Sampson, J. R., & Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033-1037.
Sierzputowska-Gracz, H., Sochacka, E., Malkiewicz, A., Kuo, K., Gehrke, C. W., & Agris, P. F. (1987) *J. Am. Chem. Soc. 109*, 7171-7177.
Singh, H., Herbut, M. H., Kee, C.-H., & Sarma, R. H. (1976) *Biopolymers 15*, 2167-2184.
Sprinzl, M., Hartmann, T., Weber, J., Blank, J., & Zeidler, R. (1989) *Nucleic Acids Res.* (Sequences Suppl.) *17*, rl-r67.
Sundaralingam, M., Mizuno, H., Stout, C. D., Rao, S. T., Liebman, M., & Yathindra, N. (1976) *Nucleic Acids Res. 3*, 2471-2484.
Watanabe, K., Oshima, T., Saneyoshi, M., & Nishimura, S. (1974) *FEBS Lett. 43*, 59-63.
Watanabe, K., Shinma, M., Oshima, T., & Nishimura, S. (1976) *Biochem. Biophys. Res. Commun. 72*, 1137-1144.
Watanabe, K., Yokoyama, S., Hansske, F., Kasai, H., & Miyazawa, T. (1979) *Biochem. Biophys. Res. Commun. 91*, 671-677.
Yamamoto, Y., Yokoyama, S., Miyazawa, T., Watanabe, K., & Higuchi, S. (1983) *FEBS Lett. 157*, 95-99.
Yokoyama, S., Yamaizumi, Z., Nishimura, S., & Miyazawa, T. (1979) *Nucleic Acids Res. 6*, 2611-2626.
Yokoyama, S., Inagaki, F., & Miyazawa, T. (1981) *Biochemistry 20*, 2981-2988.
Yokoyama, S., Watanabe, T., Murao, K., Ishikura, H., Yamaizumi, Z., Nishimura, S., & Miyazawa, T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4905-4909.
Zmudzka, B., & Shugar, D. (1970) *FEBS Lett. 8*, 52-54.

Site-Specific Binding Constants for Actinomycin D on DNA Determined from Footprinting Studies†

Jerry Goodisman,* Robert Rehfuss, Brian Ward, and James C. Dabrowiak*
Department of Chemistry, Center for Science and Technology, Syracuse University, Syracuse, New York 13244-4100
Received March 28, 1991; Revised Manuscript Received September 17, 1991

ABSTRACT: We report site-specific binding constants for the intercalating anticancer drug actinomycin D (Act-D), binding to a 139-base-pair restriction fragment from pBR 322 DNA. The binding constants are derived from analysis of footprinting experiments, in which the radiolabeled 139-mer is cleaved using DNase I, the cleavage products undergo gel electrophoresis, and, from the gel autoradiogram, spot intensities, proportional to amounts of cleaved fragments, are measured. A bound drug prevents DNase I from cleaving at ~7 bonds, leading to decreased amounts of corresponding fragments. With the radiolabel on the 3′ end of the noncoding strand (A-label), we measured relative amounts of 54 cleavage products at 25 Act-D concentrations. For cleavage of the 139-mer with the label on the 3′ end of the coding strand (G-label), relative amounts of 43 cleavage products at 11 Act-D concentrations were measured. These measurements give information about ~120 base pairs of the restriction fragment (~12 turns of the DNA helix); in this region, 14 strong and weak Act-D binding sites were identified. The model used to interpret the footprinting plots is derived in detail. Binding constants for 14 sites on the fragment are obtained simultaneously. It is important to take into account the effect of drug binding at its various sites on the local concentration of probe elsewhere. It is also necessary to include in the model weak as well as strong Act-D sites on the carrier DNA which is present, since the carrier DNA controls the free-drug concentration. As expected, the strongest sites are those with the sequence (all sequences are 5′ → 3′) GC, with TGCT having the highest binding constant, 6.4×10^6 M^{-1}. Sites having the sequence GC preceded by G are weak binding sites, having binding constants approximately 1 order of magnitude lower than those of the strong sites. Also, the non-GC-containing sequences CCG and CCC bind Act-D with a binding constant comparable to those of the weak GGC sites. The analysis may reveal drug-induced structural changes on the DNA, which are discussed in terms of the mechanism of Act-D binding.

Actinomycin D, Act-D (Figure 1), is one of the most intensely studied anticancer drugs (Gale et al., 1981). Numerous investigations have shown that the agent exhibits its antitumor effects by binding to double-stranded DNA, thereby blocking transcription. The binding mechanism involves intercalation of the phenoxazone ring system of the drug via the minor groove of DNA at GC-rich sites (Gale et al., 1981).

Footprinting studies on Act-D using DNase I (Lane et al., 1983; Scamrov & Beabealashvilli, 1983; Fox & Waring, 1984) and Fe–MPE (van Dyke et al., 1982) showed that the highest affinity sites have the dinucleotide sequence 5′-GC-3′. Model building and a single-crystal X-ray structural analysis (Sobell 1973) indicated that the specificity of Act-D for this sequence is due to hydrogen bonding between the 2-amino group of guanine of DNA and the threonine moiety located in the cyclic peptide of the drug.

Although 5′-GC-3′ is the preferred binding site, there are other sites. The duplex $d(CGTCGACG)_2$, which does not contain 5′-GC-3′, is able to strongly bind two actinomycin D molecules in a highly cooperative manner (Snyder et al., 1989). Further, in an effort to measure Act-D binding specificity, Rill et al. (1989) examined the DNA cleavage sites of the photoaffinity probe 7-azidoactinomycin D on several DNA re-

† We acknowledge the American Cancer Society, Grant NP-681, for supporting this research.

0006-2960/92/0431-1046$03.00/0

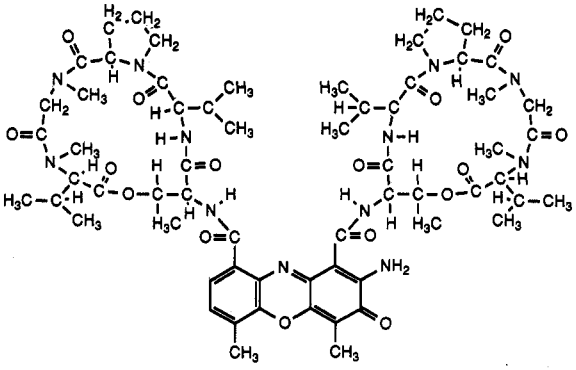

FIGURE 1: Structure of the drug actinomycin D.

striction fragments and found that GC doublets were strongly preferred only if the 5′-flanking base was a pyrimidine and the 3′-flanking base was *not* cytosine. In addition, the central GG doublet in the sequence TGGG is a strong Act-D site.

We have shown that binding constants for drugs on random sequence DNA can be obtained by quantitative analysis of footprinting data (Dabrowiak & Goodisman, 1989; Rehfuss et al., 1990a,b; Dabrowiak et al., 1991). In the footprinting experiment, we measure the amounts of the various fragments produced by cleavage of a DNA polymer by a cleavage agent, the "probe", in the presence of various amounts of the drug. The amount of fragment of a particular length is proportional to the amount of cleavage at a particular position on DNA. A plot of this quantity as a function of the amount of drug present, called the "footprinting plot", reflects the ability of the drug molecule to inhibit (or, sometimes, enhance) cleavage by the probe. To analyze these "footprinting plots", one must take into consideration overlapped drug sites, cleavage enhancements, and the way in which the drug affects cutting by the probe. The analysis makes it possible to measure binding constants for overlapping sites and, if measurements are made at different temperatures, to obtain a full thermodynamic profile of a drug: $\Delta G°$, $\Delta H°$, and $\Delta S°$ (Dabrowiak et al., 1990).

The quantitative footprinting method is the only method capable of giving *simultaneous* values for a number of binding constants, corresponding to binding to different sites, and it is a valuable tool in investigating sequence specificity, selectivity, and cooperativity in drug binding to DNA. Knowledge of these factors is important in the understanding of the mechanism of action of drugs and is crucial to drug design.

In a previous report, we measured the binding constant of Act-D toward a small oligonucleotide duplex containing the sequence GCGC using DNase I footprinting (Rehfuss et al., 1990a). In a preliminary study, we also analyzed the binding of actinomycin D to a 139-base-pair fragment of pBR 322 DNA (Figure 2), using DNase I, Fe–MPE, and the cationic porphyrin Mn–T4 as footprinting cleavage probes (Ward et al., 1988; Rehfuss et al., 1990b). Our analysis focused on the initial loading events on the 139-mer, for drug concentrations from 0 to 7.64 μM, the upper limit corresponding to roughly 50% occupation of the strong sites by drug. We derived equilibrium binding constants for the highest affinity sites and discussed the ability of different probes to reveal drug-induced structural changes on DNA.

In this report, we derive binding constants for both the high- and low-affinity sites. We must ascertain whether the simple model used so far is sufficient to explain the footprinting results obtained for the complete range of actinomycin D concentrations or whether additional assumptions (such as drug-induced structural effects) must be made. We show below that

FIGURE 2: The 139-base-pair fragment of pBR 322 DNA used in this study. Actinomycin D binding sites are indicated by rectangles, filled for strong sites and open for weak sites. Locations of the radiolabel: for the A-labeled experiment, position 33; for the G-labeled experiment, position 172.

both high- and low-affinity sites on the carrier DNA must be included in the model, even if one is interested in calculating binding constants only for the high-affinity sites on the fragment. In addition, by considering drug concentrations large enough to produce a high degree of drug loading on the DNA fragment, we have an opportunity to detect drug-induced structural changes in DNA, which have been previously reported to occur for Act-D (Fox & Waring, 1984; Huang et al., 1988; Bishop et al., 1991).

Materials and Methods

The *Hin*dIII/*Nci*I, 139-base-pair restriction fragment from pBR 322 DNA was isolated, purified, and end-labeled for one set of experiments at position 33(A), and for a second set of experiments it was labeled at 172(G) with reverse transcriptase as previously described (Lown et al., 1986; Bromley et al., 1986). Figure 2 shows the sequence of this fragment, as well as the drug-binding sites. The DNase I, obtained from Sigma Chemical Co., was used without purification.

