Binding Characteristics of Hoechst 33258 with Calf Thymus DNA, Poly[d(A-T)], and d(CCGGAATTCCGG): Multiple Stoichiometries and Determination of Tight Binding with a Wide Spectrum of Site Affinities[†]

Frank G. Loontiens,[‡] Peter Regenfuss,[§] Annelies Zechel,[§] Lieve Dumortier,[‡] and Robert M. Clegg*,[§]

Laboratory of Biochemistry, Faculty of Sciences, Gent State University, Ledeganckstraat 35, B-9000 Gent, Belgium, and Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, D-3400 Göttingen, FRG

Received October 5, 1989; Revised Manuscript Received May 1, 1990

ABSTRACT: Equilibrium binding experiments using fluorescence and absorption techniques have been performed throughout a wide concentration range (1 nM to 30 μM) of the dye Hoechst 33258 and several DNAs. The most stable complexes found with calf thymus DNA, poly[d(A-T)], d(CCGGAATTCCGG), and d(CGCGAATTCGCG) all have dissociation constants in the range $(1–3) \times 10^{-9}$ M^{-1}. Such complexes on calf thymus DNA occur with a frequency of about 1 binding site per 100 base pairs, and evidence is presented indicating a spectrum of sequence-dependent affinities with dissociation constants extending into the micromolar range. In addition to these sequence-specific binding sites on the DNA, the continuous-variation method of Job reveals distinct stoichiometries of dye–poly[d(A-T)] complexes corresponding to 1, 2, 3, 4, and 6 dyes per 5 A-T base pairs and even up to 1 and 2 (and possibly more) dyes per backbone phosphate. Models are suggested to account for these stoichiometries. With poly[d(G-C)] the stoichiometries are 1–2 dyes per 5 G-C pairs in addition to 1 and 2 dyes per backbone phosphate. Thermodynamic parameters for formation of the tightest binding complex between Hoechst 33258 and poly[d(A-T)] or d-(CCGGAATTCCGG) are determined. Hoechst 33258 binding to calf thymus DNA, chicken erythrocyte DNA, and poly[d(A-T)] exhibits an ionic strength dependence similar to that expected for a singly-charged positive ion. This ionic strength dependence remains unchanged in the presence of 25% ethanol, which decreases the affinity by 2 orders of magnitude. In addition, due to its strong binding, Hoechst 33258 easily displaces several intercalators from DNA.

The dye Hoechst 33258 [a bis(benzimidazole) derivative, Figure 5] becomes brightly fluorescent when bound to A-T-rich sequences of double-stranded DNA. It produces a bright microscope image of chromosomal DNA with high contrast due to minimal interference from dye bound to other cell components. It is useful for visualization of newly synthesized DNA, since the fluorescence of Hoechst 33258 is quenched by 5-bromouridine which has been incorporated into chromosomes. These fluorescence properties have been exploited to determine the amount of A-T base pairs in DNA samples (Weisblum & Haenssler, 1974; Commings, 1975), and binding to these sequences has been termed "specific". The absorption and circular dichroism spectra of Hoechst 33258 complexes with A-T-rich DNAs differ from those lacking the A-T sequences (Latt & Stetten, 1976; Steiner & Sternberg, 1979; Mikhailov et al., 1981), and this difference has led to the terminology of "unspecific" binding. Unspecific binding of Hoechst 33258 at high concentrations is associated with a quenching of the specifically bound dye fluorescence (Latt & Wohlleb, 1975; Latt & Stetten, 1976; Mikhailov et al., 1981; Stokke & Steen, 1985). It has been reported that increased ionic strength dissociates unspecifically bound Hoechst 33258 but has no effect on the specifically bound complexes (Latt & Wohlleb, 1975; Latt & Stetten, 1976). Relatively low concentrations of ethanol abolish unspecific binding (Stokke & Steen, 1987). A minimum of three consecutive A-T pairs in a group of four B-DNA base pairs has been proposed to

constitute a specific binding site in the minor groove (Zimmer & Waehnert, 1986); the presence of a 2-amino group of guanine precludes binding of the dye (Teng et al., 1988). X-ray diffraction data on the complex of Hoechst 33258 and the self-complementary dodecamer d(CGCGAATTCGCG) presented by Pjura et al. (1987) and Teng et al. (1988) show that the dye fits snugly in the minor groove of B-DNA. Surprisingly, both of these research groups propose different assignments (shifted by one base pair) for the binding contacts. The footprinting technique (Martin & Holmes, 1983; Harshman & Dervan, 1985; Jorgenson et al., 1988; Portugal & Waring, 1988) indicates a complementarity of the dye with 5 base pairs (including 4 A-T pairs, 3 of which are consecutive) when low ratios of dye per base pair are used; at higher dye to base pair ratios a complementarity with 6 base pairs is found.

