports in the literature of the synthesis of DNA containing c^7G in both strands using PCR (McConlogue et al., 1988; Seela & Röling, 1991). Through investigating the complexes of echinomycin with this modified DNA compared to normal DNA, we aimed to address the following questions: Upon binding to full-length DNA, does echinomycin cause purine nucleotides adjacent to the binding site to adopt a *syn* conformation, as would be required for the formation of Hoogsteen pairs? Moreover, if Hoogsteen pairs are indeed formed, how important are they to the overall stability of the complex? To answer these questions, complexes of echinomycin with both normal and modified DNA were subjected to quantitative DNase I footprinting as well as modification with DEPC and OsO_4.

```
5'-  AATT  CCGGTTACCT  TTAATC -3'  (Watson primer)
5'-  AATT  CCGGTTACCT  TTAATCCGTT  ACGGATGAAA  ATTACGCAAC
          0          10          20          30          40
3'-       GGCCAATGGA  AATTAGGCAA  TGCCTACTTT  TAATGCGTTG


          CAGTTCATTT  TTCTCAACGT  AACACTTTAC  AGCGGCGCGT
                    50          60          70          80
          GTCAAGTAAA  AAGAGTTGCA  TTGTGAAATG  TCGCCGCGCA


          CATTTGATAT  GATGCGCCCC  GCTTCCCGAT  AAGGGAGCAG
                    90         100         110         120
          GTAAACTATA  CTACGCGGGG  CGAAGGGCTA  TTCCCTCGTC


          GCCAGTAAAA  AGCATTACCC  CGTGGTGGGG  GTTC-3'
                   130         140         150         158
          CGGTCATTTT  TCGTAATGGG  GCACCACCCC  CAAGGGCT   -5'
            (Crick primer) 3'-GG  GCACCACCCC  CAAGGGCT   -5'
```

FIGURE 4: Sequence of the *tyrT* fragment and the two 20-mer oligonucleotide PCR primers. Numbers indicate the bond numbering system for this fragment, with a given bond having the number of the nucleotide on the 5′ "end" of the bond. For example, in the case of bond 58, which is a CpG, the cytidine (in the Watson strand) is numbered 58.

MATERIALS AND METHODS

Materials. Echinomycin was obtained from Ciba-Geigy or prepared in our own laboratory (Gauvreau & Waring, 1984). No differences were detected between different preparations. The antibiotic was dissolved to concentrations of 50 or 100 μM in a mixture containing 6 mM Tris-HCl, pH 7.5, 6 mM NaCl, and 40% methanol (60/40 buffer) and stored at 4 °C. Methanol was included because of the low aqueous solubility of echinomycin (5 μM) (Waring, 1979). This stock was diluted to working concentrations with appropriate volumes of 10 mM Tris–HCl, pH 7.5, 10 mM NaCl (TN buffer), and 60/40 buffer to yield a final methanol concentration of 10% or 20% in the footprinting reactions. To investigate whether or not the presence of methanol affected the footprinting reactions, control experiments have been performed by previous workers (Low et al., 1984; Portugal et al., 1988) and in the current study. In all cases, no differences were found between reactions performed in purely aqueous conditions and in the presence of methanol. Oligonucleotides for PCR amplification were a generous gift from Dr. M. McLean. *Taq* polymerase was obtained from Beckman. *Ava*I, *Eco*RI, and T4 polynucleotide kinase were obtained from New England Biolabs. Deoxyribonuclease I (DNase I) from bovine pancreas was obtained as a solid from Sigma and was dissolved to a concentration of 7200 units/mL in 0.15 M NaCl containing 1 mM $MgCl_2$. Immediately before use, DNase I was diluted to working concentrations with 20 mM NaCl containing 2 mM $MgCl_2$ and 2 mM $MnCl_2$.

Preparation of tyrT Fragment. The *tyrT* restriction fragment was isolated as previously described from the plasmid pKMΔ-98, which was extracted from transformed *Escherichia coli* NF 58 (Drew & Travers, 1984). The transformed *E. coli* was a generous gift from Dr. Andrew Travers. The sequences of the *tyrT* fragment as well as the oligonucleotide primers are shown in Figure 4.

Strategy for End-Labeling PCR Products. Footprinting experiments require that the DNA templates be radioactively end-labeled on one strand only, and therefore a strategy for so labeling the blunt-ended PCR products was needed. The chosen strategy was simply to ensure that one of the two primers used in the PCR amplifications contained an amino group at its 5′ end. Thus, the 5′ end of one of the strands of the resulting PCR product was blocked by an amino group, allowing only the other strand to be phosphorylated by T4 polynucleotide kinase. Moreover, it is important to note that the original *tyrT* template already bore a 5′-phosphate due to the action of *Eco*RI, which was used to excise the *tyrT* fragment from

the pKMΔ-98 plasmid. Thus, only newly synthesized DNA (and therefore only modified DNA in reactions containing dc^{7}NTPs) should be labeled by the kinase.

Preparation of NH_2-Blocked PCR Primers. Because of the possibility that the NH_2-blocked primers might be contaminated by small amounts of shorter oligonucleotides possessing 5′-OH groups, it was necessary to prevent any PCR products resulting from these contaminants from being labeled by the kinase. Thus, crude primer (30 μg) was exposed to phosphorylation by T4 polynucleotide kinase (20 units) in the presence of 2 mM cold ATP. The reaction was carried out at 37 °C for 30 min in a buffer containing 70 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, and 5 mM DTT. This resulted in phosphorylation of contaminants and thus blocked them from further phosphorylation by the kinase. After phosphorylation, the primer was collected by ethanol precipitation and resuspended in TN buffer at a concentration of 1 mg/mL.

Polymerase Chain Reaction (PCR). PCR reaction mixtures contained 10 ng of *tyrT* template, 1 μg each of the appropriate pair of primers (one with a 5′-OH and one with a 5′-NH_2) required to allow 5′-phosphorylation of the desired strand, 250 μM each of appropriate dNTPs and c^{7}dNTP, and 5 units of *Taq* polymerase in a volume of 50 μL containing 50 mM KCl, 10 mM Tris-HCl, pH 8.3, and 1.5 mM $MgCl_2$. To prevent unwanted primer–template annealing before the cycles began, the reactions were heated to 60 °C before adding *Taq* polymerase (Bloch, 1991). Finally, 50 μL of paraffin oil was added to each reaction to prevent evaporation. After an initial denaturing step of 3 min at 94 °C, 30 amplification cycles were performed, with each cycle consisting of the following segments: 94 °C for 1 min, 37 °C for 2 min, and 72 °C for 3 min. After the last cycle, the extension segment was continued for an additional 10 min at 72 °C, followed by a 5-min segment at 55 °C and a 5-min segment at 37 °C. The purpose of these final segments was to maximize annealing of full-length product and to minimize annealing of unused primer to full-length product. The reaction mixtures were then extracted with chloroform to remove the paraffin oil, and parallel reactions were pooled. If necessary, to reduce the volume prior to purification, the DNA was collected by ethanol precipitation. The resulting pellets were redissolved in TN buffer, and the DNA was purified by 6% polyacrylamide gel electrophoresis. In most cases the amount of PCR product was too small to be determined spectrophotometrically, and thus the DNA was quantified by comparing the ethidium bromide fluorescence of spotted samples to that of salmon sperm DNA standards (Maniatis et al., 1982). It should be noted that in the case of c^{7}G-containing DNA, this method of quantification yielded apparent quantities approximately one-third of those implied by the radioactive concentration of 5′-end-labeled DNA, based on repeated labeling experiments with normal and c^{7}G-containing DNA in parallel. The reason for this is probably related to the fluorescence properties of ethidium intercalated into c^{7}G-containing DNA.

5′-Labeling of tyrT Fragment. 5′-Labeling with T4 polynucleotide kinase was performed according to a standard procedure for labeling blunt-ended DNA fragments (Maniatis et al., 1982). Approximately 100 ng of DNA was labeled in each reaction. After completion, the reaction mixture was heated to 65 °C to denature the enzyme and extracted with phenol/chloroform. Yeast tRNA was added as a carrier to a final concentration of 100 μg/mL, and the DNA was collected by two successive precipitations with ethanol, followed by two washes with 70% ethanol. The DNA pellet was dried and resuspended in 50 μL of TN buffer. The DNA was then diluted to working concentrations with TN buffer containing 100 μg/mL tRNA and stored frozen at −20 °C.

Characterization of PCR Products. Because PCR was being used as a preparative procedure, the optimal conditions for the PCR reactions were those that would maximize the total yield of the amplification. It has been observed that amplification generally stops once the concentration of DNA strands reaches approximately 10^{-7} M (Bloch, 1991), and thus the initial concentration of template should be well below this value to allow amplification to proceed. However, in general it was found that smaller amounts of template led to lower total yields. In the case of synthesis of normal DNA, 10 ng of template (2×10^{-9} M) produced a total yield of 2.5 μg (5×10^{-7} M) representing a 250–fold amplification. When the amount of template was reduced to 1 pg (2×10^{-13} M), the total yield was only 250 ng (5×10^{-8} M), even though the amplification was 250 000–fold. This effect was even more pronounced for reactions containing 7-deazapurines, which generally gave lower levels of amplification and total yields. For example, reaction mixtures containing c^{7}A and using 10 ng of template produced amplifications of 10- to 20-fold, whereas reaction mixtures using 1 pg of template produced no product visible by ethidium bromide staining. Reaction mixtures containing c^{7}G produced intermediate results, with amplifications of approximately 50-fold when 10 ng of template was used. The optimal conditions were found to be PCR reactions where 10 ng of template was amplified for 30 cycles. These yielded acceptable product purities of >97% and >90% for c^{7}G-DNA and c^{7}A-DNA, respectively.