The DNase I footprinting experiments involving the 139-mer and Act-D were carried out as previously described (Ward et al., 1988). The photograph of the autoradiogram along with the drug concentrations (from 0 to 38.8 μM) used in the A-label experiment was published previously (Ward et al., 1988). The experiments involving the G-labeled 139-mer were carried out under conditions identical to those of the A-labeled series except that only 10 drug concentrations were used, having final values in the reaction medium of 2.48, 3.40, 4.86, 6.93, 9.89, 14.1, 20.2, 28.8, 41.0, and 58.7 μM. The autoradiogram for the G-labeled experiment is shown in Figure 3. All experiments were carried out in buffer, containing 50 mM Tris-HCl, 8 mM $MgCl_2$, and 2 mM $CaCl_2$, at pH 7.5, for 10 min. The final concentration of DNA in each of the reactions was 194 μM base pairs: 193 μM base pairs sonicated calf thymus DNA and $\sim$1 μM base pairs radiolabeled fragment. Control reactions were carried out to determine the amount of enzyme necessary to leave $\sim$70% of the fragment uncleaved at the end of the digest time of 10 min (single-hit regime). The autoradiograms resulting from both the A- and G-labeled experiments were scanned with a microdensitometer/computer system to yield cross-sectional areas of bands, proportional to oligonucleotide concentrations (Dabrowiak et al., 1986). Plots of spot intensity as a function of drug concentration for each DNase I cleavage site (footprinting plots) were prepared in the manner previously described (Dabrowiak & Goodisman, 1989). This involved using the "total-cut plot", shown in Figure 4 for the A-label experiment, to correct the data for gel-loading errors and slightly different digest times.

Theory: Model for the System and Footprinting Plots

Binding of Act-D to the 139-mer. It is known that actinomycin D binds to DNA via the minor groove, mainly at GC-rich sites (Gale et al., 1981). Our footprinting analysis confirmed that the strongest binding sites on the polymer were

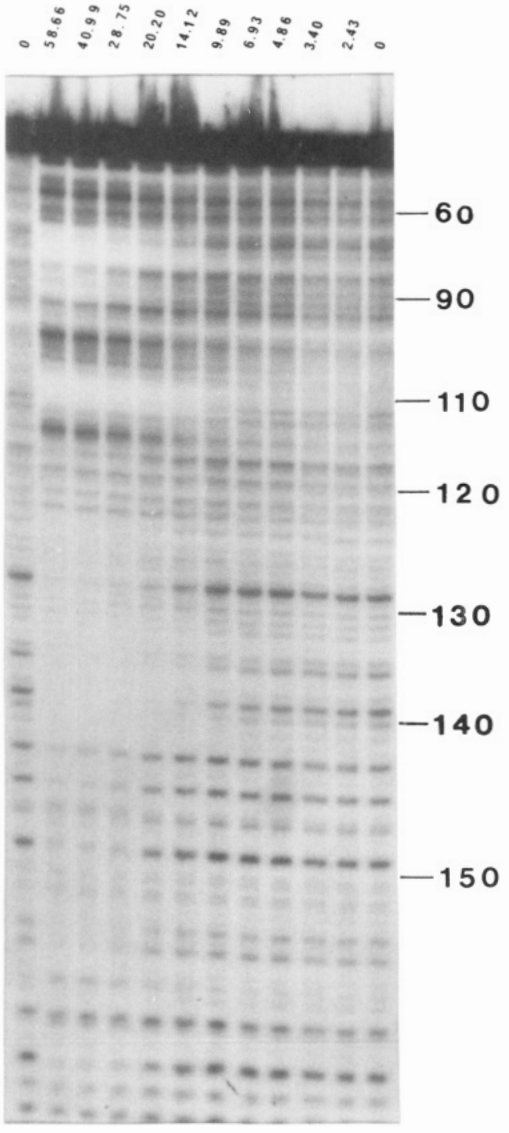

FIGURE 3: Autoradiogram for the G-labeled experiment. The Act-D micromolar concentrations are shown at the top of each lane.

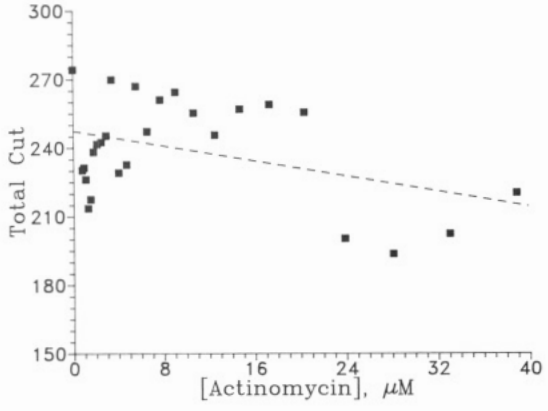

FIGURE 4: Total-cut plot for the A-labeled experiment. The sum of the spot intensities corresponding to 93% of all cleavage products is plotted as a function of total drug concentration. The broken line is a least-squares fit to a linear function; the ratio of this function to the measured total cut multiplies all spot intensities for that drug concentration to correct for gel loading and other errors.

5′-GC-3′, but the binding constant depended markedly on the bases on either side of the GC (Rehfuss et al., 1990b; Ward et al., 1988). In particular, sites having the sequence 5′-GGC-3′ were much weaker than other sites having the sequence 5′-GC-3′ (the 5′-GGC-3′ sites are considered among the weak sites in the present work). Furthermore, 5′-CCGT-3′, which does not contain the classical GC-binding sequence, appeared to be a binding site, with a binding constant comparable to that of the GGC sites.

The model used assumes that the drug bound to DNA blocks the probe from binding and cutting in a region near the position of the drug molecule. The actinomycin D molecule "covers" about 4 base pairs, including the G, the C, and one base pair on either side (Gale et al., 1981). Considering the size and shape of DNase I (Suck et al., 1988), we conclude that when an actinomycin molecule is bound at the GC site, DNase I cannot cleave at sites from about 3 base pairs on the 3′ side of the GC to about 2 base pairs on the 5′ side. Thus, the inhibition region for a single drug molecule is about 7 base pairs in size; footprinting plots for all sites in such a region should be identical, within a multiplying constant.

These footprinting plots show intensities monotonically decreasing with drug concentration. Footprinting plots for other sites show enhanced cutting, leveling off, and then further enhancement of cutting as drug concentration is increased; we assume that these are sites at which no drug binds. The explanation for the enhancement is given below. Other plots show initial increases of intensity with increased drug concentration and a subsequent intensity decrease. These are the weaker binding sites, for which drug binds at a sufficiently high drug concentration, causing a decrease in intensity which overcomes the enhancement. At drug concentrations high enough to cause binding at the weak binding sites on the 139-mer, weak-binding sites on the carrier DNA used in the experiment must also accept the drug. Thus, the analysis of footprinting data for higher drug concentrations requires modification in the model previously used for quantitative footprinting analysis of the initial loading events (strong binding sites), to include the activity of weak sites on the carrier DNA (Rehfuss et al., 1990b).

The equilibrium between drug bound at drug-binding site i and free (unbound) drug is governed by the equilibrium constant K_i, according to

$$K_i = D_i/D_0(S_i - D_i) \tag{1}$$

where S_i is the concentration of sites i (equal to the concentration of DNA polymers), D_i is the concentration of polymers having drug bound at site i, and D_0 is the concentration of free (unbound) drug. For cutting sites within the ith drug-binding site, the probability that the probe can cut is thus determined by

$$K_iD_0 = \nu_i/(1 - \nu_i) \tag{2}$$

where ν_i is the fraction of DNA polymers having a drug bound at drug-binding site i (and thus not cleavable at all sites in a 7-base-pair inhibition region centered near site i). D_0 is determined by equilibria like that given in eq 1 between free drug and drug bound to sites on the labeled DNA and the unlabeled calf thymus carrier DNA. Thus, the equation

$$c_t = D_0 + c_c\frac{K_cD_0}{1 + K_cD_0} + c_w\frac{K_wD_0}{1 + K_wD_0} + \frac{1 \times 10^{-6}}{139}\sum_i \frac{K_iD_0}{1 + K_iD_0} \tag{3}$$

to be solved for D_0 for each total drug concentration c_t, by successive iteration. Here, c_c and c_w are the concentrations of strong and weak sites on the carrier, K_c and K_w are the respective binding constants, and the concentration of any site i on labeled fragment is the fragment concentration in base pairs, 1×10^{-6}, divided by the number of sites per fragment.

If carrier is present in large excess, c_c and c_w are probably much larger than $(1 \times 10^{-6})/139$, and equilibria with carrier dominate (eq 3).

Modeling the carrier would not be necessary if it consisted, instead of calf thymus DNA, of unlabeled copies of the fragment studied. The theoretical analysis would be simplified (determination of c_c, c_w, K_c, and K_w would be unnecessary), but one would have to have larger amounts of the fragment and accurately determine its concentration. Most drug–DNA footprinting experiments to date have not used unlabeled fragment as the carrier.

The probability of cutting at a site near i, if the drug-binding site is an isolated one, i.e., not overlapped with other drug sites, is proportional to $1 - \nu_i$, which is equal to $(1 + K_iD_0)^{-1}$. If the cleavage site is near two drug sites, j and k, so that drug bound at either site will inhibit cleavage by DNase I, $1 - \nu_i$ is equal to $[(1 + K_jD_0)(1 + K_kD_0)]^{-1}$. This assumes that drug binds independently at the two sites j and k. If the sites are close enough together so that a drug molecule bound at one prevents drug from binding at the other but that a drug bound at either one prevents cutting at site i, $(1 - \nu_i)$ is replaced by $(1 + K_jD_0 + K_kD_0)^{-1}$. Such anticooperative binding occurs for the strong sites (GC's) at 101–102 and 103–104 of the 139-mer (Figure 2).