We have investigated the binding properties of Hoechst 33258 in solution with and without 25% ethanol over a wide range of concentrations, dye/DNA ratios, and ionic strength, and as a function of temperature with calf thymus DNA, chicken erythrocyte DNA, and poly[d(A-T)]; some of our studies include poly[d(G-C)], d(CCGGAATTCCGG), and d(CGCGAATTCGCG). A detailed investigation of the equilibrium binding parameters was necessary because our preliminary fast-kinetic measurements of Hoechst 33258 binding to DNA (Regenfuss et al., 1989) indicated that the previously published equilibrium binding data could not be completely correct; indeed, apparent affinities reported earlier vary considerably and differ from our data by more than 1 order of magnitude.

Two methods were used to determine the binding parameters: (1) a series of titrations with different but constant dye

concentrations were carried out to determine the dissociation constant, K_{diss}, and the number of base pairs per site; (2) continuous variation of the dye/DNA ratio at high concentration to determine the stoichiometry (Job, 1928; Cantor & Schimmel, 1980; Huang, 1982). An important finding of this study is the existence of sites on calf thymus DNA that have K_{diss} values of the order of 1 nM, much lower than previously supposed; this is confirmed by similar K_{diss} values for the two dodecamers with a central AATT sequence and also for poly[d(A-T)]. Also, contrary to the prevalent findings in the literature, we find the strong and sequence-selective binding to be dependent upon ionic strength, as expected for a monocationic ligand interacting in the minor groove with a favorable charge potential (Pullmann, 1985).

MATERIALS AND METHODS

The buffers were either 50 mM Tris–100 mM NaCl (pH 7.5) or 20 mM Bis-Tris–100 mM NaCl (pH 7.0); in some cases the NaCl concentration was varied, or 25% ethanol (w/v) was present. Concentrations were determined spectrophotometrically; nucleic acid concentrations were expressed as base pairs with ϵ values (M^{-1} cm^{-1}) as indicated: calf thymus DNA (sodium salt, type V, Sigma) and chicken erythrocyte DNA [prepared from nucleosomes according to Simon and Felsenfeld (1979)], ϵ_{258} = 13 600; poly[d(A-T)] (Boehringer), ϵ_{262} = 13 200; poly[d(G-C)] (Pharmacia), ϵ_{256} = 16 800; d-(CCGGAATTCCGG) (Serva) and d(CGCGAATTCGCG), ϵ_{260} = 15 270, an estimate calculated from the values for poly[d(A-T)] and poly[d(G-C)]. In some Job plots, nucleic acid concentrations were expressed as phosphate. The following ϵ values (M^{-1} cm^{-1}) were used to determine concentrations of the dyes: Hoechst 33258 (gift from Dr. Loewe, Frankfurt, FRG), ϵ_{338} = 42 000 at pH 7 for absorbance readings between 0.7 and 0.75 in a 1-cm cuvette (it is slightly concentration dependent); ethidium bromide (Serva), ϵ_{480} = 5850; quinacrine (Sigma), ϵ_{422} = 9570; proflavin (BDH), ϵ_{444} = 3920; actinomycin D (Sigma), ϵ_{442} = 25 200; adriamycin and daunomycin [Farmitalia (Carlo Erba)], ϵ_{480} = 11 500.

Absorption and difference absorption spectra were measured on a Uvikon (Kontron) 820 or 810 instrument with a bandwidth of 1 or 2 nm. Many fluorescence measurements were made on an SLM 8000S (Urbana, IL) instrument. Excitation and emission spectra were corrected. For all titrations in the SLM λ_{ex} = 355 nm and λ_{em} = 465 nm; the fluorescence readings were corrected for lamp fluctuations. Polarization artifacts were avoided by employing "magic angle" conditions (i.e., the angle of the linear polarizer relative to the vertical position is 54.5° for excitation and 0° for emission). The fluorescence titrations with ~1 nM dye were done under constant stirring on either an Aminco SPF 500 instrument or a sensitive fluorescence instrument constructed in the laboratory with a chopped light source, lock-in phase detection, and a stable tungsten lamp. In the latter case fluorescence of >470 nm was selected with a KV-470 cutoff filter. The fluorescence signal was digitized and processed in a LSI DEC computer; 30-s averaging was used.