The normal and c^{7}G-DNA PCR products comigrated with the *tyrT* template, indicating a successful synthesis of full-length fragment. This was also the case with c^{7}A-DNA, but in most cases this DNA was contaminated by shorter fragments visible only in denaturing gels. In these cases, the c^{7}A-DNA was first purified on a denaturing gel and then renatured prior to 5′-end-labeling. The fact that the shorter contaminating fragments present in c^{7}A-containing reactions were only observed in denaturing gels suggests that these short species were originally annealed to a full-length strand under native conditions. Such species could easily be the result of premature chain termination by *Taq* polymerase. Indeed, the lengths of the principal contaminants were found to be consistent with the hypothesis that they resulted from premature termination at sequences of repeated adenosines, without doubt the most troublesome sites for *Taq* polymerase attempting to incorporate c^{7}A. Recently, another group was unable to produce DNA containing c^{7}A using *Taq* polymerase (Seela & Röling, 1992). While we also observed that PCR reactions containing c^{7}A proceeded poorly compared to reactions containing either normal nucleotides or c^{7}G, it is likely that our considerably shorter template as well as the act of pooling several parallel reactions aided in visualizing full-length products.

Verification of the Correct Incorporation of 7-Deazapurines. Four chemical probes were employed to verify that 7-deazapurines had been correctly incorporated into the DNA by PCR. These were DMS/piperidine, hydrazine/piperidine (Maxam–Gilbert G- and C-tracks), OsO_4 (McLean & Waring, 1988), and DEPC (Portugal et al., 1988). The Maxam–Gilbert G-track and C-track from normal and c^{7}A-DNA both showed the typical G- and C-specific patterns, respectively. With c^{7}G-DNA a small amount of reaction was observed with DMS/piperidine at the positions of putative c^{7}G residues. It could be argued that this reactivity was due to cleavage of contaminating molecules of the original template, which contain normal guanosine. However, this

idea can be refuted since the template comprises less than 5% of the c^{7}G-DNA. Moreover, the template has a 5′-phosphate and thus should not be labeled by T4 polynucleotide kinase. A more likely explanation is that, in the absence of the N7, DMS reacted to a lesser extent at another guanosine nitrogen. Such reaction would presumably not be seen in normal DNA because of the greater reactivity of the N7.

With regard to the C-track, the reaction with c^{7}G-DNA also generated an unexpected result. In addition to reacting at the cytidine residues, hydrazine also reacted at almost all positions corresponding to guanosines in normal DNA. The probable explanation of this reactivity lies in the fact that hydrazine reacts by a nucleophilic attack on the 5–6 double bond of the pyrimidines (McMurry, 1984). A similar mechanism can easily be written for nucleophilic attack by hydrazine on the 7–8 double bond of c^{7}G (see Figure 3), thus leading to the "extra" bands observed on the autoradiogram. Such a reaction would be impossible with normal guanosine.

In initial experiments where c^{7}G-DNA and c^{7}A-DNA were exposed to OsO_4 for 10 min (as for the normal DNA), an overwhelming amount of reaction was observed, so much so that the upper halves of the lanes for these reactions on the autoradiograms were entirely devoid of bands. A more acceptable amount of reaction was produced by drastically reducing the reaction time, yielding the result shown in Figure 5. Here the normal DNA was reacted for 10 and 15 min, while the modified DNAs were reacted for only 1 and 5 min. The results in the case of normal DNA are as previously observed (McLean & Waring, 1988), with several thymidines displaying hyperreactivity in the presence of echinomycin. However, in the case of c^{7}G-DNA, strong reactions are seen at each of the putative c^{7}G residues in c^{7}G-DNA (compare to the G-track of normal DNA). Likewise for c^{7}A-DNA, strong reactions are seen at each of the putative c^{7}As (right arrowheads). Interestingly, echinomycin seems to have no effect on OsO_4 reactivity in the case of c^{7}G-DNA, while the antibiotic seems to increase the reactivity further in the case of c^{7}A-DNA. These differences can surely be traced to the presence of 7-deazapurines which contain a 7–8 double bond lying relatively unhindered in the DNA major groove. Since OsO_4 is a powerful oxidizing agent which reacts preferentially with alkenes, the modified bases present an ideal substrate for this probe. It is easy to see why no enhanced reaction is observed at thymidines in the modified DNAs in the presence of echinomycin: the reaction at thymidine, which is normally observed over a 10-min reaction, is totally overwhelmed by the reactions at the 7-deazapurines, which are well advanced after only 1 min. That echinomycin is able to increase this reactivity further in the case of c^{7}A-DNA is fully consistent with the well-documented model of the unwinding of AT-rich sequences caused by the bis-intercalation of echinomycin (Low et al., 1984; McLean & Waring, 1988; McLean et al., 1989). Moreover, the observation that no full-length modified DNA remained uncleaved after 10 min of exposure to OsO_4 confirms that the vast majority of the labeled modified DNA truly contained 7-deazapurines and that these analogs had been incorporated accurately.

DEPC modification of normal DNA, c^{7}G-DNA, and c^{7}A-DNA in the absence and presence of echinomycin is illustrated in Figure 6. The typical pattern of echinomycin-induced purine hyperreactivity is clearly seen in the normal DNA lanes, occurring at many adenosines and fewer guanosines. However, this pattern changes drastically in the modified DNAs. With c^{7}A-DNA, echinomycin-induced hyperreactivity is again observed at the same guanosines, but none of the c^{7}A residues

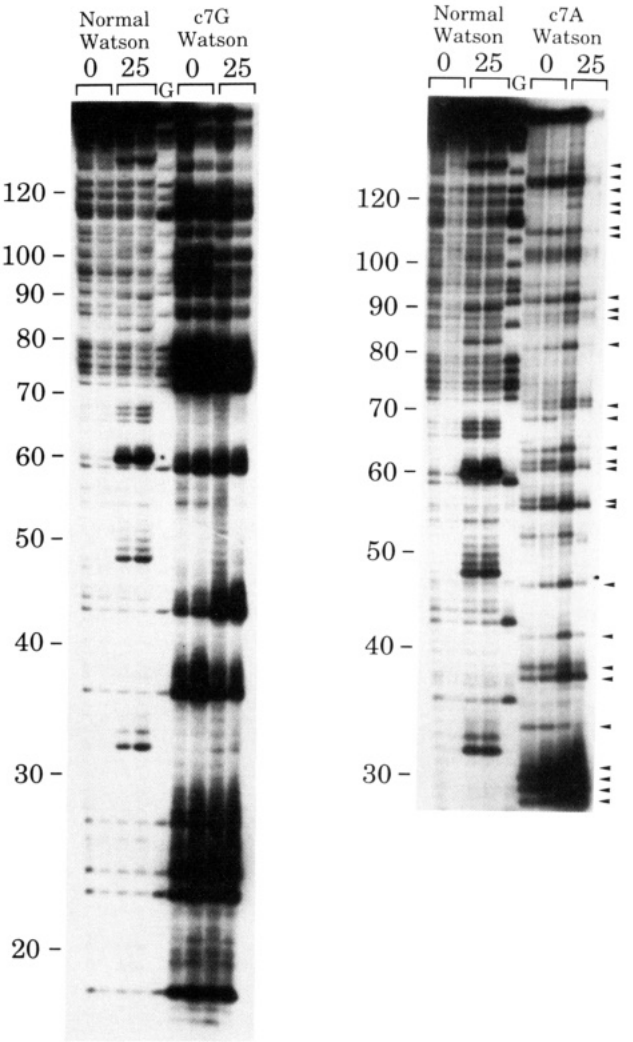

FIGURE 5: Results of OsO_4 modification performed on the Watson strand of normal DNA and c^{7}G- and c^{7}A-DNA. The concentrations of echinomycin (μM) included in the reactions are indicated by numbers above each pair of lanes, while lanes marked with a "G" are the Maxam–Gilbert DMS/piperidine tracks specific for guanosine performed on normal DNA. In the case of normal DNA, the left and right lanes of each pair correspond to reaction times of 10 and 15 min, respectively. In the case of the two modified DNAs, the left and right lanes of each pair correspond to reaction times of 1 and 5 min, respectively. Bond numbers are indicated along the left side of each autoradiogram, while the positions of c^{7}A residues are indicated by arrowheads along the right side of the c^{7}A-DNA lanes.

show any increased reactivity in the antibiotic-containing lanes. In addition, fairly strong reactivity is observed around bond 51 and at some unresolved positions around bonds 100–110. Likewise, the usual hyperreactivity at adenosines is retained in c^{7}G-DNA, while the c^{7}G residues at the positions of the hyperreactive guanosines in normal DNA show no increased reactivity when antibiotic is present. Rather, it appears that *all* of the c^{7}G residues react strongly with DEPC in both the absence and the presence of drug (compare with the G-track) and that this reactivity is not much affected, if at all, by echinomycin. The source of this reactivity is not likely to be the other nitrogens, since they react only minimally in guanosine in the absence of echinomycin. A more plausible possibility is that DEPC may be capable of reacting at the C7 of c^{7}G, as it is known that electrophilic substitutions such as halogenation, sulfonation, azo-coupling, and the Mannich reaction occur at the C7 in most 6-substituted 7-deazapurines (Lunt, 1979). In the case of c^{7}G, a 6-oxo-7-deazapurine, electrophilic attack by DEPC at C7 would result in positive