Enhancements in Cleavage Rate. The amount of cutting at site i, in addition to being proportional to $(1 - \nu_i)$, is proportional to the local concentration of DNase I. If the amount of DNase I bound to the polymer is small and remains constant as drug is added, the binding of drug at its binding sites must redistribute the enzyme to sites where no drug is bound, leading to enhanced cutting. Evidence for this is the constancy of "total cut" with drug concentration, the total cut being the sum of the measured intensities, corresponding to 93% of all fragments produced by cleavage (see Figure 4). The local concentration of enzyme, and hence the cutting rate, at a site where no drug is bound is inversely proportional to the fraction of the sites on the DNA polymer which are *not* blocked by drug (Ward et al., 1988). To take this into account, the local concentration of probe, the factor $(1 - \nu_i)$, is multiplied by $(1 - K_ec_b)^{-1}$, where K_e is the enhancement constant and c_b is the total concentration of drug bound on the fragment. The value of K_e can be estimated from the fraction of the sites on the DNA polymer which are binding sites, but we generally allow its value to be determined by the minimization of D (see below).

Of course, the number of DNase I molecules bound to the fragment cannot remain constant for indefinitely large concentrations of bound drug; eventually, enough drug will be bound to remove probe molecules from the fragment. Therefore, the enhancement factor should behave as $(1 - K_ec_b)^{-1}$ only for small c_b, but it should increase less rapidly for large c_b, never becoming infinite. It can be shown (Goodisman & Dabrowiak, 1992) for a simplified situation that the enhancement factor behaves like $(1 - K_ec_b)^{-1}$ for small c_b and remains finite even when all sites are blocked. Since $(1 + K_ec_b)$ behaves like $(1 - K_ec_b)^{-1}$ for small c_b but increases less rapidly for larger c_b, one might use $(1 + K_ec_b)$ as the enhancement factor. In the example studied (Goodisman & Dabrowiak, 1992), it is found to work *less* well than $(1 - K_ec_b)^{-1}$, but the truncated power series $\sum_{i=0}^{n}(K_ec_b)^i$ are found to work better.

Spot Intensities. The intensity of the spot on the autoradiogram due to a fragment produced by cutting at site i, when the drug concentration is c_t, is

$$\tilde{I}_i(c_t) = k_i(1 - \nu_i)(1 - K_ec_b)^{-1} \quad (4)$$

The tilde refers to a theoretical value; $I_i(c_t)$ is the measured intensity. In the expression for $\tilde{I}_i$, k_i includes a cutting rate constant and various proportionality constants, and it is independent of c_t. For a site which cannot be blocked by drug binding (enhancement site), $\tilde{I}_i(c_t) = k_i(1 - K_ec_b)^{-1}$.

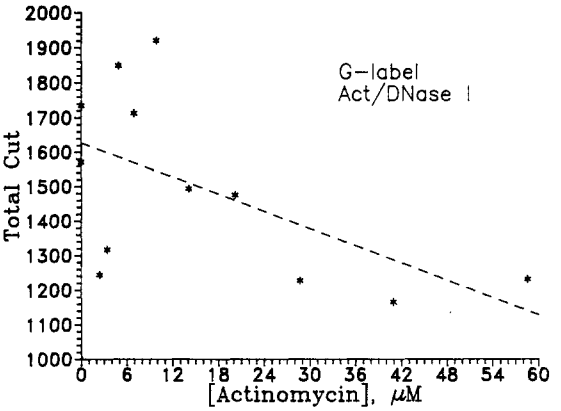

FIGURE 5: Total-cut plot for the G-labeled experiment. See caption for Figure 4.

The theoretical intensities $\tilde{I}_i(c_t)$ can now be calculated in terms of the various parameters: binding constants K_i, enhancement constant K_e, and carrier properties c_c, c_w, K_w, and K_c. We determine the values of these parameters by minimizing the mean-square deviation

$$D = \sum_{ij}[I_i(c_{tj}) - \tilde{I}_i(c_{tj})]^2 \quad (5)$$

Here, c_{tj} is the jth total drug concentration. The minimization is performed using the Simplex search technique (Fletcher, 1980).

The totals of the cut-fragment intensities for the A-label gel (26 values, corresponding to total drug concentrations from 0 to 38.8 μM), plotted in Figure 4, were fitted to a constant term, to a linear function, and to a quadratic function of drug concentration c_t. Since the linear and quadratic functions gave no significant improvement in the fit, total cleavage on the 139-mer is essentially constant with added Act-D, justifying our use of the enhancement model discussed above. The linear function $247.1 - 0.825c_t$ (dashed line in Figure 4) was used to smooth the intensities in the individual footprinting plots (correcting for errors in loading DNA into the gel, etc.) by multiplying all $I_i(c_t)$ for a given c_t by $(247.1 - 0.825c_t)/\sum_i I_i(c_t)$. The total-cut results for the G-label gel, shown in Figure 5, were treated similarly. The linear fit, $J(c_t) = 1627.1 - 8.29c_t$, was used to smooth intensities for this gel in the way this was done for the A-label gel, by multiplying all $I_i(c_t)$ by $J(c_t)/\sum_i I_i(c_t)$.

For the A-label gel, we rejected data from sites weakly cleaved by DNase I, retaining data for 54 sites. For the G-label gel, data for 43 sites were retained. Some of the footprinting plots used in the analysis for the A-label gel are shown in Figure 6 (after total-cut correction).

In our previous work (Rehfuss et al., 1990b), we considered only the strong binding events, using footprinting plots for the following sites: 54–56, 60, 62–69, 71–72, 85, 87, 98–99, 102–103, 106, 112, 114, 120, 124, 128, 133, 136, 138, 143, 145, and 161. Because the position of the bands is not linear in site number and because there is not a spot for each band, assignment of the bands to sites is problematic, especially for sites above 110, corresponding to larger DNA oligomers, because resolution decreases. However, the G-labeled fragment, which has the ^{32}P atom on the opposite strand and at the opposite end of the 139-mer from the A-labeled fragment (Figure 2), gives a gel for which site resolution is best for the higher numbered sites, which now correspond to smaller oli-

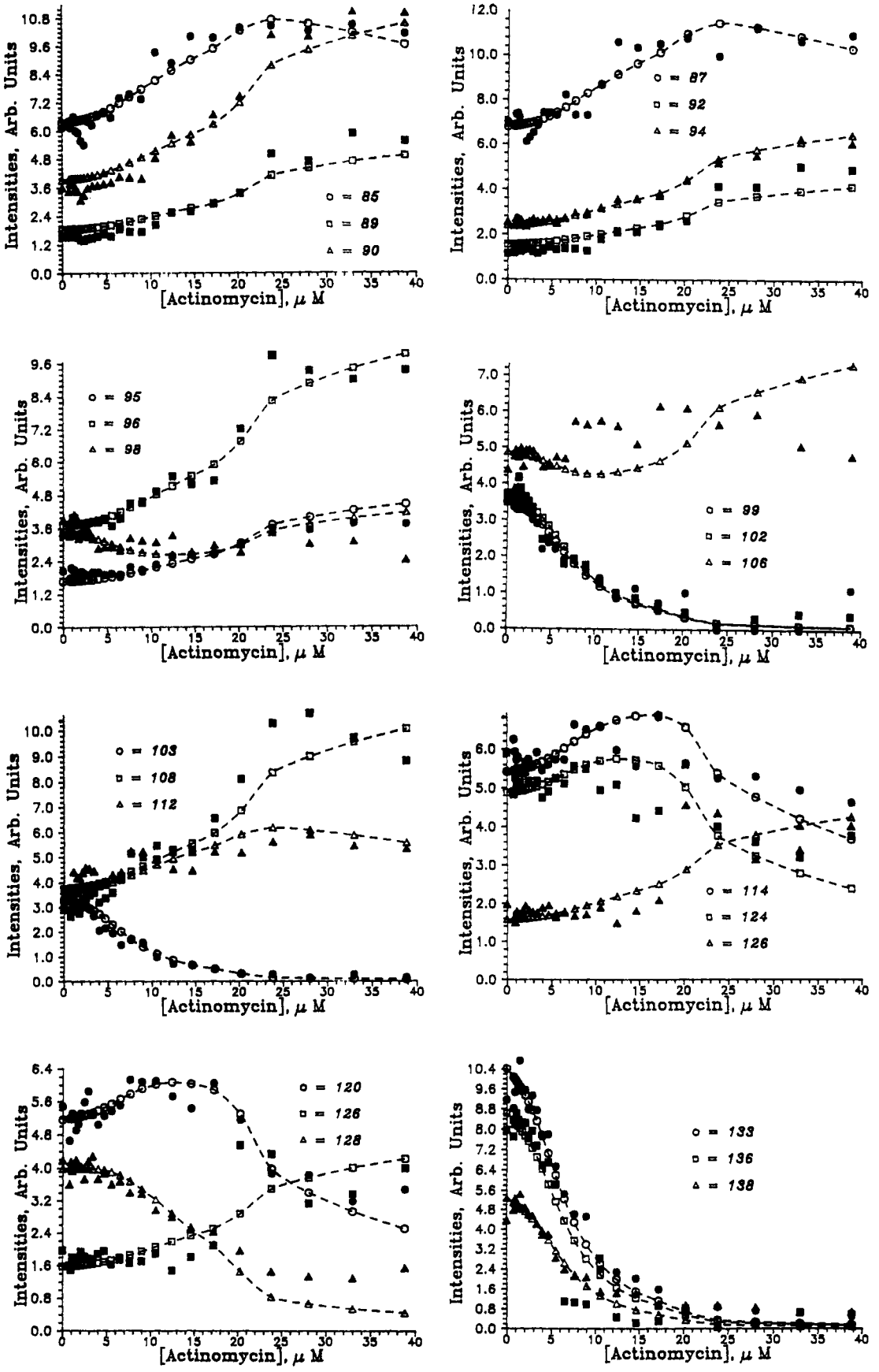

FIGURE 6: Footprinting plots from the A-label gel for selected sites in the region 85–138 of the fragment. Solid symbols are experimental gel spot intensities; fits calculated according to our best model (column III of Table III) are shown as open symbols and dashed lines. Sites 89, 90, 95, 96, and 108 are enhancement sites. Sites 133, 136, and 138 show strong drug binding.

gomers. Comparing the gels from the A-labeled and the G-labeled fragment facilitates indexing of these bands. The numbering used previously is apparently low by several site numbers for sites above about 120.