The adsorption of Hoechst 33258 from neutral aqueous solution onto glass, quartz, and some plastic (Parafilm) surfaces was prevented as follows: (1) Working solutions of the dye (>125 nM) were prepared daily by dilutions of a concentrated stock solution (1–5 mg in water) with 5 mM HCl. This provided a stable environment for the free dye for at least 24 h. The dye was then diluted from a 5 mM HCl solution into a solution with sufficient buffering capacity (pH 7 or 7.5); (2) Glass and quartz cuvettes were silanized with dimethyldichlorosilane (2% in 1,1,1-trichloroethane from BDH) to

minimize adsorption. The dye does not adsorb to polystyrene or silanized quartz cuvettes. (3) The cuvette content was stirred with a silanized glass stirrer. Neutral-pH solutions of the dye were kept in polystyrene tubes. Neutral dye solutions with 25% ethanol do not adsorb to surfaces.

Fluorescence titrations were done in two ways: (1) Solutions with 10 nM or greater dye concentrations were individually prepared; the low concentrations of the dye were mixed immediately with the DNA. (2) To minimize background instabilities in the fluorescence measurements with 1 nM dye, the cuvette was not removed from its holder during consecutive additions of the titrant. The DNA solution was added from a micrometer-actuated syringe, with constant stirring by a motor-driven silanized glass stirrer. All measurements were corrected for a blank titration without dye. Difference absorption titrations were obtained similarly, using microburets and stirrers for the two cuvettes.

Job's continuous-variation method (Job, 1928; Cantor & Schimmel, 1980; Huang, 1982) was applied to determine binding stoichiometries. The concentration of both reactants is varied, while the sum of their concentrations is constant. A signal change, absorption or fluorescence, related to the amount of complex formed is plotted against the input mole fraction of the dye (χ). Ideally, at concentrations much higher than K_{diss} inflection points yield the stoichiometry of the complexes formed. The method has seldom been applied to the binding of ligands to natural or synthetic DNAs (Chou et al., 1987). We have verified that Job's continuous-variation method can be applied to the statistically excluded binding (McGhee & von Hippel, 1974; Zasetatelev et al., 1971) of a multivalent ligand to a repeating linear polymer with independent and noninteracting elements. Simulations of statistically excluded binding show that plot maxima and minima are less acute and shift less toward the center of the plot upon dilution of the reactants than those of normal bimolecular binding. Job plots with ethidium bromide and DAPI (Loontiens et al., 1989a) can be fitted well if simulated according to the excluded-site binding model. The entire curve of the Job plot for the excluded-site model can be fitted to determine the true value of the mole fraction at equivalence (χ_t). In practice, our χ values at the crossing of the neighboring slopes and at the experimental maximum are very close, and their average (χ_e) is almost identical with χ_t. The nucleic acid concentrations were expressed either as phosphate or as base pairs. For instance, for 1 dye per 5 base pairs and 1 dye per phosphate, the χ_t values are 0.091 and 0.5, with concentrations expressed as phosphate; with concentrations expressed as base pairs, χ_t = 0.167 and 0.667. Concentrated stock solutions (equal concentrations) of the two reagents were added from microburets (total added volumes of 45 μL) into buffer (1.00 mL) in fresh 1 × 0.4 cm polystyrene cuvettes; mixing was done with a motor-driven silanized stirrer. Absorbance readings were made consecutively on the same cuvette (which was never removed from the cell holder), first filled with buffer, then after addition of the dye, and finally with the dye–DNA mixture. Differences due to binding were obtained after correction for the contribution by the free dye for dilution by DNA. Job plots measuring fluorescence were obtained in a similar way.