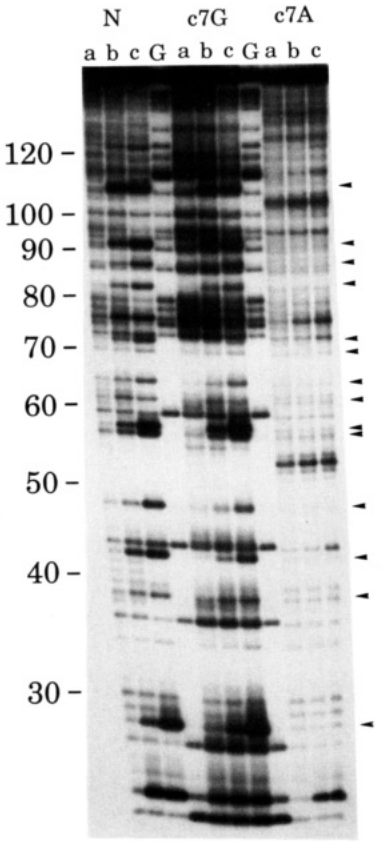

FIGURE 6: Results of DEPC modification performed on the Watson strand of normal DNA (N) and c^{7}G- and c^{7}A-DNA. In each case, lanes labeled "a", "b", and "c" correspond to reactions containing 0, 5, and 25 μM echinomycin, respectively. Lanes marked with a "G" are the Maxam–Gilbert DMS/piperidine tracks specific for guanosine performed on normal DNA. Bond numbers are indicated along the left side of the autoradiogram, while the positions of adenosines hyperreactive to DEPC in normal DNA are indicated by arrowheads along the right side of the autoradiogram.

charge being shifted to the N9, labilizing the glycosidic bond toward cleavage by base. Moreover, this reaction would not be possible with guanosine, as observed.

DNase I Footprinting Reactions. Preliminary "saturation" experiments were performed as previously described (Low et al., 1984). For quantitative footprinting experiments this protocol was modified as follows. A 2-μL sample of 5′-end-labeled DNA (approx. 4 pmol of base pairs) was incubated with 4 μL of either a mixture containing 80% TN buffer and 20% methanol (80/20 buffer) or an appropriate solution of echinomycin in 80/20 buffer for at least 30 min at 37 °C. Following this incubation, a chosen amount of DNase I was added in a volume of 2 μL (typical final concentrations of enzyme were 0.05–0.2 unit/mL). The final concentrations of echinomycin were 0–10 μM in an 8-μL reaction volume containing 11.5 mM NaCl, 6.5 mM Tris-HCl, pH 7.5, 0.5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 10% methanol, and 25 μg/mL yeast tRNA. The reaction was then allowed to proceed at 37 °C for 1 min and was quenched by the addition of 4 μL of a stop solution containing 80% formamide, 10 mM EDTA, and 0.1% each of the dyes bromophenol blue and xylene cyanol. The samples were heated at 95 °C for 5 min prior to electrophoresis.

Gel Electrophoresis, Autoradiography, and Densitometry. Gel electrophoresis, autoradiography, and densitometry were performed as previously described (McLean & Waring, 1988). In the case of the quantitative footprinting experiments, the densitometry was performed three times for each autoradiogram.

Band Assignment. The identity of each DNase I digestion product was assigned from the known sequence and the Maxam–Gilbert sequencing track specific for guanosine. Because the *tyrT* fragment was labeled on the 5′ end in the present studies rather than on the 3′ end as in past work, the guanosine reaction products contained an additional phosphate group and thus migrated approximately 1.5 bonds faster than their DNase I counterparts. Band assignment was adjusted accordingly.

Linearity of the Film Response. The linearity of the film response was determined by preparing serial dilutions of 5′-end-labeled DNA and then carrying out electrophoresis, autoradiography, and densitometry on the unmodified fragment. The response was found to be linear up to an area value which exceeded that of the vast majority of observed bands. The only exceptions were a few of the bands where large enhancements of cutting occurred. For these cases, a corrected area value was calculated from the best-fit curve to the linear response data.

Analysis of the DNase I Footprinting Data. The method of analyzing the data from the titration experiments was based largely on a previous method (Ward et al., 1987). The first step was the normalization of the data to correct for the inherent unevenness of gel loading, thereby reducing "noise" in the data. Since each autoradiogram was scanned three times, each of the three resulting data sets was considered separately in this initial step. The normalization was achieved by adjusting the individual integrated band areas in a lane corresponding to drug concentration, c, by a common factor χ_c. This factor was determined by constructing a "total cut plot", which is a plot of A_c vs c, where A_c is the sum of all band areas in lane c. The factor χ_c was defined as the factor required to adjust the value A_c so that it fell on a smooth curve resulting from a least-squares fit of the points on the total cut plot. The normalized data were then obtained by multiplying each band area in lane c by χ_c. To normalize the three data sets (from the three scanning trials) corresponding to the same drug concentration, the data were weighted so that the A_c values were equal for each value of c. This was done by weighting the data of the two trials which resulted in the two lowest A_c values so that their A_c values equaled the largest A_c value for a given value of c. The normalized area values were then used to construct "footprinting plots", from which values of C_{50} were determined for those sites where drug-induced protection from DNase I cleavage was observed. Footprinting plots were constructed by plotting R vs c, where $R = a_c/a_o$; a_c is the area of the given band in lane c, and a_o is the area of the given band in the control lane. The error bars shown in these plots (see Figures 9 and 10) were derived solely from the triplicate scan/normalization procedure and were not intended to represent the true extent of scatter in the raw data, which in many cases must be substantially greater than the bars drawn. This is particularly true of sites which were cut weakly by DNase I. In order to avoid imposing upon the subsequent analysis any assumptions as to the behavior of the experimental system, especially regarding the potential existence of positive cooperativity, a nonlinear least-squares algorithm (GraphPad InPlot, GraphPad Software, San Diego, CA) was generally used to construct the best-fit curve through the points. While we acknowledge that a function having four variables, which will fit almost any kind of curve, has no relationship to the physical system under study (i.e., drug binding to DNA), we chose this method to avoid any possible introduction of bias.

From each fitted curve a value of C_{50} was determined as follows. For those sites where R reached an identifiable end point, R_{min} (type I), C_{50} was defined as the value of c where R reached half of its total change in the titration (i.e., where $R = (\Delta R/2) + R_{min}$). For sites which exhibited enhancement followed by protection (type II), C_{50} was the value of c where R reached half of its total change during the *protection* phase (i.e., where $R = ((R_{max} - R_{min})/2) + R_{min}$). For sites where no end point was observed (type III), C_{50} was estimated by extrapolating to 0 a line having the initial slope of the protection. Using this line, two "extreme" values of C_{50} were computed. The low extreme was found by assuming that $R_{min} = R$ at 10 μM, and the C_{50} was found as described above for simple protection. The upper extreme was found by assuming $R_{min} = 0$, and again the C_{50} was found as above. The result was a range of C_{50} values in which the "real" C_{50} value most likely exists. The reported C_{50} for such curves was simply the midpoint of this possible range.

Error Analysis for C_{50} Determination. For type I plots, the errors reported for the C_{50} values are estimates generated by a standard analysis of variance. For those sites where a value of R_{min} was determined by nonlinear curve fitting, the software returned an error estimate for R_{min}. Because the value R_{50} (the value of R corresponding to $c = C_{50}$) was determined from R_{min} by simple additions, the error in R_{min} is also the error in R_{50}. Since R can be considered a function of c, the error in c (δc) can be calculated if the error in R (δR) is known as follows:

$$(\delta c)^2 = \frac{(\delta R)^2}{(\delta R/\delta c)^2} \qquad (1)$$

It should be emphasized that the error values generated by this method are only estimates, and as the computer-estimated error is designed to be conservative, they are very likely underestimates. Turning to type II plots, since the R_{min} value used to compute C_{50} values for these plots was not the result of a curve-fitting algorithm, but was rather found by simple visual location of an endpoint, no error estimates were available for these R_{min} values. Thus, no error estimation could be performed for type II plots as was done for type I plots, and all data for type II plots were thus displayed without error bars. Finally, as described above, the C_{50} values for type III plots were calculated as the midpoints of the range of possible C_{50} values. The error for these C_{50} values was simply defined as the extent of the possible range of values. For example, if the range of C_{50} values was 8–12 μM, then the C_{50} was calculated to be 10 μM with the error estimate given as ± 2 μM.

RESULTS

DNase I Footprinting Saturation Experiments. The first question posed in the introduction is whether Hoogsteen pairs must form in order for echinomycin to bind to DNA. To address this question, DNase I footprinting experiments were performed using saturating concentrations of antibiotic. At high concentrations, where all of the stronger binding sites should be maximally protected from DNase I cleavage, any significant reduction in footprinting observed in modified DNAs would indicate a dramatic change in the strength of binding to the affected sites. Previous footprinting studies with echinomycin used antibiotic concentrations of 2–16 μM, where it was found that protection from DNase I cleavage was not observed below 2 μM and did not increase above 8 μM (Low et al., 1984). On the basis of those studies, a concentration of 25 μM was chosen to ensure saturating conditions. The DNAs used were initially only normal DNA and c^{7}G-DNA; the analog c^{7}G was studied first because it was known that it could be successfully incorporated into DNA by PCR (McConlogue et al., 1988; Seela & Röling, 1991).