To identify strong binding sites, we found it useful to calculate the initial slopes of the footprinting plots (fragment intensities vs total drug concentration) and divide the slopes by the intercepts (intensities for zero drug concentration) to get "initial relative slopes" (Ward et al., 1988). Stronger binding sites are easily identified by negative slopes, while positive slopes characterize weaker or nonbinding (enhancement) sites. The constancy of initial relative slopes over enhancement sites provides further evidence for the validity of the redistribution–enhancement model (Ward et al., 1988). Deviations from constancy may indicate drug-induced structural changes in DNA, manifesting themselves in a change in cleavage rate constant at the site (Rehfuss et al., 1990b).

In the previous analysis (Rehfuss et al., 1990b), sites were classified as either enhancement or binding sites. Some of the sites considered as enhancements at that stage (for example, 85, 87, 112, 114) are revealed to be influenced by weak-binding drug sites when results for drug concentrations greater than 7.64 μM are considered. For lower drug concentrations, however, they behave like typical enhancements, as can be seen from the goodness of fit by the theoretical enhancement–site curves (dashed lines in Figure 6). Clearly, examination of initial relative slopes is not useful when one considers weaker sites, since intensities for such sites show an initial enhancement (followed by a decrease in intensity as c_t increases further. It is necessary to consider the entire footprinting plot in order to classify a site as binding or enhancement.

Role of Carrier DNA. Since appreciable binding to the weaker sites takes place only at higher drug concentrations, we modeled the carrier in our previous work (Rehfuss et al., 1990b) as consisting of only one kind of site, with concentration c_c and binding constant K_c, K_c representing the average of the strong binding sites on the carrier. We carried out the minimization of D using several different values of c_c, and we found that the minimum value of D was lowest near $c_c = 4.7$ μM. The enhancement factor was written as $(1 - K_e c_b)^{-1}$, where c_b is the concentration of drug bound to fragment. The value found for the enhancement constant K_e was 4.9×10^6 M^{-1}. K_e should be equal to the reciprocal of the concentration of drug–binding sites on the fragment. Since the fragment is 139 base pairs long and each site occupies 7 base pairs, there are 20 sites per fragment; multiplying by the fragment concentration of 1/139 μM should make K_e about 20/139 μM^{-1} = 7×10^6 M^{-1}.

With 19 drug concentrations and 32 sites, the calculation of the deviation D involved 608 data points. The lowest value of D (Rehfuss et al., 1990b) was 126, making the root-mean-square deviation $\sqrt{126/608} = 0.46$, about the same as the error in the data points, as judged by their scatter. However, certain footprinting plots were noticeably less well fitted than others, indicating inadequacies in the model. The low-concentration part of the G-label gel was analyzed similarly to the A-label gel. The determined binding constants for strong sites were in agreement with those from the A-label gel.

In the present calculations, designed to determine weak- and strong-site binding constants, we consider footprinting plots from the A-label gel for all sites for which appreciable intensities are available. We develop the model used to interpret the A-label gel below, and we derive the strong-site and weak-site binding constants. Then, the model is used to interpret footprinting plots from the G-label gel which have appreciable intensities and to obtain binding constants. Although the data from the G-label gel show a lot of scatter, its autoradiogram has the best site resolution for the higher numbered sites, as mentioned above. By comparing the plots derived from the G-label gel with those obtained from the A-label gel to check on site numbering, we conclude that the A-label site numbers are too low by about 3 for sites after about 120. The original reference numbering is used in the present paper.

Since we are now dealing with the full range of drug concentrations (to 38.8 μM), we must include the additional parameters c_w and K_w, representing the concentration and average binding constant of weak sites on carrier DNA. These enter the determination of the free-drug concentration D_0. We use a value of 5 μM, as found in earlier calculations (Rehfuss et al., 1990b), for the concentration of strong sites on carrier and (initially) 13 μM for the weak-site concentration c_w since, on the fragment, there are about $2^1/_2$ as many weak sites as strong ones. We determine the value of K_w in the minimization, along with the strong-site binding constant K_c, the enhancement constant K_e, and the site-binding constants for the fragment.

Analysis: Development of the Model

We now discuss the development of the model as it applies to the A-label gel data. We are able to use data for 54 sites, and there are 26 drug concentrations, giving 1404 data points to be fit. Table I summarizes the behaviors of the 54 footprinting plots used (uncorrected site numbers are given), and it gives the assumptions of the model used to explain them. To limit the number of parameters, several unknown weak binding sites have been assigned the same binding constant, K_7, although one might consider assigning them different weak binding constants, and K_9 has been used to describe the binding to the sequence CCC at 129–131, even though this binding constant need not be the same as that for the sequence CCGT at 123–126.

The two GC sites at 101–102 and 103–104 are so close together that binding of drug at one excludes binding at the other. This exclusion explains why the intensity for cutting at site 98 (Figure 4) never decreases to zero (cutting here is blocked by drug binding at 101–102, which is in turn blocked by drug binding at 103–104) and why the intensities for sites 99, 102, and 103 decrease more sharply with drug concentration than the intensity for site 106 (they are affected by binding at either GC site). An attempt to include a similar kind of exclusion between other nearby weak binding sites (e.g., the GGC at 119–121 and the CCG at 123–126) failed: D increased when the exclusion-binding model was used for sites in this region. Presumably, drugs bound at these two positions are not close enough to interfere with each other. We also tried to put in a second binding constant for the weak sites in the region 160–165, which is 5′-GCCG-3′, but minimization of D returned a very small value for this binding constant. This means that 5′-CG-3′ is so much weaker than 5′-GC-3′ as a binding site that binding to the former site can be neglected.

Weak Act-D Sites on Carrier DNA. With the carrier-site concentrations given above, and an enhancement factor of $(1-K_e c_b)^{-1}$, we minimize the sum of the squared deviations D with respect to K_e and the binding constants on fragment and carrier. This represents a total of 16 nonlinear parameters. The binding constants obtained, with c_c and c_w fixed respectively at 5 and 13 μM, are given in column I of Table II. The logarithms of the fragment binding constants K_1–K_{13} are given, followed by the logarithms of the carrier binding constants K_c

Table I: Behavior of Footprinting Plots and Their Interpretation from the A-Label Gel[a]

site no.[b]	behavior of spot intensities	interpretation
54, 55	inc with c_t	very weak bind., K_7
56–59	rap. inc with c_t	enh
62–66	rap. dec	bind. to GC at 63–64,[c] K_1
67	very rap. dec	bind. to GC's at 63–64 and 69–70, K_2
68–69	rap. dec	bind. to GC at 69–70, K_2
71–72	rap. dec	bind. to GC at 69–70
75–78	inc, then dec	bind. to GGC at 76–78, K_3
79	inc, then dec	bind. to 76–78 and CCG at 80–82, K_4
80	unchanged with c_t, then dec	bind. to CCG at 80–82, K_4
81, 83	inc, then dec	bind. to CCG at 80–82
85, 87	slowly inc	unknown weak bind. site, K_4
89, 90	inc	enh
92, 94	inc	enh
95	slowly inc	enh
96	inc	enh
98	dec with c_t	bind. to GC at 101–102, K_5
99, 102	rap. dec with c_t	bind. to GC at 101–102 or 103–104, K_6
103	rap. dec	bind. to GC at 101–102 or 103–104
106	rap. dec	bind. to GC at 103–104, K_6
108	inc	enh
112	inc, then dec	bind. to nonclassical site, K_7
114	inc, then dec	bind. to GGC at 119–121, K_8
120	unchanged with c_t, then dec	bind. to CCGT at 123–126, K_9
124	inc, then dec	bind. to CCGT at 123–126
126	rap. inc	enh
128	inc, then dec	bind. to CCC at 129–131, K_9
133, 136	dec with c_t	bind. to GC at 137–138, K_{10}
138	dec with c_t	bind. to GC at 137–138
143, 145	inc, then dec	bind. to GGC at 143–145, K_{11}
147	inc, then dec	bind. to GGC at 149–151, K_{12}
150	inc	enh
154	inc, then dec	bind. to GC at 160–161, K_{13}
158, 161	inc, then dec	bind. to GCCGG at 160–164, K_{13}

[a] Abbreviations: rap., rapid; inc, increase; dec, decrease; bind., binding; enh, enhancement. [b] Site numbers refer to the sequence shown in Figure 2. Numbers greater than ~110 are probably low by ~3. [c] All sequences are 5′ to 3′.

and K_w (for strong and weak sites), the value of the enhancement constant K_e in the factor $(1 - K_e c_b)^{-1}$, and the value of D.

All of the binding constants for strong sites on the fragment are significantly higher than the corresponding binding constants determined previously (Rehfuss et al., 1990b). The reason for this is the inclusion of weak binding sites on the carrier, as is shown by carrying out the minimization of D with c_w reduced from 13 to 0.05 μM (essentially zero). The fragment binding constants are all reduced, as shown in column II. (Of course, the value of K_w must increase markedly when the value of c_w is reduced markedly.) With a low or zero value of c_w, the maximum concentration of drug that can be bound to the carrier is 5 μM (the value of c_c), so that when the total drug concentration exceeds 5 μM, the free-drug concentration rises rapidly with total drug concentration and smaller fragment binding constants are required to explain the measured binding isotherms. Since D for the low value of c_w (calculation II) is much larger than for c_w = 13 μM (627 vs 517), we have to accept an appreciable value of c_w as correct, with the higher binding constants that result.

To ascertain whether the value of c_w can be determined from the minimization of D and, if it can, what the value of c_w is, we performed minimizations of D assuming c_w = 10 and 7 μM. The results are shown in columns III and IV of Table II. The values of D from columns I, III, and IV show a shallow minimum as a function of c_w. Fitting them to a parabola, we find the minimum at c_w = 10.3 μM, with D = 507. In all subsequent calculations, we maintain c_w at 10 μM. It may be noted that increasing c_w leads in general to increases in fragment binding constants, but that the changes when c_w varies from 7 to 13 μM are small. The binding constant K_w must of course decrease with increased c_w, while K_e hardly changes with c_w.

Enhancement Factor. As mentioned above, the enhancement factor $(1 - K_e c_b)^{-1}$ cannot be correct for large values of c_b, since this factor becomes infinite for $c_b = 1/K_e$, whereas the enhancement should never become infinite (Goodisman & Dabrowiak, 1992). When $(1 + K_e c_b)$, which behaves like $(1 - K_e c_b)^{-1}$ for small c_b and remains finite for all values of c_b, was used for the enhancement factor, the results of column V of the table were obtained. The resulting binding constants differ only slightly from those of column III, but D is much larger (554 instead of 507), so that the factor $(1 - K_e c_b)^{-1}$ fits our data better, and this form of the enhancement is used in succeeding calculations.