RESULTS

Evidence for the Aggregation of the Free Dye in Neutral Solution. The ϵ value of Hoechst 33258 at pH 7.0 or 7.5 near the absorption maximum (338 nm) decreases with dye concentrations up to about 30 μM, whereas ϵ increases in the 370–400-nm region. However, this concentration-dependent effect of the free dye at pH 7.0 is absent when 25% ethanol
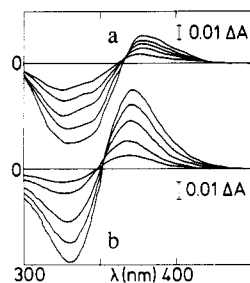
FIGURE 1: (a) Concentration-induced difference absorption spectra of the dye free in solution at pH 7.0 (24 °C) with [dye] of 6.0, 10.2, 14.3, 19.4, and 24.2 μM in the 1-cm reference cuvette; the concentration in the 0.5-cm sample cuvette was higher by a factor of 2.000. (b) Difference absorption spectra for 25 μM calf thymus DNA binding with increasing dye concentrations corresponding to dye/base pair ratios of 0.019, 0.038, 0.075, 0.113, and 0.151. In (a) and (b), the red shifts of the minima, maxima, and near-isosbestic points with concentration indicate the complexity of the equilibria between multiple forms both of (a) free dye and (b) of the DNA-bound dye at low dye/base pair ratios.

is added to the solution or when the dye is dissolved in 5 mM HCl. At pH 7.0, the concentration-induced difference absorption spectra (Figure 1a) also shift to the red as the concentration of the dye increases (6–24 μM), resulting in maxima at 376 to 380 nm, minima near 330 nm, and a near-isosbestic point shifting from 362 to 366 nm. The concentration dependence of the signal change is complex and probably results from dye aggregates higher than dimers. These wavelengths shifts of the free dye are similar to shifts in the difference absorption spectra upon binding of the dye to DNA (Figure 1b). At pH 7 in aqueous buffer we measure a fluorescence anisotropy of a freshly prepared solution of dye to be 0.17; this is independent of time [in contrast to Stokke and Steen (1985)] or of concentration from 0.1 to 10 μM.

Difference Absorption Spectra of the Dye upon Binding to DNA and Competition with Intercalators. When Hoechst 33258 binds to 25 μM calf thymus DNA at dye/base pair values smaller than 0.15, the difference absorption spectra indicate a deviation from simple binding. This follows from a small but reproducible and systematic red shift in the difference absorption spectra of Figure 1b, noticeable at the maxima (368–374 nm), at the minima (326–352 nm), and at the crossings at the base line (347–352 nm). These spectroscopic changes, although complex, allow a convenient monitoring of Hoechst 33258 in the presence of other molecules, dyes, or nonchromophoric ligands (both intercalators and peripherally bound ligands) that interact with DNA. Hoechst 33258 offers some advantages over methyl green (Krey & Hahn, 1975) or ethidium bromide (Baguley, 1982) in this respect due to the lack of spectral interference and high affinity of binding. Even with almost maximal occupancy of the DNA lattice with ethidium bromide, Hoechst 33258 binding at dye/base pair ratios smaller than 0.08 can be followed (Figure 2b) at the isosbestic point for ethidium bromide (390 nm). The number of dissociated ethidium bromide molecules is almost equal to the number of bound Hoechst 33258 molecules, and the bound Hoechst 33258 molecules are decreased by less than a factor of 2. Thus, Hoechst 33258 binds more tightly to DNA than ethidium bromide. Fluorescence energy transfer from Hoechst 33258 to ethidium bromide is observed, confirming the data by Langlois and Jensen (1979).

Other intercalators such as quinacrine, acridine orange, daunomycin, adriamycin, and especially proflavin are released from DNA by Hoechst 33258 at low dye/base pair values. The estimated quantity of released intercalator often differs with the wavelength (e.g., to the left or right of the isosbestic
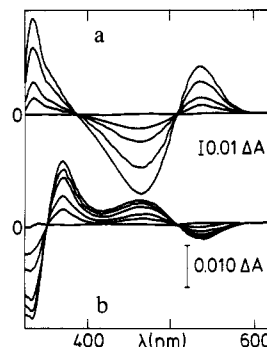

FIGURE 2: Release of intercalated ethidium bromide from calf thymus DNA by Hoechst 33258. (a) Binding of ethidium bromide: the base line in (a) was set with 81.2 μM ethidium bromide (2.022 mL) in the sample and reference cuvettes; aliquots of calf thymus DNA were added to the sample cuvette, and an equal volume of buffer to the reference cuvette. The successive additions produced concentrations of DNA equal to 6.25, 12.5, 24.8, and 37.2 μM. (b) Hoechst 33258 releases ethidium bromide: after the total addition in (a) the total volume was 2.048 mL, and the ethidium bromide/DNA (bp) ratio was 2.16; the base line was adjusted to zero, and aliquots of concentrated Hoechst 33258 were added to the sample and reference cuvettes. The concentrations of Hoechst 33258 in (b) were 0.47, 0.94, 1.86, 2.31, and 2.76 μM with Hoechst 33258 to base pair ratios from 0.0126 to 0.074.

point); this indicates the tendency of many of the intercalators to aggregate.