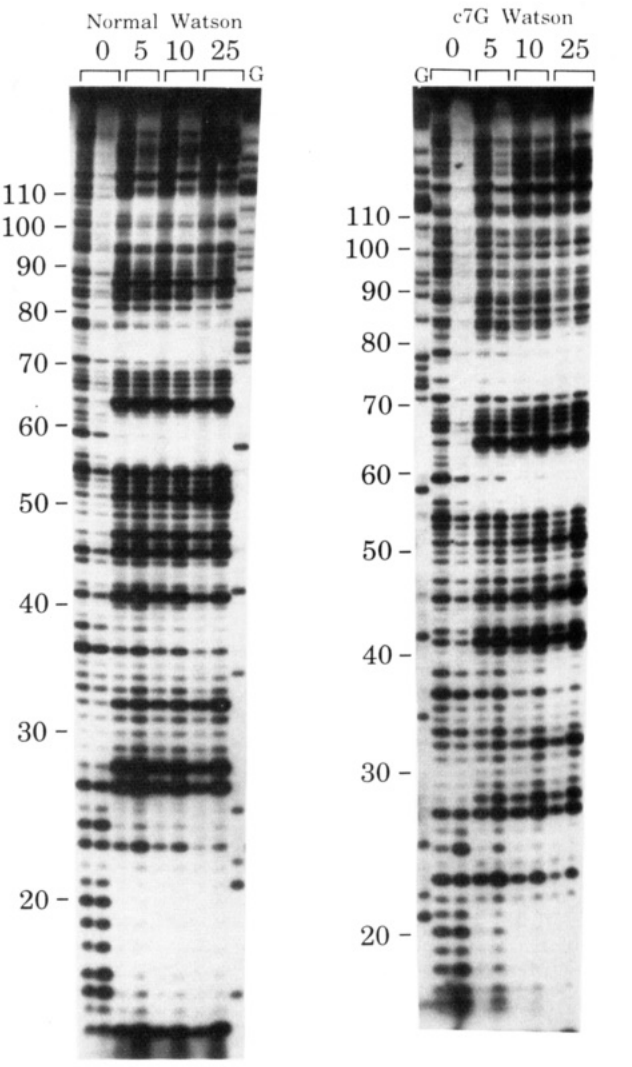

FIGURE 7: Results of DNase I footprinting saturation experiments performed on the Watson strand of normal DNA and c^{7}G-DNA. All reactions were performed at 37 °C. The concentrations of echinomycin (μM) included in the reactions are indicated by numbers above each pair of lanes. The left and right lanes in each pair correspond to reaction times of 1 and 5 min, respectively. Lanes marked with a "G" are the Maxam–Gilbert DMS/piperidine tracks specific for guanosine performed on normal DNA. Bond numbers are indicated along the left side of each autoradiogram.

The autoradiograms which resulted from these saturation experiments are shown in Figure 7. Several sites of strong footprinting are clearly visible, for example, at bonds 15–25, 54–62, 71–80, and 94–99. Also clearly evident are regions of enhanced cleavage bordering many of these sites. The strong effects are evident in both normal and c^{7}G-DNA; indeed, there appear to be very few differences between the DNase I cleavage patterns seen with either of the two DNAs. A preliminary densitometric analysis of these data reinforced the conclusion that the effects on both DNAs were very similar (results not shown). It therefore appears that replacing guanosine with c^{7}G, thereby preventing the formation of Hoogsteen GC pairs, has practically no effect on the binding of echinomycin at saturating concentrations. Therefore, Hoogsteen GC pairs cannot be required in order for echinomycin to bind to DNA. It has previously been shown that Hoogsteen AT pairs are not required (McLean et al., 1989).

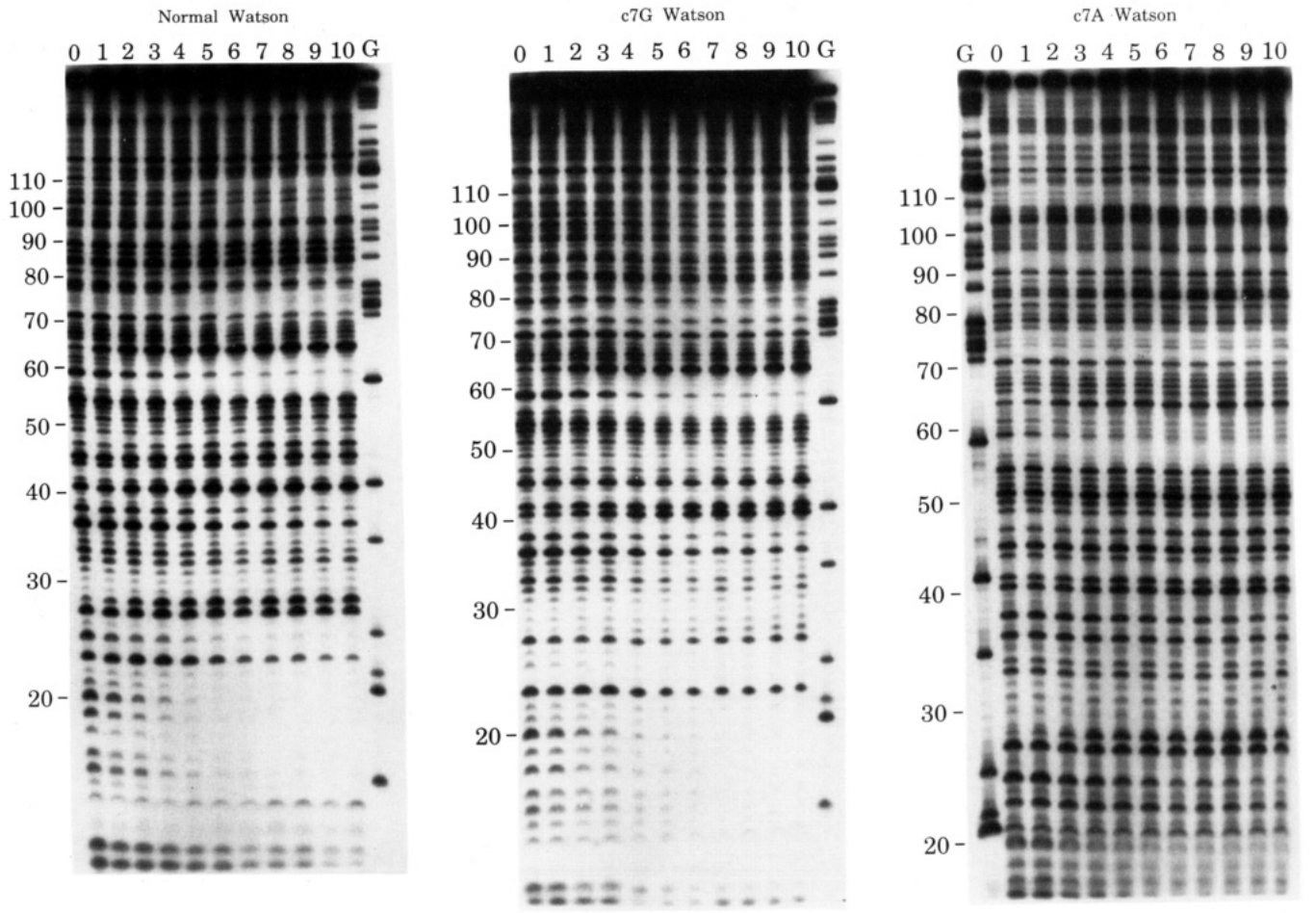

FIGURE 8: Results of DNase I footprinting titration experiments performed on the Watson strands of normal DNA and c^7G- and c^7A-DNA. All reactions were carried out for 1 min at 37 °C. The concentration of echinomycin (μM) included in each reaction is indicated by the number above each lane. Lanes marked with a "G" are the Maxam–Gilbert DMS/piperidine tracks specific for guanosine performed on normal DNA. Bond numbers are indicated along the left side of each autoradiogram.

Quantitative DNase I Footprinting Experiments. The next question to be addressed is, if Hoogsteen pairs are not mandatory, then exactly how important are they to the DNA-binding ability of echinomycin? The difference between this question and the previous one is that this question demands a quantitative answer. Therefore, in order to answer it properly, quantitative information concerning the DNA-binding ability of echinomycin must be obtained. Fortunately, the technique of DNase I footprinting has recently been improved so that a quantitative estimate of the strength of ligand binding can be obtained. In these experiments, increasing concentrations of ligand are employed so that the observed reductions in band intensity may be plotted as a function of ligand concentration. Using a computer minimization, Dabrowiak and co-workers have used such footprinting plots to determine microscopic binding constants for a given ligand (Rehfuss et al., 1990a; Goodisman & Dabrowiak, 1991). However, as will be discussed later, there are several pitfalls in this process which complicate the job of determining accurate K_a values. Since the present study is primarily concerned with detecting any differences in echinomycin binding behavior between normal and modified DNA, only relative estimates of K_a are required. Thus, the goal of the present titration experiments was to determine a quantity hereafter referred to as C_{50}, defined as the antibiotic concentration at which half of the maximal reduction in band intensity occurred. To answer the proposed question, it was first necessary to analyze thoroughly the binding behavior of echinomycin to normal DNA using this footprinting titration technique. After such an analysis was repeated for c^7G-DNA and c^7A-DNA, the resulting C_{50} values could then be compared.