The lowest value of D so far is from calculation III, D = 507. The root-mean-square deviation is $\sqrt{[507/(26 \times 54)]} = 0.60$, which is somewhat greater than our estimate for the precision of the experimental intensities. Thus, we cannot state that our model fits the data to within experimental error. This implies that some effects are not yet included in our model,

Table II: Effect on Calculated Binding Constants of Weak-Site Concentration on Carrier and Expression Used for Enhancement

log of binding constant	calculation I c_w = 13 μM, $(1 - K_e c_b)^{-1}$	II c_w = 0.05 μM, $(1 - K_e c_b)^{-1}$	III c_w = 10 μM, $(1 - K_e c_b)^{-1}$	IV c_w = 7 μM, $(1 - K_e c_b)^{-1}$	V c_w = 10 μM, $(1 + K_e c_b)$
K_1	6.37	5.60	6.15	6.31	6.16
K_2	6.01	5.48	6.07	6.01	6.07
K_3	5.17	4.88	5.15	5.09	5.11
K_4	5.13	4.92	5.18	5.06	5.12
K_5	6.05	5.83	6.15	6.01	6.05
K_6	5.95	5.65	5.96	5.82	5.97
K_7	4.31	4.23	4.41	4.33	4.08
K_8	4.97	4.87	4.95	4.94	4.85
K_9	5.23	5.01	5.19	5.16	5.20
K_{10}	6.61	6.01	6.60	6.56	6.68
K_{11}	5.35	5.00	5.29	5.19	5.26
K_{12}	4.40	4.42	4.47	4.51	4.35
K_{13}	5.55	5.16	5.46	5.39	5.43
K_c	7.13	6.89	7.21	7.01	6.92
K_w	5.78	7.10	5.93	6.20	6.12
K_e	7.9×10^6	8.2×10^6	8.2×10^6	8.0×10^6	1.58×10^7
D	517	627	507	522	554

Table III: Effect on Calculated Binding Constants of Inclusion of Cooperative Binding and Changed Enhancement Expression

	calculation		
log of binding constant	I c_w = 10 μM, $(1 - K_e c_b)^{-1}$, 17 parameters	II c_w = 10 μM, $(1 - K_e c_b)^{-1}$, cooperativity	III c_w = 10 μM, $\sum_{i=0}^{10}(K_e c_b)^i$, cooperativity
K_1	6.17	6.16	6.16
K_2	5.94	5.90	5.87
K_3	5.14	5.00	4.95
K_4	5.16	5.10	4.89
K_5	6.23	6.07	6.01
K_6	6.13	5.90	5.92
K_7	4.33	4.14	4.17
K_8	4.94	4.81	4.71
K_9	5.10	4.95	4.89
K_{10}	6.63	6.48	6.43
K_{11}	5.25	5.05	5.02
K_{12}	4.45	4.33	4.30
K_{13}	5.46	5.27	5.24
K_{14}	5.65	5.31	5.60
K_c	7.28	7.26	7.03
K_w	5.92	5.82	5.67
value of K_e	7.2×10^6	7.2×10^6	7.1×10^6
D	495	477	435
Parameters in tanh Factor			
a_1		1.71	1.70
a_2		5.0×10^7	4.0×10^8
a_3		6.9×10^{-8}	5.9×10^{-8}

for example structural changes. Sites for which the fit of theoretical to experimental footprinting plots is worst at this stage are 78, 87, 112, 128, and 147.

The solution to the problem for site 128 has already been noted: cutting at this site is inhibited by drug binding at the CCC sequence located at base pairs 129 to 131, and the drug binding constant for this sequence may be different from K_9, which describes drug binding to the CCG sequence located at 123–125. By including an additional constant K_{14} to describe drug binding to the CCC, we minimize D with respect to 17 nonlinear parameters, obtaining the results in column I of Table III. We note a reasonable improvement in D, from 507 to 495, when the additional parameter is inserted; values of other parameters are not changed much.

Increased Binding. Comparison between theory and experiment shows that, in general, the footprinting plots calculated from our model for binding sites do not decrease sharply enough with total drug concentration for concentrations near 20 μM. Correspondingly, footprinting plots calculated for enhancement sites do not increase sharply enough in this region. Apparently, there is a rather rapid increase in the concentration of drug bound to the weak sites on the fragment when the total drug concentration increases past 20 μM. (Because intensities for cutting at sites blocked by drug binding at strong sites are already near zero at this point, no information is available about the amount of drug bound to strong sites.) This could be due to a structural change in the fragment induced by the bound drug or to cooperative binding. It may be noted that there is a sudden (although relatively small in magnitude) drop in total cut (Figure 4) after [Act–D] = 20 μM. This affects the correction factors used in the total-cut smoothing, but this cannot cause both sharp increases and sharp decreases in individual footprinting plots; these are present in the unsmoothed plots.

We model the increase in the effective drug binding constant by multiplying each binding constant K_i (i = 1–14) by the same function of bound drug concentration:

$$K_i^{\text{eff}} = K_i(a_1 + (a_1 - 1)\tanh[a_2(c_b - a_3)]) \qquad (6)$$

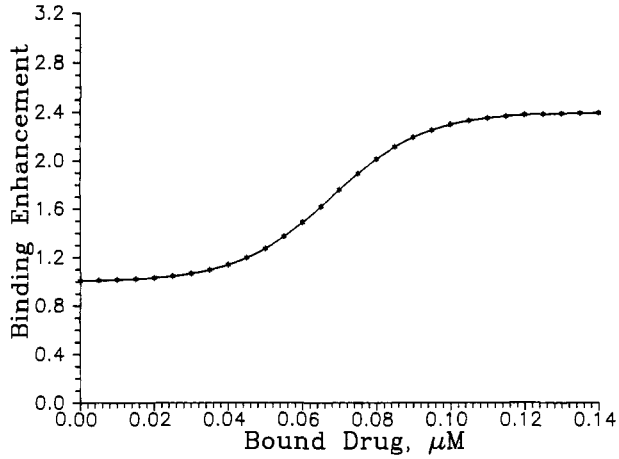

FIGURE 7: Increased binding factor, K_i^{eff}/K_i (eq 6), which multiplies binding constants to give effective binding constants, as a function of the bound drug concentration. Optimum (minimum D) values of the parameters are used.

According to the above function, the effective binding constant is K_i in the absence of bound drug, when c_b, the calculated concentration of drug bound to the fragment, is much less than a_3. It becomes $(2a_1 - 1)K_i$ when c_b is much greater than a_3, the parameter a_2 giving the sharpness of the change. With eq 6, three new parameters are introduced, making the total number 20. As shown in column II of Table III, there is a small but significant improvement in the fit.

The enhancement factor for binding (eq 6) is plotted as a function of c_b in Figure 7, using the best values of the parameters a_1, a_2 and a_3. For the five highest total drug concentrations (20.3, 23.8, 28.0, 33.0, and 38.8 μM), c_b is calculated as 0.059, 0.075, 0.080, 0.084, and 0.087 μM, respectively. Then, the enhancement in binding occurs mostly between total drug concentrations of 20 and 24 μM, which is where the sharp intensity decreases occur. In fact, D is found to be insensitive to the value of a_2. This is because we know only that most of the increase in drug binding (by a factor of 2.4) occurs between 20 and 24 μM, so that $a_2(24–20)$ μM must be much larger than unity, and $a_2 \gg 2.5 \times 10^5$ M^{-1}.

On studying the observed and calculated footprinting plots, we observe one remaining problem: calculated plots for enhancement sites show intensities which rise sharply with total drug even for the highest drug concentrations, whereas the experimental intensities apparently level off. This is related to the fact mentioned above that $(1 - K_e c_b)^{-1}$, which becomes infinite for $c_b = 1/K_e$, cannot be the correct enhancement factor for large c_b, but only for small c_b. Functions behaving like $(1 - K_e c_b)^{-1}$ for small c_b but increasing less for large c_b are the power series' $\sum_{i=0}^{n}(K_e c_b)^{-1}$ for finite values of n. The linear function (n = 1) has already been tried and found unsatisfactory (column V, Table I). After some experimentation, we chose n = 10, although the exact value of n is not critical; the results, which include a significant decrease in D from $n = \infty$, corresponding to $(1 - K_e c_b)^{-1}$, are given in column III in Table III.

D is reduced to 435 and the root-mean-square deviation is now $\sqrt{[435/(54 \times 26)]}$, or 0.56. The footprinting plots now seem to reproduce the experimental ones quite well, with a few exceptions (see Figure 6). To improve our model by increasing the number of variable parameters would be unjustified and would require more work, with probably little decrease in D. The three parameters in eq 6 already have little chemical basis except that they produce the increased binding shown by the data. Note that the various improvements in

Table IV: Binding Constants of Act-D on the 139-mer ($\times 10^{-5}$ M)

position[a]	sequence	A-label gel	G-label gel
62–65	TGCT	35.0	4.8
68–71	CGCA	18.0	
76–79	GGCA	2.1	1.3
80–83	CCGT	1.8	1.7
100–103	TGCG	25.0	22.0
102–105	CGCT	20.0	8.8
various	various	0.3	
118–121	CGGC	1.2	0.3
123–126	CCGT	1.8	0.9
128–131	ACCC	9.4	4.8
136–139	TGCT	64.0	30.0
143–146	GGCA	2.5	2.0
149–152	GGCT	0.5	0.8
159–162	TGCC	4.2	4.5

[a] Site numbers refer to the sequence in Figure 2.