Terminology for Binding Modes Observed with Job's Continuous-Variation Method. It has long been suggested (Steiner & Sternberg, 1979; Mikhailov et al., 1981; Stokke & Steen, 1985) that there are two types of binding for Hoechst 33258 and DNA: a strong and A-T-"specific" binding with an intense increase in fluorescence ("type 1" binding) and, at high dye/base pair ratios, a weaker and "unspecific" binding ("type 2") related to the quenching of the dye fluorescence. *Several* complexes with distinct stoichiometries are formed between Hoechst 33258 and a DNA helix (Figures 3 and 4); therefore, the terms specific and unspecific binding are no longer adequate. We propose the following terminology for four classes of distinct interactions between DNA structural units and the dye: (1) sequence mediated, (2) charge mediated, (3) dye mediated, and (4) structure mediated. This terminology anticipates our models (in Table I and Figure 5) and facilitates their presentation and discussion.

Sequence-mediated (specific) binding refers to the complex from X-ray diffraction (Pjura et al., 1987; Teng et al., 1988) and footprinting (Harshman & Dervan, 1985; Portugal & Waring, 1988) data. The binding site is in the minor groove and requires a specific sequence (often 4 A-T base pairs), and the complex is highly fluorescent. It is the most affine (specific) interaction with DNA. We find a broad range of affinities for different sequences on calf thymus DNA. Charge-mediated binding (unspecific) is the primary mode of interaction between the monocationic Hoechst 33258 molecule and the charged backbone phosphate; this produces a stoichiometric ratio of dye/phosphate = 1:1 (see Figure 5). The electrostatic interactions play a role in all associations involving ions, especially in minor groove binding. Dye-mediated binding (also unspecific) occurs whenever a free dye molecule binds to another dye molecule, or a group of these dye molecules, already bound to the DNA, whether these dye molecules are already bound within a groove or to a backbone phosphate (charge mediated). Dye-mediated binding could involve both edge-to-edge contacts and stacking interactions between dyes and is common to all DNAs. Structure-mediated binding (unspecific or specific) confines the dye to a structural region
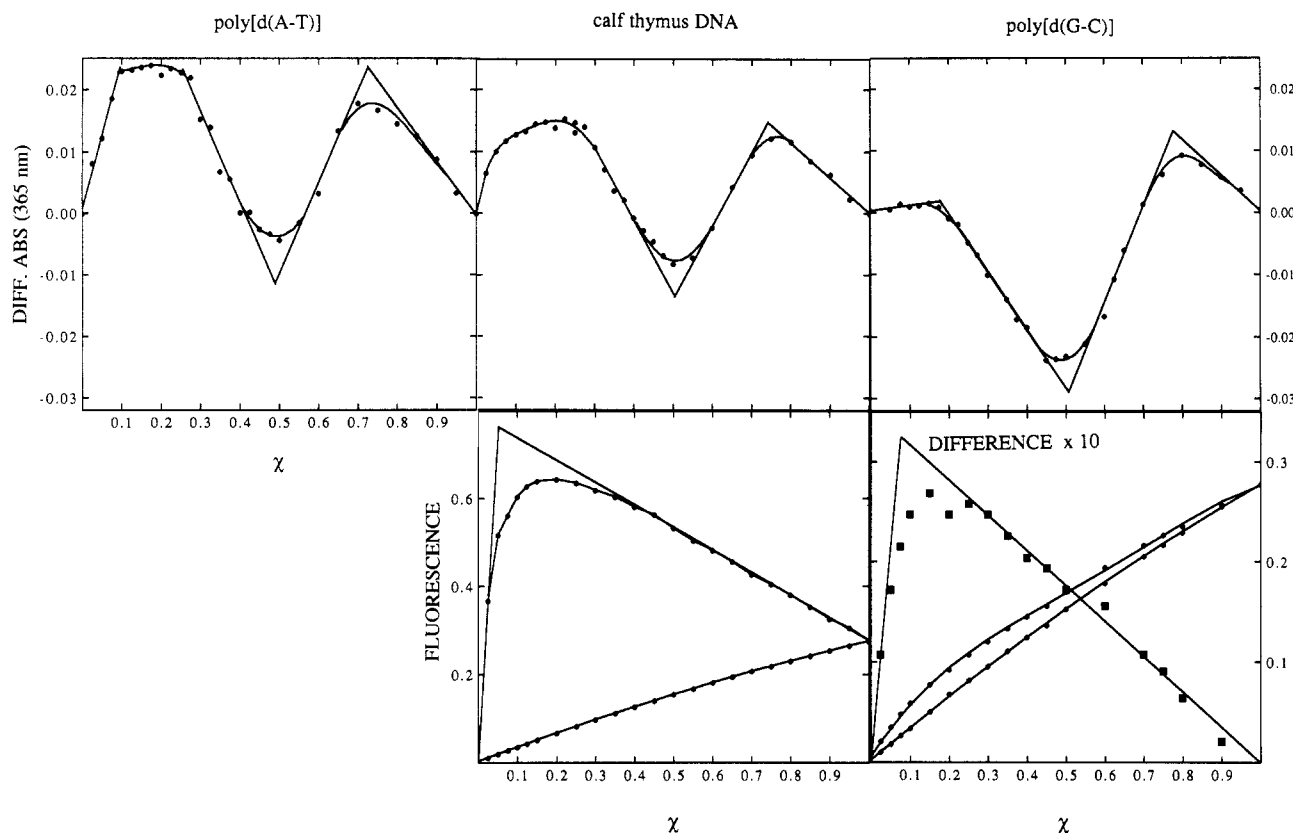
poly[d(A-T)] calf thymus DNA poly[d(G-C)]