Before these studies were initiated, an appropriate concentration range of echinomycin had to be found. Examination of previous footprinting data suggested that the majority of the C_{50} values lie in the range 2–5 μM (Low et al., 1984). Indeed, preliminary experiments using various concentration ranges revealed that the CpG steps were mostly saturated by 10 μM. Thus, the range 0–10 μM was determined to be the ideal range over which to observe the DNA-binding behavior of echinomycin.

Autoradiograms resulting from footprinting titration experiments with the Watson (upper) strand of normal and c^7G- or c^7A-substituted DNA are shown in Figure 8. The Watson strand was chosen for analysis since the majority (9 out of 12) of the CpG steps could be reasonably resolved on a gel when this strand was 5′-end-labeled. As is clearly obvious from this figure, all of the echinomycin binding sites (and also the regions of enhanced cleavage) show a marked dependence on the concentration of echinomycin. It is this concentration dependence that the footprinting titration experiment exploits to obtain quantitative binding data. The first step in the procedure involved converting the autoradiogram into a series of footprinting plots for each gel band. Such plots are simply graphs of band intensity vs ligand concentration. In this experiment, 94 bands were analyzed for each DNA lane. Figure 9 illustrates a selected group of such plots which are normalized in terms of R, the relative band intensity (control = 1), so that

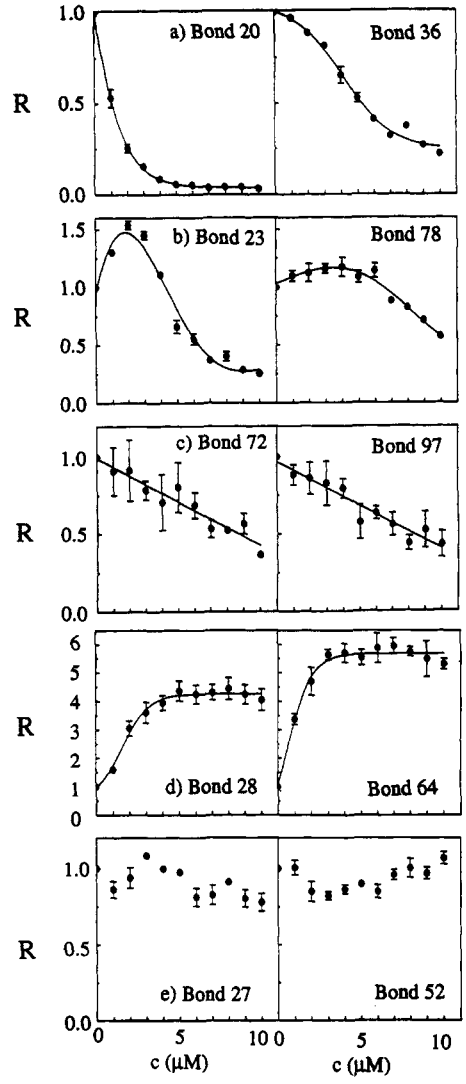

FIGURE 9: Footprinting plots for selected bonds in normal DNA which exemplify the types of response to echinomycin: (a) type I, (b) type II, (c) type III, (d) enhanced cleavage, and (e) no change. All plots display data collected at 37 °C. The quantity R is the relative band intensity as defined in the text, and c is the concentration of echinomycin. In all cases, the points shown are the mean R values of three scanning trials, and the error bars indicate the standard error of the mean. The curves shown were fitted by either a linear or a nonlinear least-squares algorithm as appropriate.

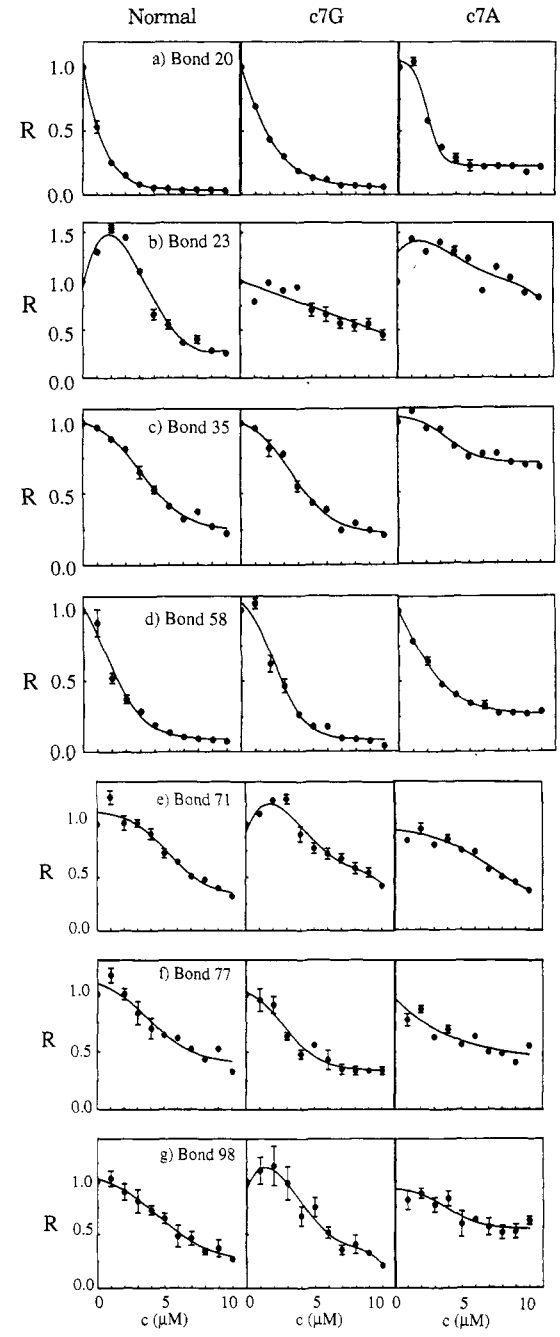

FIGURE 10: Footprinting plots showing the behavior of selected bonds in normal DNA and c7G- and c7A-DNA. All of these bonds are part of or lie very close to a CpG step. All plots display data collected at 37 °C. The details of the data presentation are the same as those used in Figure 9.

all plots have the same scale. Included with several of the plots are fitted curves from which a C_{50} value was calculated as follows.

The first kind of plot, type I, is depicted in Figure 9a and is characterized by a simple reduction in R. Among protected sites, type I plots were by far the most common, comprising 75% of such sites in normal DNA and 67% of such sites in both modified DNAs. In all cases, the curve which best fit these plots was the "logistic" curve, having the general form $R = Ac^n/(B + c^n)$, where A, B, and n are constants. Although equations of this form are generally used to fit sigmoid curves to data such as seen for bond 36, this equation yielded excellent fits to data such as seen for bond 20, which at first glance would appear to be best fit by a rectangular hyperbola. In fact, none of the cases such as bond 20 were well fit by a hyperbola. The most useful piece of data generated by this curve fitting was the quantity A, which is the extreme value reached by the curve at large c. Thus, A is the R_{min} required for determining C_{50} as described in Materials and Methods.

The second kind of plot, type II, is depicted in Figure 9b and is characterized by enhanced cleavage followed by protection. Among protected sites, such plots were relatively rare, amounting to 16% and 13% of protected sites in normal and modified DNAs, respectively. Type II plots probably indicate weak sites which are interacting with nearby strong sites, as will be discussed below. The curves used to fit such data were fourth-order polynomials and resulted in two basic subtypes of curves. The first, illustrated by bond 23, reached an identifiable end point during the titration, while the second, illustrated by bond 78, reached no such end point.

Type III plots, the third and final type of plot where protection was observed, are depicted in Figure 9c. Type III plots are actually a special case of type I plots, as they are also characterized by a simple reduction in R. The distinguishing