Table V: Behavior of Footprinting Plots and Their Interpretation from the G-Label Gel[a]

site no.[b]	behavior of spot intensities	interpretation
47, 49, 56	inc with c_t	enh
65	rap. dec	strong bind. to GC at 63–64,[c] K_1'
75	inc, then dec	bind. to GGC at 76–78, K_3'
82, 84	level, then dec	weak bind. site, K_4'
93, 95	inc	enh
101	rap. dec with c_t	bind. to GC at 101–102, K_4'
105	rap. dec	bind. to GC at 101–102 or 103–104, K_5'
110	inc with c_t	enh
119	inc, then dec	bind. to GGC at 119–121, K_6'
122	inc, then dec	bind. to GGC at 119–121
123–126	inc, then dec	bind. to CCGT at 123–126, K_7'
127, 128	inc, then dec	bind. to above and CCC, 129–131
130–132	inc, then dec	bind. to CCC at 129–131, K_8'
134, 135	dec with c_t	bind. to CCC at 129–131
136–141	dec with c_t	bind. to GC at 137–138, K_9'
142	inc, then dec	bind. to GGC at 143–145, K_{10}'
144, 146	inc, then dec	bind. to GGC at 143–145, K_{10}'
148	level, then dec with c_t	bind. to GGC's at 143–145, 149–151
149–150	level, then dec with c_t	bind. to GGC at 149–151, K_{11}'
152, 153	level, then dec with c_t	bind. to GGC at 149–151
154	inc with c_t	enh
158	inc, then dec	bind. to GCCGG at 160–161, K_{15}'

[a] Abbreviations: rap., rapid; inc, increase; dec, decrease; bind., binding; enh, enhancement. [b] Site numbers refer to the sequence shown in Figure 2. [c] All sequences are 5′ to 3′.

our model, while lowering D significantly, have had little effect on the calculated values of the binding constants for fragment sites. Thus, further changes in our model will probably not change the K's much. The binding constants for our best calculation (lowest D) are given in Table IV.

Footprinting with G-Labeled 139-mer. We now consider the data from the G-label gel, analyzing it by the model developed for analysis of the A-label gel. This means that we use the same values for the concentrations of strong and weak sites on the carrier rather than determine the best values for these concentrations by minimization of the mean-square deviation D. For the enhancement factor, we use $\sum_{i=0}^{10}(K_e c_b)^{-1}$, and no correction for cooperative binding is included.

Intensities for nine actinomycin concentrations (2.4, 3.4, 4.9, 6.9, 9.9, 14.1, 20.2, 28.7, and 40.9 μM) plus two controls were used. Having two sets of intensities for zero drug concentration allows an assessment of the error in the intensities. In general, the data from the G-label gel show more scatter than those from the A-label gel. This is evident in the plot of total cut vs drug concentration (Figure 5). The linear fit to the total-cut plot, $1627.1-8.29c_t$, was used to smooth intensities, as discussed above for the A-label gel. There appears to be a drop-off in the total cut at a drug concentration of 20–24 μM, as in the A-label gel. Several footprinting plots are shown in Figure 8. The scatter in the data should be noted, as well as the

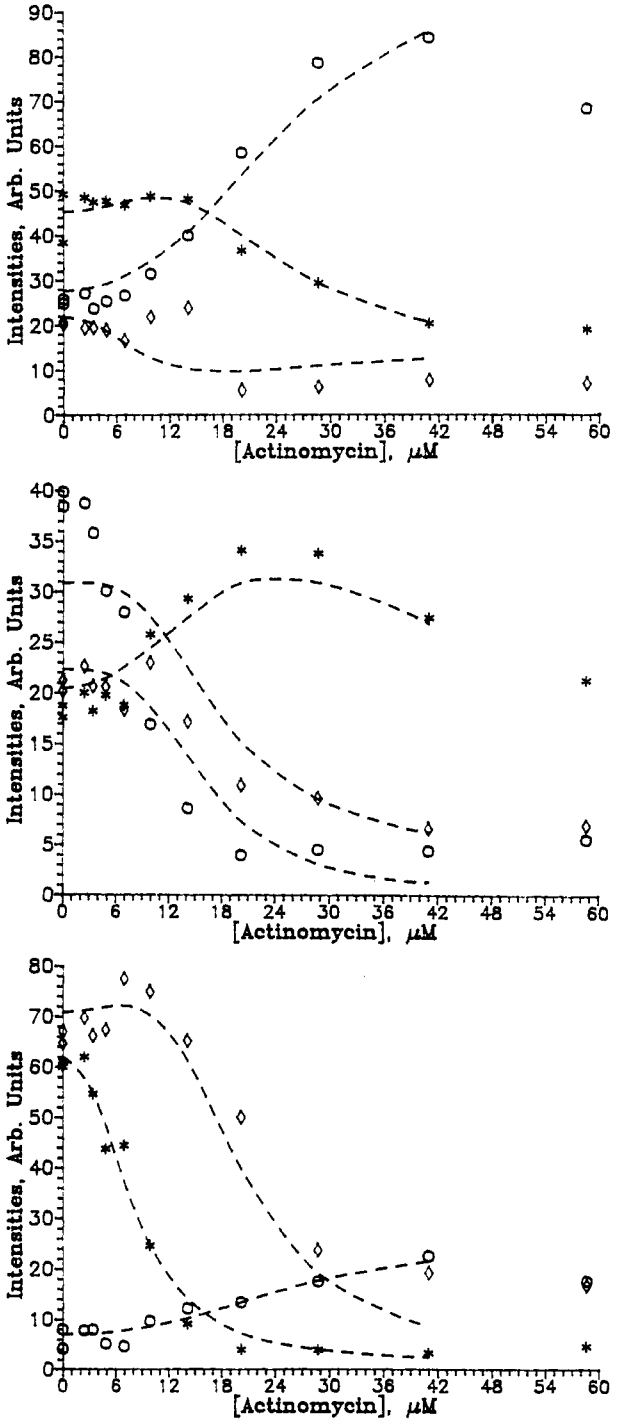

FIGURE 8: Footprinting plots from the G-label gel, after total-cut smoothing. Symbols are measured spot intensities; broken curves are fits calculated from the model. In the top plot, *, ◇, and ○ are for sites 84, 101, and 110; in the middle plot, *, ◇, and ○ are for sites 122, 127, and 135; in the bottom plot, *, ◇, and ○ are for sites 138, 148, and 154.

apparent large changes in intensities for some sites near 20 μM drug concentration.

We used data for all sites for which intensities were appreciable and the fluctuations were significantly smaller than the intensities themselves. This means that 43 sites were considered, as listed in Table V, and 473 total data points. Table V indicates how each footprint was modeled. Note that the K_i' for this gel do not always correspond to the K_i for the A-label gel.

There are a total of 15 parameters in this model: 12 binding constants on the fragment, the strong- and weak-site binding

constants on the carrier, and the enhancement constant K_e. Minimization of D with respect to these parameters gives a value of 1.03×10^4. Dividing by 473 and taking the square root gives a mean-square deviation of 4.7, which seems slightly bigger than the fluctuations in the data points. The carrier binding constants were found to be 89.3×10^5 and 8.9×10^5 M^{-1} for the strong and weak carrier sites, and the enhancement constant was 7.7×10^6 M^{-1}.

The values of the fragment binding constants are given in Table IV. Note that, for this gel, resolution is low and site numbering is uncertain for the lower numbered sites. This probably accounts for the constant for the first site being much lower than that calculated from the A-label gel: spots for the single-site 65 considered to be responding to binding to the GC at 63–64 were actually superpositions of spots arising from cutting at several sites. Thus, the value of this binding constant is very poorly determined from the G-label gel. Similarly, the constants for the overlapped sites having GC's at 101–102 and 103–104 were determined from data from only two footprinting plots, as compared to five on the A-label gel, so one cannot put much confidence in the smaller value. On the other hand, the value found for the binding constant for the site having GC at 137–138 is probably more reliable from the G-label than from the A-label gel because six footprinting plots were used in the former case and only three were used in the latter.

Summary. It was assumed that when a molecule of actinomycin D is bound to DNA, it blocks about 7 base pairs from cleavage by the probe DNase I. The size of the inhibition region follows from the structures of the drug and probe as well as from previous footprinting work on this system. In setting up our model, we must decide which drug-binding events, if any, affect cleavage at each site for which we have a footprinting plot. A consistent picture could only be made if the numbering of the footprints for the A-label gel was understood to be too low by about 3 base pairs for sites after about 110, where the numbering for the A-label gel becomes uncertain. That the original numbering was incorrect in this way was confirmed by observation of footprinting data obtained from G-labeled fragment, for which assignment was more accurate for the higher numbered sites.

The binding of drug to a site on DNA is governed by the simple chemical equilibrium

$$K_i = \frac{c_i}{(c - c_i)D_0} \tag{7}$$

where K_i is the binding constant (whose value our analysis attempts to derive), c_i is the concentration of fragments which have drug bound at site i, c is the total concentration of fragments (so $c - c_i$ is the concentration of fragments having no drug bound at site i), and D_0 is the concentration of free or unbound drug. Where two strong binding sites are adjacent, drug binding at one prevents drug binding at another and eq 7 must be modified. The free-drug concentration is related to the known total drug concentration c_t by

$$c_t = D_0 + c_b + c_{ca} \tag{8}$$

where c_b is the concentration of drug bound to the fragment ($c_b = \sum_i c_i$) and c_{ca} is the concentration of drug bound to sites on the carrier DNA. The value of c_{ca}, which was expected to be much larger than c_b because there was about 200 times more carrier than labeled fragment, was determined by equilibria like that in eq 7. For simplicity, we assumed that carrier sites could be represented by two kinds: strong sites of concentration c_c and binding constant K_c, and weak sites of concentration c_w and binding constant K_w. This leads to eq 3.

By solving all the equilibrium equations simultaneously, we determined c_i for each site i and for each c_t. For any cleavage site within the inhibition region of drug site i, the probability that cleavage can occur is the probability that no drug is bound at site i, $1 - c_i/c$. Some cleavage sites may be blocked by two independent drug-binding events; if the drug sites are j and k, the probability of cleavage is $(1 - c_j/c)(1 - c_k/c)$.

In addition, we noted that if bound drug prevents cleavage at some sites on a fragment by blocking access of the enzyme to part of the DNA, the concentration of enzyme at other places on that fragment, where no drug is bound, should increase. Thus, the cleavage rate, which is proportional to the local concentration of enzyme, should increase at those places. This is the origin of the enhancement factor, which multiplies all intensities. The factor was written as $E = (1 - K_e c_b)^{-1}$, where K_e is a constant and c_b is the concentration of drug bound to fragment, or as a finite power series, $\sum_{i=0}^{n} (K_e c_b)^i$. Finally, at total drug concentrations high enough for appreciable binding to the weak sites, we noted an abrupt decrease in amounts of cleavage products with drug concentration. This is modeled by multiplying all binding constants by an increasing function of c_b (eq 6). The values of the parameters in this function are determined, along with the binding constants to 139-mer and carrier and the enhancement constant K_e, by variation.