FIGURE 3: Job plots for Hoechst 33258 binding to poly[d(A-T)], calf thymus DNA, and poly[d(G-C)]; [DNA] expressed as phosphate; T = 23 °C and pH 7.5 (50 mM Tris, 100 mM NaCl). In the three top plots, the Δabs was measured at 365 nm for (from left to right) poly[d(A-T)], calf thymus DNA, and poly[d(G-C)]; [dye] + [DNA] was 22.6 μM; χ is the mole fraction of dye. In the two bottom figures, fluorescence was measured for the dye alone (bottom curves) and for [dye] + [DNA] (calf thymus DNA and poly[d(G-C)]); the sum of the concentrations was 1.1 μM. The fluorescence plot for poly[d(G-C)] is enlarged by 10×.
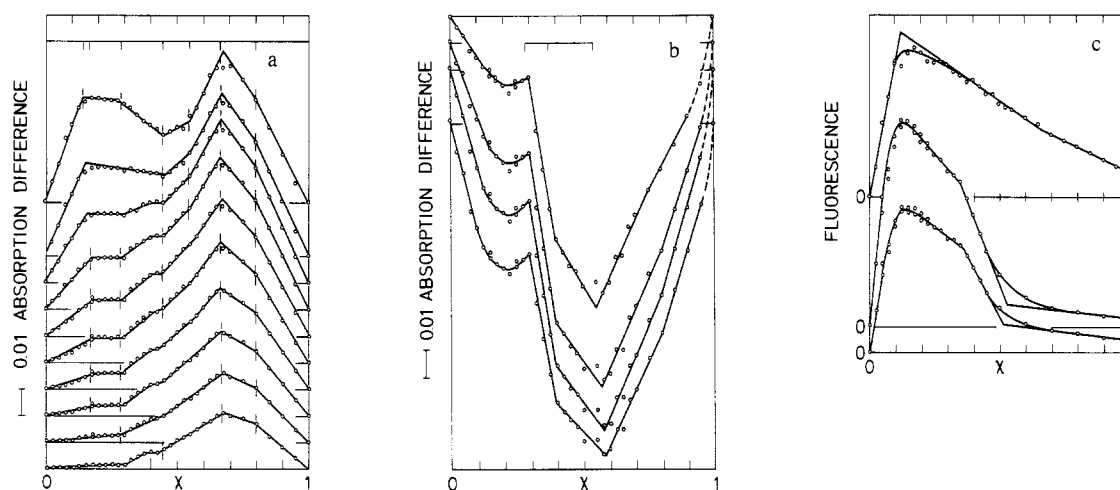