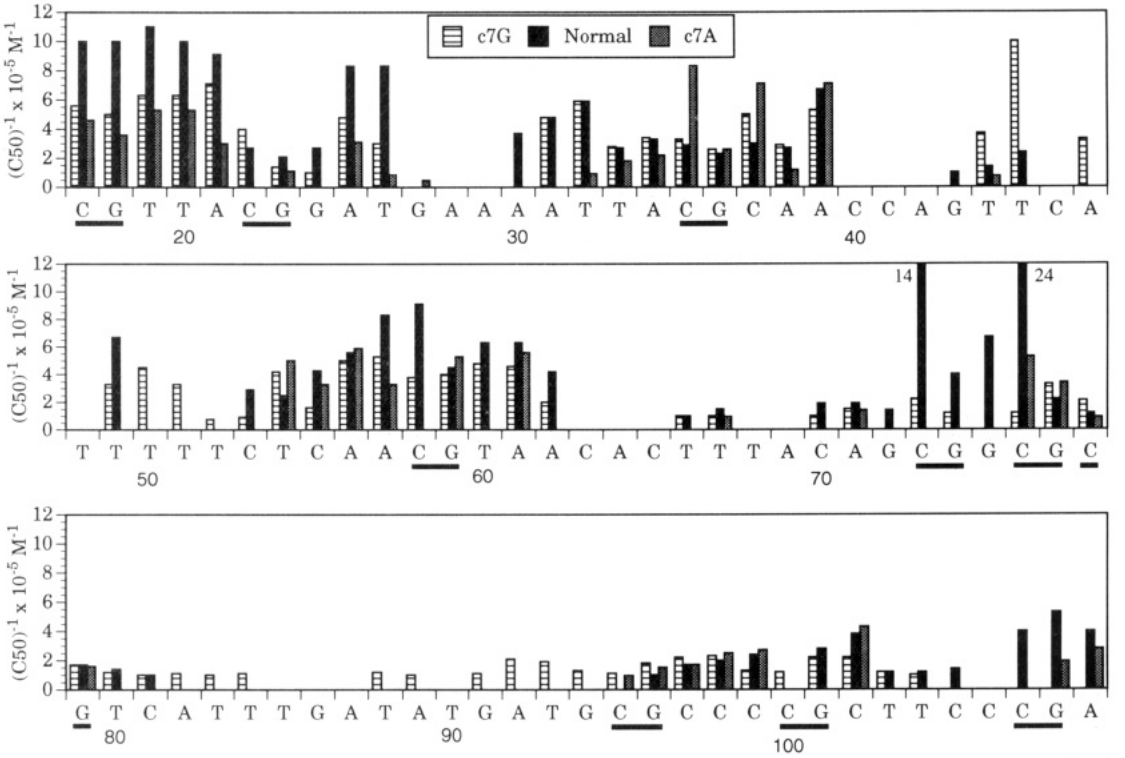

FIGURE 11: Calculated $1/C_{50}$ values for bonds in the Watson strands of normal DNA and c7G- and c7A-DNA showing protection from DNase I cleavage at 37 °C. Bond numbers are indicated below the printed sequence, and all CpG steps have been underscored.

features of type III plots are that the best-fit curve is simply a straight line and there is no discernible end point below 10 μM. Like type II plots, type III plots are relatively uncommon, comprising 8% and 20% of protected sites in normal and modified DNAs, respectively. These plots also indicate weak sites, but unlike type II plots, they do not appear to be affected by antibiotic binding nearby. It is, of course, certain that these curves would indeed reach an end point if the concentration range were extended to higher values, but the increase in accuracy would be quite small, with some risk of spurious and/or nonspecific perturbations of cleavage intervening. C_{50} values could only be estimated for these sites, as described in Materials and Methods.

The remaining two types of plots, depicted in panels d and e of Figure 9, show behavior other than protection. Those in Figure 9d represent the well-documented regions of enhanced cleavage which occur immediately adjacent to echinomycin binding sites. Indeed, both of the plots shown correspond to sites immediately 3′ to a region of protection. Of the 94 bonds analyzed for normal DNA, 25 showed this type of behavior. Finally, the plots shown in Figure 9e are examples of sites where the rate of DNase I cleavage was not affected by echinomycin. Eight such sites were observed in normal DNA.

Footprinting plots with the modified DNAs are shown in Figure 10, alongside those measured with normal DNA for comparison. Again, a variety of behavior was observed, but some consistent patterns can be discerned. First of all, protection clearly occurred with all DNAs at each site. It also appears that sites which either are very strong (such as bond 20) or are isolated from other sites (such as bonds 35 and 58) gave rise to generally similar plots on the three DNAs. Larger differences appear between the normal and modified DNAs at sites which are close to another site (such as bonds 23, 71, 77, and 98). It also appears that the minima reached by curves for c^7A-DNA tend to be slightly higher than those for the other DNAs. Little more can be said from these graphs, since they represent only one bond in a binding site containing at least three bonds. To construct a more complete picture, the data must be analyzed in terms of C_{50} values for each individual bond.

In Figure 11 the full data sets are plotted in terms of reciprocal C_{50} values, so that larger bars indicate stronger binding. Bonds for which no bar is shown either were not protected or showed enhanced cleavage. It is obvious that major areas of protection centered around the CpG steps. In addition, there are some sporadic sites of protection, such as at bonds 43–45, 49–51, and 66–67. It is interesting to note that all of these sites lie very near CpA steps, which recently were found to provide weak binding sites for echinomycin (Phillips et al., 1990). However, it must be said that in most cases the CpA bond itself was not protected. Comparing the three DNAs shows that protection was consistently observed at the CpG sites. Moreover, the $1/C_{50}$ values calculated for each type of DNA seem to be very similar when a protected region is considered as a whole, even though there are some specific bonds where the calculated $1/C_{50}$ values for the three DNAs are considerably different. It should be pointed out that in several cases such discrepancies were clearly caused by inherent limitations of this footprinting method. For example, in regions of DNA that were cleaved poorly by DNase I (such as bonds 71–75) the integrated band areas were inevitably low, giving a low signal-to-noise ratio and resulting in a high degree of scatter in the data. Thus, any reduction in signal due to protection would be partially (or even totally) obscured by noise. This effect gave rise to the sporadic nature of the data for bonds 71–75, as well as to "holes" such as at bonds 58 and 60 in c^7A-DNA, where no protection could be calculated even though strong protection was observed at all surrounding bonds which were better cleaved by DNase I. Finally, the most salient results of the titration experiments are summarized in Figure 12, where each protected region is presented separately in an expanded form, together with error estimates and an indication of the type of curve measured for each bond.

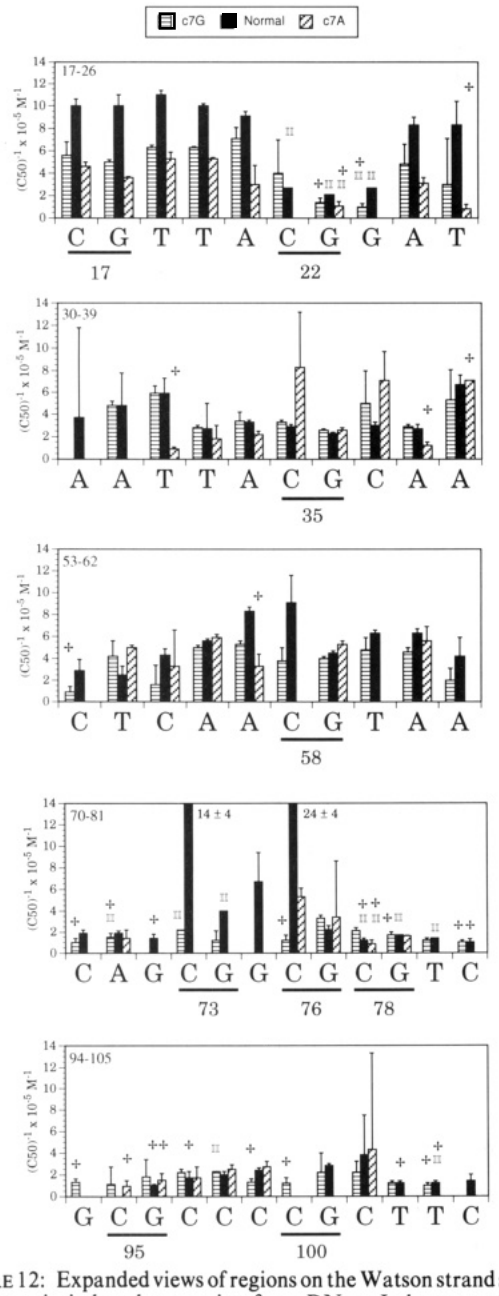

FIGURE 12: Expanded views of regions on the Watson strand showing echinomycin-induced protection from DNase I cleavage at 37 °C. All CpG steps have been underlined, and the CpG bond number indicated. Error values were calculated as described in Materials and Methods. Bonds which displayed type II behavior are indicated by a II, while those which displayed type III behavior are indicated by a +. Bonds labeled with both symbols displayed type II behavior but did not reach an identifiable end point during the titration. Hence, the bars plotted for all bonds labeled with a + are only estimates of the true values. All other bonds not labeled by either symbol displayed type I behavior.

DISCUSSION

Characteristics of the Binding of Echinomycin to Different Types of Sites in Normal DNA. Before any meaningful comparisons can be made, it is necessary to characterize fully the DNA-binding behavior of echinomycin as revealed by the footprinting titration experiments. This behavior is indicated by both the plot type and the $1/C_{50}$ observed for each DNA bond. From these data, a full picture of both the sequence specificity of the antibiotic and the relative strengths of the various binding sites can be obtained.

To begin with, the data in Figure 12 show how the three plot types are distributed. It appears that type I curves represent typical protection behavior, while types II and III seem to be exceptions. Type III behavior is generally associated with weak sites, as expected since the putative end points of these plots lie beyond 10 μM echinomycin. As for type II behavior, without exception such plots resulted where two or more CpG steps lie close together. On the *tyrT* fragment, there are three such areas: bonds 17–23, 73–79, and 95–108. Thus, these plots most likely indicate the existence of interacting binding sites where cleavage is initially enhanced due to binding at a nearby site and which become filled with bound antibiotic only at higher concentrations.