We found the set of values of the parameters which minimized the sum of the squared deviations between calculated and experimental intensities:

$$D = \sum_{i,j} (\tilde{I}_i(c_{tj}) - I_i(c_{tj}))^2 \tag{9}$$

Here, $I_i(c_{tj})$ is the measured intensity or amount of fragment resulting from cleavage at enzyme site i when the total drug concentration is c_{tj} and $\tilde{I}_i(c_{tj})$ is the corresponding intensity calculated from our model. The measured cut-fragment intensity is proportional to the probability of cleavage. For the A-label gel, there were 54 footprinting plots, each corresponding to a particular cleavage product, and 26 drug concentrations, giving 1404 points. Our best calculation involved 20 nonlinear parameters and led to $D = 435$. The root-mean-square deviation was the same size as the fluctuations in measured intensities from point to point.

Discussion

In this article, we present measured footprinting plots for actinomycin D binding to DNA, as well as the model and analysis which allow us to derive binding constants for strong and weak sites, as well as other parameters. This is the only quantitative footprinting study of which we are aware which deals simultaneously with such a large number of footprinting plots and such a large number of ligand concentrations. The model involves the same assumptions as did models used in previous work (Ward et al., 1988; Rehfuss et al., 1990b), but modifications were necessary to take into account the behaviors of the footprinting plots revealed at high drug concentrations. The footprinting plots $\tilde{I}_i(c_{tj})$ calculated from our model reproduce the behavior of the observed footprints quite well in general (see Figure 6). A measure of the experimental error or scatter in the intensities is obtained from deviations from the expected continuous variation of intensity with drug concentration, as estimated by fitting a line through several points, and is about 0.5 unit.

However, there are exceptions, identified by inspection of the footprinting plots, examples of which are shown in Figure 6, or by calculation of the contribution of each site to D: sites

59, 106, 108, 124, 126, 136, 143, and 161. For sites 108 124, 126, 136, and 161, the problem seems to be in the experimental intensities themselves, which vary irregularly with concentration. Site 59 is modeled as an enhancement, whereas the spot intensity actually remains roughly constant with drug concentration for concentrations up to 20 μM and then increases rapidly to a new constant value approximately double the first. The only nearby drug-binding site is the GC at base pairs 63–64, which is a strong site, giving binding-type plots for sites 62–65. Possibly, the special behavior of site 59 is associated with its being on the boundary of the inhibition region for this binding site. The data for site 106, which is modeled as responding to drug binding to the GCGC at 101–104, can be better fit by a curve (see Figure 6) such as those for sites 85 and 87, which were assumed to be showing inhibition due to a very weak and unspecified binding site ($K = 3 \times 10^4$). However, site 108 appears to be an enhancement site, so it is difficult to postulate a weak site to which 106 could respond.

Sites 143 and 145 are modeled as responding to binding at the GGC at 143–145. However, experimental intensities seem to drop much more abruptly near 20 μM, before leveling off, for site 143 than for site 145. This cannot be explained with our model, unless it simply reflects large errors in the experimental intensities. Also, as noted previously (Goodisman & Dabrowiak, 1991, 1992), many footprinting plots show a small rapid decrease in intensity with total drug concentration, followed by an increase, near 2 μM. Examples in Figure 6 are sites 85, 87, and 89. This behavior cannot be explained in the context of the present model. We have investigated (Goodisman & Dabrowiak, 1992) whether this indicates a structural change in the DNA like that suggested by other workers (Fox & Waring, 1984; Huang et al., 1988; Bishop et al., 1991).

Values of Parameters. We now consider the values of the parameters determined by the minimization process. According to our model, the enhancement E at sites where no drug is present is due to the increased probe concentration there, caused by drug at the other sites, so that E should be inversely proportional to the fraction of the total number of probe sites available. The concentration of probe sites (base pairs) is about 1 μM. Letting c_b be the molar concentration of drug bound to the 139-mer and assuming each drug blocks seven base pairs, we have

$$E = 10^{-6}/(10^{-6} - 7c_b)$$

If this is written as $(1 - K_e c_b)^{-1}$, we find $K_e = 7 \times 10^6$, close to what is determined by minimization. In previous work (Ward et al., 1988), we noted larger enhancements at sites adjacent to (strong) drug-binding sites. Now that many footprinting plots previously considered to be enhancements are seen to respond to weak binding sites, only a single site, 96, appears to show this effect (GC is found at 101–102).

The carrier, calf thymus DNA, is modeled as having two kinds of sites: the strong sites are given a concentration of 5 μM, as determined from previous calculations which considered strong sites only, and the weak-site concentration is determined as 10 μM from a series of calculations to determine the value that makes D smallest. These values are reasonable, although perhaps low, since the 139-mer contains about 6 strong sites and 11 weak sites in 137 base pairs and the concentration of carrier is 190 μM. The binding constants for strong and weak sites on carrier were determined by minimization of D as 1.1×10^7 and 4.7×10^5. That these values are higher than the average strong- and weak-site binding constants on the fragment may be due to the carrier-site concentrations being too low. Using higher values for these concentrations will decrease the carrier binding constants and decrease the binding constants for the fragments, but only slightly.

The use of carrier DNA makes it unnecessary to determine the concentration of radiolabeled fragment precisely, which would be difficult, but it also complicates the analysis. Since the structure of the carrier DNA is different from that of the radiolabeled fragment, only a fraction of the individual loading events taking place in the system are seen on the sequencing autoradiogram. However, other analyses have shown that a simple model of drug binding to the carrier is sufficient to obtain reliable site-specific binding constants toward the fragment. For small oligonucleotide duplexes which possess a single drug-binding site and which are readily available, unlabeled duplex may be used as the carrier. Then, quantitative analysis of footprinting data is simplified since the only loading event taking place is observed on the sequencing autoradiogram, and there are fewer parameters to determine. As it becomes easier to purify significant amounts of DNA restriction fragments, it will be more practical to use the unlabeled fragment as carrier. This will be advisable since it ensures that most (not all) of the loading events are seen on the autoradiogram, making the modeling of unobserved events less important.

Binding Constants. The binding constants for Act-D at specific sites on the 139-mer are given in Table IV. Calculated values for log K_i are probably reliable to 0.1 in most cases, which means K_i is good to 25% and the free energy of binding to 0.6 kJ/mol. Values of K_i for the G-label gel are generally less reliable because there is more scatter in the data. The value for the site at 62–65 is particularly uncertain because the site is in a poorly resolved region of the gel and because only one footprinting plot is used in its determination. As discussed above, the binding constant for the CGCT at 102–105 is much better determined from the A-label gel than from the G-label gel, while the reverse is true for the TGCT at 136–139. Aside from these cases, binding constants derived from analyses of the two gels generally agree.

The binding constants are generally higher than what we reported in previous work on this system (Rehfuss et al., 1990b), which analyzed only footprinting data at low drug concentrations, but the order of binding constants is not changed. As shown above, the reason for larger binding constants now is that previous work neglected the weak carrier sites, using a total carrier-site concentration of only 5 μM. Then, the carrier is saturated with drug when c_t gets much past 5 μM, and the free-drug concentration rises rapidly with total drug concentration.

The strong sites, with binding constants greater than 10^6 M^{-1}, all possess the 5′-GC-3′ sequence without a preceding G, and every such sequence within the region of the fragment for which we have footprinting data is a strong binding site, except possibly the 5′-TGCCGG-3′ at sites 159–164 ($K = 4 \times 10^5$ M^{-1}). The binding constant for this last site may appear low because the site is close to the end of the measurable region of the gel, making the experimental data unreliable. More likely, it is low because the sequence is 5′-GCC-3′ on the coding strand and therefore 5′-GGC-3′ (a typical weak binding site) on the noncoding strand.

The highest binding constant we find on the fragment is 6×10^6 M^{-1}, for the sequence 5′-TGCT-3′ at 136-139. However, the site at 62–65, which has the same sequence, has a binding constant of only 4×10^6 M^{-1}. This indicates that bases flanking the tetramer are important in the interaction. Since

the drug can span only 4 base pairs of DNA, it is implied that Act-D affinity may be affected by changes in DNA structure induced by the bases flanking the tetrameric contact sequence. Flanking bases may also influence the binding of Act-D to the sequence CGCT. On the 139-mer, the 5′-GCGC-3′ sequence (sites 101–104) has a binding constant of about 2×10^6 M^{-1} (Table IV), whereas quantitative footprinting studies of Act-D binding to the fragment $d(TAGCGCTA)_2$ (Rehfuss et al., 1990) return a binding constant of about 4×10^6 M^{-1}. The difference may also be the result of end effects associated with short segments of DNA compared to polymeric DNA.

The site in the vicinity of 130 on the fragment, with sequence CCC, is also a relatively strong binding site, with $K = 9 \times 10^5$ M^{-1} (this value is somewhat more uncertain than others because it comes from data for only one cleavage product). Unlike the examples above, it does not possess a core sequence having a C_2 axis of symmetry, like GC, and it is difficult to see how the drug, which has a pseudo-C_2 symmetry axis, interacts with the sequence. In the classical model for binding to GC, the two threonine moieties of the drug hydrogen bond to the 2-amino groups of both guanines at the intercalation site. The fact that CCC is a strong binding site was previously noted in studies with 7-azidoactinomycin D, which is able to photocleave DNA (Rill et al., 1989).

The weak sites, with binding constants between 0.5×10^5 M^{-1} and 9×10^5 M^{-1}, mostly have the 5′-GGC-3′ sequence. Exceptions are the "nonclassical" sequences near 80–84 and 123–126, discussed below. The reason for the weak binding at the GGC sequence is unknown, but it may be related to the enhanced stacking interactions of the GG doublet, which reduces the ability of the drug to intercalate at the adjacent GC site (Ward et al., 1988). Previous footprinting studies (Ward et al., 1988) and studies of cleavage by 7-azidoactinomycin D (Rill et al., 1989) also showed that GC doublets were strongly preferred only if the 5′-flanking base was a pyrimidine and the 3′-flanking base was not cytosine.