FIGURE 4: Job plots for Hoechst 33258 binding to poly[d(A-T)] (bp) at 23 °C, 50 mM Tris, pH 7.5, and 100 mM NaCl. Δabs was determined at 5-nm intervals (a) between 375 (top) and 420 nm (bottom) and (b) between 330 (bottom) and 345 nm (top) with [dye] + [DNA] = 30.2 μM. The nonlinear scale as an inset is for χ_t = 0.143, 0.167, 0.285, 0.444, 0.545, 0.667, and 0.8 in (a) and for χ_t = 0.285, 0.375, and 0.545 in (b). In (c), fluorescence was measured; for the top curve [dye] + [DNA] = 1.4 μM with λ_{ex} = 360 nm and λ_{em} = 460 nm, and the two corresponding curves at the bottom were obtained at 15.1 μM with λ_{ex} = 405 nm and λ_{em} = 460 or 500 nm; the latter wavelength yielded the smaller signal. Fluorescence inflection points are close to χ_t = 0.143, 0.375, and 0.545.

(e.g., a groove of the helix) by some geometrical requirement; this may be responsible for the binding of one Hoechst 33258 molecule to 5 G-C pairs in poly[d(G-C)].

These four classes of binding interactions described above are not distinctly separate entities, since more than one type of interaction contributes to the affinity of each bound dye. Our experiments provide only the composition, not the detailed

structures, corresponding to the multiple stoichiometries. However, the schematic presentations in Figure 5 are suggested by the series of χ_e values we have obtained and offer a testable view of the interactions.

Common Charge-Mediated and Dye-Mediated Stoichiometries Found by Difference Absorption. The three top plots in Figure 3 have a sequence-independent minimum close to

Table I: Stoichiometries for Interaction of Hoechst 33258 and Poly[d(A-T)][a]

| dye + BP | χ_t | χ_e found in | | | | binding mode |
| | | fluorescence | | difference absorption | | |
		1.4 μM	15 μM[b]	30 μM	nm range	
1 + 5	0.167	0.12–0.15	0.12–0.16	0.17	390–410	sequence mediated
or	or		0.10–0.15	0.14–0.17	385–370	
1 + 6	0.143			0.14	375–380	
2 + 5	0.285	NO	NO	0.30	350–360; 375	dye and structure mediated
				0.28–0.30	415–390	
3 + 5	0.375	NO	0.38	0.38	345–330	dye and structure mediated
			0.35–0.36			+little quenching
4 + 5	0.444	NO	NO	0.45	375–395	dye and structure mediated
5 + 5	0.500	NO	NO	wide	370–350	
6 + 5	0.545	uncertain	0.55	0.55	375–390	dye and structure mediated
			0.52–0.53	wide	370–350	+strong quenching
1 + 0.5	0.667 (a)	uncertain	NO	0.67	375–420	charge mediated
				0.67–0.75	350–370	
2 + 0.5	0.800 (b)	NO	NO	0.8	375–420	dye mediated

[a] χ_t and χ_e are the theoretical and experimental mole fractions of dye: (a) 0.5 or (b) 0.667 for 1 (a) or 2 (b) dyes + phosphate. NO, not observed. Each concentration is the sum of poly[d(A-T)], in base pairs, plus dye and is constant in each series. [b] Two independent determinations.
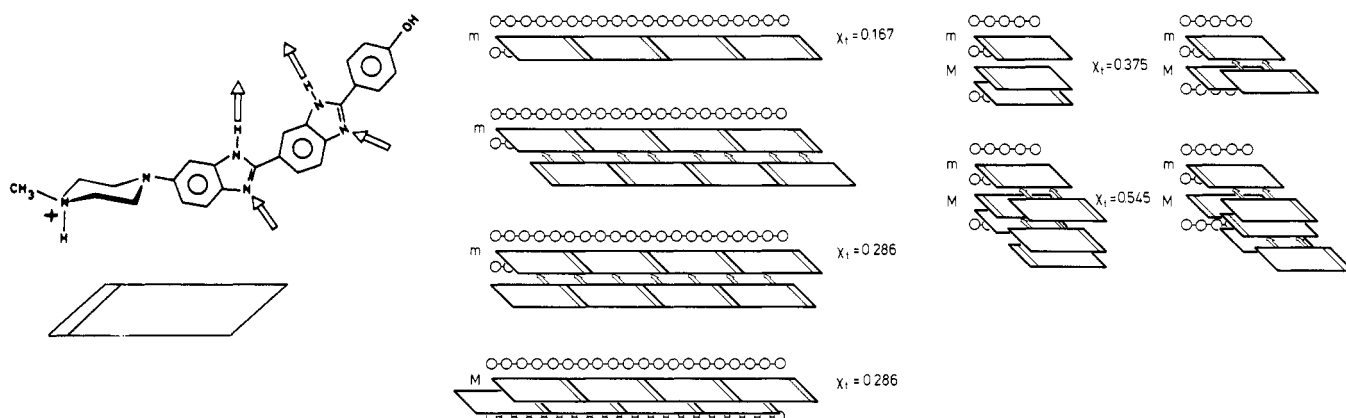