Next we turn to the problem of extracting pertinent constants out of the footprinting data. One of the great advantages of the footprinting titration experiment is the very detailed view of binding behavior which it reveals. The strength of the method lies not only in its ability to describe the behavior of echinomycin with respect to each bond but also in an improved method of data normalization. Because in previous footprinting protocols (such as that used for the present saturation studies) the data were normalized by forcing the summed band area to be the same for each lane (even though the summed band area of drug lanes should be less due to footprinting), the intensity of footprints was reduced and bands whose rate of cleavage actually remained constant were portrayed as enhancements. This problem can be rectified by the total cut plot (see Materials and Methods), which will accommodate any dependence of the summed band area on ligand concentration. The problem can also be avoided by normalizing the data with respect to the density of bands unaffected by ligand binding, as practiced by the Ackers group (Brenowitz et al., 1986) and by Chaires et al. (1990), provided that such "control" bands can be identified with confidence. Another significant advantage of the titration method is its ability to determine the relative strength of interaction with different binding sites. This is due to the fact that the method can determine both C_{50} and R_{min} values, unlike previous methods which can only determine R_{min} values, and then only if saturating conditions are employed. As discussed below, R_{min} is wholly unrelated to the thermodynamics of drug binding, but rather describes the ability of DNase I to cleave the drug–DNA complex (Goodisman & Dabrowiak, 1985, 1992). This point may be illustrated by reference to the behavior around bond 20 at the bottom of the autoradiograms in Figure 8. The data for c^7A-DNA look different from the results obtained with the other two DNAs, chiefly in that the intensity of band 20 at high echinomycin concentrations is greater than observed with the normal and c^7G-DNA species. The quantitated data (Figure 10) reflect this difference: R_{min} for c^7A-DNA (0.2) is larger than for normal and c^7G-DNA (0.05). However, the C_{50} value for c^7A-DNA is very similar to the C_{50} values for both normal and c^7G-DNA, and thus the apparent binding affinity of echinomycin for each of the DNA species is much the same, *regardless* of the different R_{min} values. It is the C_{50} value which holds the thermodynamic data, and the important question is how such information may best be harvested.

A great deal of work has been done by Dabrowiak and co-workers on various mathematical models of footprinting behavior, with the ultimate goal of determining accurate microscopic affinity constants. The first and simplest model, proposed for a single, noninteracting site, is summarized by the following expression:

$$\nu = \frac{I - I_o}{I_\infty - I_o} = \frac{Kc}{1 + Kc} \qquad (2)$$

where ν is the fractional occupancy of the site, I is the band area for a free ligand concentration c, I_0 is the area for the control, I_∞ is the area at saturating ligand concentration, and K is the binding constant. This expression defines a simple binding isotherm, and K can be determined from a plot of ν *vs* $-\log c$ by finding the point on the $-\log c$ axis where $\nu = 0.5$. This $-\log c$ would then be equal to log K (Goodisman & Dabrowiak, 1985, 1992; Rehfuss et al., 1990b). In the present case where $R = I/I_0$ is plotted *vs* c (yielding plots with a vertical range of R_{min} to 1 rather than of 0 to 1), the equivalent procedure would be to determine the value of c where R reached half of its total change. This C_{50} value should be $1/K$. However, there are several considerations which make this a dangerous assumption. First, the value of c in eq 2 is the *free* ligand concentration, which is not known in the footprinting experiment. Only the total ligand concentration is known, and this discrepancy can lead to underestimation of K by 1 or 2 orders of magnitude (Rehfuss et al., 1990b; Chaires et al., 1990). However, it can be calculated that in the present experiments, if one assumes that an echinomycin molecule is bound at each of the 12 CpG sites on every *tyrT* DNA molecule, then 96% of the added echinomycin would remain unbound even at the lowest concentration used (1 μM). Thus, it would seem that in the present case the approximation of free antibiotic concentration to total echinomycin concentration is a reasonable one. The second problem is that the simple model is valid only for independent, noninteracting sites. Even the most sophisticated current multisite models are limited by this assumption. However, there are several pieces of evidence which suggest that this assumption is not valid for echinomycin. The very existence of type II behavior is almost certainly a reflection of site–site interactions. Moreover, type I curves did not fit a simple binding isotherm such as eq 2, but rather the logistic equation, a classic indication of cooperative binding. We note that for a truly independent and noninteracting antibiotic binding site, the binding isotherm would require approximately 1.8 logarithmic units in free titrant concentration to span 0.1–0.9 fractional site occupancy. This means that titrant concentration must vary over at least a 100-fold range to saturate such a site. In our curves (Figures 9 and 10) saturation is completed in a much smaller span. Given that the site concentration in our experiments is in the picomolar range, well below the reciprocal of the anticipated binding constants, the likelihood of positive cooperativity cannot be overlooked. Indeed, in recent NMR studies using the duplex d(ACGTACGT), the binding of echinomycin to the two CpG sites was found to be cooperative (Gilbert & Feigon, 1991). Even though these observations do not prove the existence of cooperative binding in full-length DNA, they do suggest that it is a very real possibility. Consequently, the reciprocal C_{50} values presented for type I plots should not be considered as an accurate measure of the true K values, but only estimates (and probably underestimates). That said, the majority of the binding constants determined for echinomycin with several naturally occurring DNAs fall within the range $1–5 \times 10^5$ M^{-1} (Wakelin & Waring, 1976), which is exactly the $1/C_{50}$ range observed for the weaker CpG steps. Thus, the $1/C_{50}$ values obtained may not be much lower than the true values. In any case, they can certainly be considered as relative values, and that is sufficient for the present purposes of comparing echinomycin binding to different DNAs. It should, however, be emphasized that the values given for type II and III plots are less reliable than those for type I plots.

A broad conclusion which can be drawn from the data in Figures 11 and 12 is that protection generally extends 3 base pairs to the 3′ side of a CpG site and 5 base pairs to the 5′ side. This is especially true of the isolated sites at bonds 35 and 58, although it is observed for all CpG steps with the exception of the very weakly protected site at bond 95. As seen from Figures 7 and 8, protection does extend 5′ to bond 17, but this information was not analyzed since that region was derived from PCR primer and thus contained modified purines only in one strand. Happily, this protection pattern is exactly that predicted from the model based on the crystal structure of DNase I–DNA complexes (Suck & Oefner, 1986). This model greatly facilitates the interpretation of complex regions such as the CpG cluster at bonds 73–79. Because of the close spacing of these sites, it is highly likely that only sites 73 and 78 can be bound simultaneously by separate echinomycin molecules. Therefore, the observed behavior is most likely an average of the behavior of DNA species having an echinomycin molecule bound at one of the three sites and possibly at both bonds 73 and 78. This might explain the apparently stronger binding at the middle site (bond 76), which would be protected by binding at any of the three sites. The fact that the data fit the DNase I structural model so well strongly suggests that CpG is indeed the favored site of echinomycin. Moreover, as is apparent from Table I, which ranks the sites in Figure 12 on the basis of both $1/C_{50}$ values and plot type, the range of binding strength of the various CpG steps seems to be no greater than 1 order of magnitude. Similar results were obtained recently by assaying echinomycin-induced transcription blockade *in vitro* (Phillips et al., 1990). These studies revealed that the relative occupancy of echinomycin varied by less than 1 order of magnitude between eight 5′-NCGN-3′ binding sites. So while echinomycin slightly prefers 5′-ACGN-3′ and 5′-NCGT-3′ sequences, the data suggest that any CpG serves as a fairly good binding site for the antibiotic. These results are in agreement with recent NMR studies which demonstrated that echinomycin binds more tightly to 5′-ACGT-3′ than to 5′-TCGA-3′ (Gilbert & Feigon, 1991).

Table I: Ranking of CpG Binding Sites[a]

site	CpG position	$1/C_{50}$ (range)	plot type
CCGT	17	10 (10–10)	I
ACGT	58	7.3 (4.5–9.1)	I
ACGC	35	2.8 (2.3–3.3)	I
GCGC	76[b]	11.1 (2.2–24)	I
CCGC	100	1.7 (0.0–2.8)	I
GCGG	73[b]	6.5 (1.4–14)	I, II
ACGG	22	4.6 (2.1–9.1)	II
GCGT	78[b]	1.7 (1.2–2.2)	II
GCGC	95	0.3 (0.0–1.0)	III

[a] At each site the cutting of the three internucleotide bonds in "normal" DNA is considered, and the respective $1/C_{50}$ values are averaged. The *range* of the three values, giving some idea of the consistency of the effect spread over all three bonds, is shown in parentheses. For site 17, data were only available for the CpG and GpT bonds (Figure 12). [b] Data at this site are probably composite values of the CpG "cluster" at this region.

Importance of Hoogsteen Pairs to the DNA-Binding Ability of Echinomycin. The overall strategy in the present work centered around the incorporation of 7-deazapurines into DNA in an attempt to create DNA which was incapable of forming Hoogsteen pairs. But does the incorporation of these analogs into DNA actually achieve this goal? The formation of a Hoogsteen pair consists of two steps: first, the purine must adopt a *syn* conformation and second, a hydrogen bond must form between the purine N7 and the N3-H moiety of the pyrimidine. Clearly, the incorporation of 7-deazapurines would prevent the second step from occurring, but not the first. There is nothing to prevent a 7-deazapurine from

adopting a *syn* conformation, especially as this allows the sugar–phosphate backbone of DNA to fold more closely around echinomycin, thereby further stabilizing the complex (Quigley et al., 1986). However, if a 7-deazapurine were to adopt a *syn* conformation, it could not be stabilized in this conformation by the aforementioned hydrogen bond, as could a normal purine. Presumably, this lack of further stabilization would impart an energy penalty to the echinomycin–c^7-DNA complex. The question thus arises as to which event, the closer approach of the sugar–phosphate backbone to echinomycin or the formation of the hydrogen bond to N7, endows the complex with the greater energy benefit. The present results bear directly on the significance of a hydrogen bond to N7 to the stability of echinomycin–DNA complexes. If this hydrogen bond is critical, then its absence in echinomycin–c^7-DNA complexes should significantly reduce the stability of these complexes, causing a corresponding reduction in $1/C_{50}$. In light of this argument the importance of Hoogsteen pairs to the DNA-binding ability of echinomycin can be deduced from a simple comparison between the $1/C_{50}$ values calculated for normal DNA and the two modified DNAs.