The two weak sites near positions 81 and 124 bind drug but do not have a GC core. It is difficult to identify exact interaction sequences from the footprinting data, but the sites both have CCGT and Snyder et al. (1989) have shown that Act-D binds with a high binding constant to the duplex $d(CGTCGACG)_2$. However, the CG sequence at 109–110 shows no tendency to bind Act-D. The sequence CG has 2-fold symmetry like GC and thus might hydrogen bond to appropriate sites on the pseudo-2-fold symmetric drug, but this would involve reorientation of its cyclic pentapeptide rings relative to the phenoxazone chromophore of Act-D if the 2-amino groups of guanine are used as sequence determinants. The relative orientations of the peptide rings of the drug *outside* of DNA are stabilized by specific hydrogen bonds between the rings (Ginell et al., 1988), which might prevent reorientation of the peptide ring in DNA, discouraging binding of the drug to CG. Like CG, the isolated CC sequence (sites 115–116) shows little tendency to bind drug although CCC binds strongly. This suggests that the CGT triplet is required for binding.

Conclusions

In this work, we study the binding of the anticancer drug actinomycin D to a restriction fragment using quantitative footprinting methods. We develop the model which allows calculation of theoretical footprinting plots (intensity of each spot on the footprinting autoradiogram as a function of drug concentration), taking into account equilibria between unbound drug and drug bound to sites on the fragment and on calf thymus DNA carrier, as well as the enhanced cleavage at sites where no drug is bound. A simple model for this mass-action enhancement is given. Parameters such as binding constants of drug to fragment and carrier sites (Table IV) are determined by minimizing the deviation between theoretical and experimental footprinting plots. The binding constants for Act-D sites on a 139-base-pair restriction fragment are determined. The highest affinity sites have the dinucleotide sequence GC, and of these, those having the sequence TGCT have the highest binding constants. Sites having the dinucleotide sequence GC preceded by a G have binding constants in the range $(1–4) \times 10^5$ M^{-1}, about 1 order of magnitude lower than those of the strong sites. Certain sites which do not contain the sequence GC, e.g., CCG and CCC, also bind Act-D with a binding constant comparable to those of the weak GGC sites.

References

Bishop, K. D., Borer, P. N., Huang, Y.-Q. & Lane, M. J. (1991) *Nucleic Acids Res. 19*, 871–875.

Bromley, S. D., Ward, B. W., & Dabrowiak, J. C. (1986) *Nucleic Acids Res. 14*, 1933–1947.

Dabrowiak, J. C., & Goodisman, J. (1989) in *Chemistry and Physics of DNA–Ligand Interactions* (Kallenbach, N. R., Ed.) pp 143–174, Adenine Press, Guilderland, NY.

Dabrowiak, J. C., Skorobogaty, A., Rich, N., Vary, C. P. H., & Vournakis, J. N. (1986) *Nucleic Acids Res. 14*, 489–499.

Dabrowiak, J. C., Goodisman, J., & Kissinger, K. (1990) *Biochemistry 29*, 6139–6145.

Dabrowiak, J. C., Stankus, A., & Goodisman, J. (1991) in *Nucleic Acid Targeted Drug Design* (Propst, C., & Perun, T., Eds.) Marcel Dekker, Inc., New York, NY (in press).

Fletcher, R. (1980) *Practical Methods of Optimization*, John Wiley and Sons, Chichester.

Fox, K. R., & Waring, M. J. (1984) *Nucleic Acids Res. 12*, 9271–9285.

Gale, E. F., Cundliffe, E., Reynolds, P. E., Richmond, M. H., & Waring, M. J. (1981) *The Molecular Basis of Antibiotic Action*, John Wiley and Sons, London.

Ginell, S., Lessinger, L., & Berman, H. M. (1988) *Biopolymers 27*, 843–864.

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* (following paper in this issue).

Huang, Y.-Q., Rehfuss, R. P., LaPlante, S. R., Boudreau, E., Borer, P. N., & Lowe, M. J. (1988) *Nucleic Acids Res. 16*, 11125–11139.

Lane, M. J., Dabrowiak, J. C., & Vournakis, J. N. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3260–3264.

Lown, J. W., Sondhi, S. M., Org, C. W., Skorobogaty, A., Kishikawa, H., & Dabrowiak, J. C. (1986) *Biochemistry 25*, 5111–5117.

Rehfuss, R., Goodisman, J., & Dabrowiak, J. C. (1990a) *Biochemistry 29*, 777–781.

Rehfuss, R., Goodisman, J., & Dabrowiak, J. C. (1990b) in *Molecular Basis of Specificity In Nucleic Acid–Drug Interactions* (Jortner, J., & Pullman, B., Eds.) pp 157–166, Kluwer Academic Publ., Amsterdam.

Rill, R. L., Marsch, G. A., & Graves, D. E. (1989) *J. Biomol. Struct. Dyn. 7*, 591–605.

Scamrov, A. V., & Beabealashvilli, R. Sh. (1983) *FEBS Lett. 164*, 97–101.

Snyder, J. G., Hortmer, N. G., D'Estantoit, B. L., Kennard, O., Remets, D. P., & Breslauer, K. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3968–3972.

Sobell, H. M. (1973) *Prog. Nucleic Acid Res. Mol. Biol. 13*, 153–190.
Suck, D., Lahm, A., & Oefner, C. (1988) *Nature 382*, 465–468.
van Dyke, M. W., Hertzberg, R. P., & Dervan, P. B. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5470–5474.
Ward, B., Rehfuss, R., Goodisman, J., & Dabrowiak, J. C. (1988) *Nucleic Acids Res. 16*, 1359–1369.

Structural Changes and Enhancements in DNase I Footprinting Experiments†

Jerry Goodisman* and James C. Dabrowiak*
Department of Chemistry, Room 1-014, Center for Science and Technology, Syracuse University, Syracuse, New York 13244-4100

Received March 28, 1991; Revised Manuscript Received September 17, 1991

ABSTRACT: In footprinting experiments, an increase in DNA cleavage with addition of ligand to a system may be due to a ligand-induced structural change. Ligand binding also enhances cleavage by displacing the cleavage agent from ligand-binding sites, thus increasing its concentration elsewhere. The theory and characteristics of this mass-action enhancement are given, and it is shown how it may be recognized. Results of DNase I footprinting of small oligomers, with actinomycin D as ligand, are analyzed to reveal which enhancements are due to mass action, and which can reasonably be ascribed to structural changes. Patterns in the footprinting plots from our experiments on actinomycin D binding to a 139-base-pair DNA fragment (with DNase I as a probe) are studied in the same way. The likely origins of these patterns are discussed, as are enhancements occurring with other probes commonly used in footprinting experiments.

The binding of drugs and other ligands may induce structural changes in DNA, which may be detected by a number of techniques. Since the rate of cleavage at a particular bond by agents such as DNase I depends on the local DNA structure (Lomonosoff et al., 1981; Drew, 1984; Suck et al., 1988), a natural way to study such changes is the footprinting technique. In a footprinting experiment, one measures the amounts of DNA fragments of different lengths produced by a cleavage agent, and hence the amount of cutting taking place at various positions on a DNA oligomer, as a function of the ligand concentration (Dabrowiak & Goodisman, 1989; Dabrowiak et al., 1991). The amount of cutting at positions at or near ligand binding sites on DNA will decrease with ligand concentration because the bound ligand prevents the approach of the cleavage agent (inhibition), but, at other positions, one might expect to see changes in cutting rate due to structural changes in the DNA. It should be noted that ligand-induced structural changes may lead to *increases* or *decreases* in the observed cutting rate. Thus, Low et al. (1984) observed large enhancements in cutting by DNase I and DNase II at many sites on a 160-base-pair DNA fragment when echinomycin was allowed to bind, in addition to inhibition of cutting near the drug binding sites (having the sequence CpG). These authors noted two possible explanations for the enhancements: structural changes in DNA and an attractive interaction between the cleaving protein and the antibiotic, leading to increased concentration of the former near antibiotic-binding sites. They were able to dismiss the latter explanation.

However, it is now clear (Ward et al., 1988; Dabrowiak & Goodisman, 1989; Dabrowiak et al., 1991; Portugal, 1989) that there is a third explanation for rate enhancements in DNase I footprinting experiments. Increased cutting at sites where no ligand binds can arise from a mass-action effect, caused by the bound ligand displacing the cleavage agent away from some regions of DNA, and thus increasing the concentration of cleavage agent elsewhere. This means that one may not automatically interpret cleavage rate enhancements as ligand-induced structural changes, since the mass-action effect, due to the equilibrium between DNase I and DNA, is always present. Since mass-action and structural effects may exist simultaneously (Portugal, 1989; Ward et al., 1988), one must always consider whether observed enhancements can be explained by mass action alone or if they are the result of a structural change as well.

Below, we give a model for the mass-action enhancement and discuss how one can judge whether enhancements observed in a footprinting experiment have a structural origin. Then, we consider observed enhancements on small DNA oligomers and longer fragments, previously noted by other workers, which may be due to structural changes or mass-action effects. The experimentally observed intensity enhancements for actinomycin D interacting with a 139-mer that we believe to be structural in origin are then presented and discussed.

Experimental Procedures

The quantitative footprinting studies involving actinomycin D, the 139-base-pair *Hin*dIII/*Nci*I restriction fragment of pBR 322 DNA, and DNase I were as earlier described (Ward et al., 1988). The sequence of the restriction fragment and the location of strong and weak actinomycin D binding sites are shown in Figure 1. Autoradiographic spot intensities, corresponding to relative amounts of 54 cleavage products of different lengths, were measured for 26 actinomycin D concentrations ranging from 0 to 38.8 μM.

The model used to interpret the resulting footprinting plots, plots of spot intensity as a function of total drug concentration, is described in detail elsewhere (Goodisman et al., 1992). The analysis takes into account binding of drug at the various binding sites, the mass-action enhancement, and binding of the drug to unlabeled carrier DNA (calf thymus). Correct description of the carrier, described as a concentration of strong binding sites and a concentration of weak binding sites, is important, since it is mainly the equilibria between drug and

† We ackowledge the American Cancer Society, Grant NP-681, for supporting this research.

0006-2960/92/0431-1058$03.00/0