As discussed in the introduction, Hoogsteen pairs have been observed in ACGT and GCGC sequences, and then only at low pH for the latter case. Thus, any differences between the $1/C_{50}$ values for normal and modified DNAs would be most likely to occur at corresponding sequences in the *tyrT* fragment. These exact sequences are found at bonds 58, 76, and 95, while partial matches are found at bonds 17, 22, 35, 73, 78, and 100. Considering the $1/C_{50}$ values shown in Figure 12, it is evident that in the majority of cases the differences between the DNAs are reduced to statistical insignificance by the inclusion of error bars (also bear in mind that, as explained before, these error estimates are probably underestimates of the true error). After careful examination of the data, the only regions which show consistent differences between normal and modified DNAs are those around bonds 17 and 22. Around bond 17, the $1/C_{50}$ values for normal DNA are somewhat larger than those for both modified DNAs, while this is the case for only c^7A-DNA around bond 22. Attempts to discern any unique sequence characteristics at these two sites failed. Thus, preventing the formation of the N7 hydrogen bond at sites most likely to form Hoogsteen pairs did not significantly affect the DNA-binding behavior of echinomycin. Moreover, the only differences observed were modest reductions in $1/C_{50}$ values at two nonprime sites.

Thus, the results strongly suggest that the elimination of the N7 hydrogen bond has a very minimal effect, if indeed any, on the ability of echinomycin to bind to DNA. Therefore, the contribution of stable, intact Hoogsteen pairs *per se* to the stability of echinomycin–DNA complexes appears to be negligible. The same result was observed at 1 °C with c^7G-DNA (data not shown), and thus the importance of Hoogsteen pairs appears to be independent of temperature, although some interesting differences were observed in the DNA-binding behavior of echinomycin at low temperature. Whether echinomycin causes purine nucleotides adjacent to its binding site to adopt a *syn* conformation cannot be rigorously disproved, but the data show that such a conformational change cannot be an end unto itself. If it were, the removal of the N7, which certainly destabilizes this conformation, should have caused a more significant alteration in the DNA-binding behavior of echinomycin. Rather, any perturbation of *syn–anti* equilibrium is likely to be a mere consequence of some other echinomycin-induced conformational change under permissive conditions, which is most likely the movement of the sugar–phosphate backbone toward the antibiotic. This assertion is supported by the NMR evidence, which indicates that Hoogsteen pairs can be eliminated simply by increasing the temperature to 37 °C or by adding additional base pairs to the end of an oligonucleotide. For all intents and purposes, these same changes did not appear to disturb echinomycin in the least (Gao & Patel, 1989; Gilbert & Feigon, 1991). The discussion thus returns to the theme expressed at the beginning, which is the necessity for DNA to undergo profound structural deformation in order to accommodate the echinomycin molecule. Therefore, it would appear that Hoogsteen pairs in echinomycin–DNA complexes, if they do exist, are merely byproducts of the process by which DNA, confronted with such a highly invasive ligand, attempts to find the nearest minimum on a freshly altered free energy surface.

ACKNOWLEDGMENT

We thank Dean Gentle for excellent technical assistance and Dr. M. McLean for synthesizing the PCR primers.

REFERENCES

Bloch, W. (1991) *Biochemistry 30*, 2735–2747.
Brenowitz, M., Senear, D. F., Shea, M. A., & Ackers, G. (1986) *Methods Enzymol. 30*, 132–181.
Chaires, J. B., Herrera, J. E., & Waring, M. J. (1990) *Biochemistry 29*, 6145–6153.
Corbaz, R., Ettlinger, L., Gaumann, E., Keller-Schierlein, W., Kradolfer, F., Neipp, L., Prelog, V., Reusser, P., & Zahner, H. (1957) *Helv. Chim. Acta. 40*, 199–204.
Drew, H. R., & Travers, A. A. (1984) *Cell 37*, 491–502.
Foster, B. J., Clagett-Carr, K., Shoemaker, D. D., Suffness, M., Plowman, J., Trissel, L. A., Grieshaber, C. K., & Leyland-Jones, B. (1986) *Invest. New Drugs 3*, 403–410.
Gao, X., & Patel, D. J. (1988) *Biochemistry 27*, 1744–1751.
Gao, X., & Patel, D. J. (1989) *Q. Rev. Biophys. 22*, 93–138.
Gauvreau, D., & Waring, M. J. (1984) *Can. J. Microbiol. 30*, 439–450, 730–738.
Gilbert, D. E., & Feigon, J. (1991) *Biochemistry 30*, 2483–2494.
Gilbert, D. E., van der Marel, G. A., van Boom, J. H., & Feigon, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3006–3010.
Goodisman, J., & Dabrowiak, J. C. (1985) *J. Biomol. Struct. Dyn. 2*, 967–979.
Goodisman, J., & Dabrowiak, J. C. (1991) in *Advances in DNA Sequence Specific Agents* (Hurley, L. H., Ed.) American Chemical Society, Washington, DC.
Goodisman, J., & Dabrowiak, J. C. (1992) *Biochemistry 31*, 1058–1064.
Jeppesen, C., & Nielsen, P. E. (1988) *FEBS Lett. 231*, 172–176.
Katagiri, K., Yoshida, T., & Sato, K. (1975) In *Antibiotics III: Mechanism of Action of Antimicrobial and Antitumor Agents* (Corcoran, J. W., & Hahn, F. E., Eds.) pp 234–251, Springer-Verlag, Berlin.
Low, C. M. L., Drew, H. R., & Waring, M. J. (1984) *Nucleic Acids Res. 12*, 4865–4879.
Lunt, E. (1979) in *Comprehensive Organic Chemistry* (Sammes, P. G., Ed.) Vol. 4, pp. 493–563, Pergamon Press, Oxford.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
McConlogue, L., Brow, A. D., & Innis, M. A. (1988) *Nucleic Acids Res. 16*, 9869.
McLean, M. J., & Waring, M. J. (1988) *J. Mol. Recognit. 1*, 138–151.
McLean, M. J., Seela, F., & Waring, M. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9687–9691.
McMurry, J. (1984) *Organic Chemistry*, pp 1123–1131, Brooks/Cole Publishing Co., Monterey, CA.
Mendel, D., & Dervan, P. B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 910–914.

Mizusawa, S., Nishimura, S., & Seela, F. (1986) *Nucleic Acids Res. 14*, 1319–1324.
Muss, H. B., Blessing, J. A., & Malfetano, J. (1990a) *Am. J. Clin. Oncol. 13*, 191–193.
Muss, H. B., Blessing J. A., Baker, V. V., Barnhill, D. R., & Adelson, M. D. (1990b) *Am. J. Clin. Oncol. 13*, 299–301.
Phillips, D. R., White, R. J., Dean, D., & Crothers, D. M. (1990) *Biochemistry 29*, 4812–4819.
Portugal, J., Fox, K. R., McLean, M. J., Richenberg, J. L., & Waring, M. J. (1988) *Nucleic Acids Res. 16*, 3655–3670.
Quigley, G. J., Ughetto, G., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., & Rich, A. (1986) *Science 232*, 1255–1258.
Rehfuss, R., Goodisman, J., & Dabrowiak, J. C. (1990a) in *Molecular Basis of Specificity in Nucleic Acid–Drug Interactions* (Pullman, B., & Jortner, J., Eds.) pp 157–166, Kluwer Academic Publishers, Dordrecht.
Rehfuss, R., Goodisman, J., & Dabrowiak, J. C. (1990b) *Biochemistry 29*, 777–781.
Sato, K., Shiratori, O., & Katagiri, K. (1967) *J. Antibiot., Ser. A 20*, 270–276.
Seela, F., & Röling, A. (1991) *Nucleosides Nucleotides 10*, 715–717.
Seela, F., & Röling, A. (1992) *Nucleic Acids Res. 20*, 55–61.
Suck, D., & Oefner, C. (1986) *Nature 321*, 620–625.
Ughetto, G., Wang, A. H., Quigley, G. J., van der Marel, G. A., 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.
Wakelin, L. P. G., & Waring, M. J. (1976) *Biochem. J. 157*, 721–740.
Wang, A. H., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science 225*, 1115–1121.
Wang, A. H., Ughetto, G., Quigley, G. J., & Rich, A. (1986) *J. Biomol. Struct. Dyn. 4*, 319–342.
Ward, B., Rehfuss, R., & Dabrowiak, J. C. (1987) *J. Biomol. Struct. Dyn. 4*, 685–695.
Ward, D., Reich, E., & Goldberg, I. H. (1965) *Science 149*, 1259–1263.
Waring, M. J. (1979) in *Antibiotics V, Part 2: Mechanism of Action of Antieukaryotic and Antiviral Compounds* (Hahn, F. E., Ed.) pp 173–194, Springer-Verlag, Heidelberg.
Waring, M. J. (1990) in *Molecular Basis of Specificity in Nucleic Acid–Drug Interactions* (Pullman, B., & Jortner, J., Eds.) pp 225–245, Kluwer Academic Publishers, Dordrecht.
Waring, M. J., & Wakelin, L. P. G. (1974) *Nature 252*, 653–657.