FIGURE 5: (Left) Structure of Hoechst 33258 at neutral pH. The two top arrows indicate H bonds toward O2 of T and N3 of A in the minor groove, and the two bottom arrows indicate possible H bonds from a second dye molecule forming an edge-to-edge-dimer in head-to-tail orientation. The dye is shown as an oblong rectangle in the center and right panels. At pH 7.0–7.5, only the piperazine methyl N has a positive charge. (Center) Sequence-mediated binding and structure-mediated binding in poly[d(A-T)] and poly[d(G-C)] with χ_t = 0.167 and 0.286; the extended ribbons of juxtaposed dyes are aligned in the minor (m) and major (M) grooves, assuming a 5 base pair complementarity represented for two helical turns of DNA. Circles represent the backbone phosphates. The helical grooves and phosphate backbones are represented as linear entities for simplicity. With sequence-mediated binding, the dye forms a ribbon in the minor groove at a mole fraction of dye χ_t = 0.167, which is the fluorescent complex. Dye-mediated binding forms a "dimer" with χ_t = 0.286 onto this ribbon in poly[d(A-T)], probably by hydrogen bonds. The two middle schemes represent two possible arrangements. Structure-mediated binding is represented at the bottom as dye dimers stacked as tiles in the major groove, also at χ_t = 0.286; this arrangement could occur with poly[d(A-T)] and most probably with poly[d(G-C)]. (Right) Possible arrangements of complexes involved in the Hoechst 33258 fluorescence quenching at high dye/base pair ratios (χ_t = 0.375 and 0.545). A complex with 3 dyes per 5 A-T pairs in poly[d(A-T)] with χ_t = 0.375 causes a minor decrease in fluorescence (see Figure 4c, bottom). Two possible arrangements are represented as local entities rather than as extended ribbons; there is either a single dye in the minor groove and a stacked dimer in the major groove (left) or a hydrogen-bonded edge-to-edge dimer at the minor groove and a single dye per 5 base pairs in the major groove (right). Each of these arrangements can form a further complex of 6 dyes per 5 A-T pairs (χ_t = 0.545) responsible for the nearly total fluorescence quenching. Two possible arrangements involve three edge-to-edge dimers: (1) one dimer in the minor groove and two dimers stacked in the major groove (left); (2) one in the minor groove, one in the major groove, and a third dimer partially sandwiched in between (right).

χ_t = 0.5 (meaning a 1:1 dye–phosphate complex), suggesting charge neutralization. Dye-mediated binding through phosphate-bound dye is a second common feature in Figure 3; these maxima are between the experimental mole fractions of dye, χ_e = 0.7 and 0.8, and close to or encompassing χ_t = 0.67, 0.75, and 0.8 for 2, 3, and 4 dyes per phosphate; these values suggest additional stoichiometric aggregation through the dye bound at one backbone phosphate. The neutralized dye–DNA aggregates are insoluble (Commings, 1975), and at wavelengths larger than 400 nm, where dye absorption becomes negligible, turbidity contributes the major signal (bottom plots in Figure 4a). This dye-mediated binding or aggregation on DNA probably results from the propensity of the dye to aggregate in neutral solution. Dye-mediated aggregation, initiated by

a tightly bound (sequence mediated) dye molecule in the minor groove of poly[d(A-T)], is described below.

Sequence- and Structure-Mediated Binding from Fluorescence Job Plots. The fluorescence plot in Figure 4c (top) suggests a stoichiometry for sequence-mediated binding of 1 dye per 6 base pairs (χ_t = 0.143 for poly[d(A-T)], between the experimental maximum at χ = 0.16 and the crossing of the slopes at χ = 0.12). A stoichiometry of 1:5, with χ_t = 0.167, cannot be excluded. The deviation from linearity in Figure 4c (top) between χ = 0.5 and χ = 0.75 was examined further at higher concentrations (see below and Figure 4c, two bottom graphs). The simple fluorescence Job plot for calf thymus DNA in Figure 3 (bottom) estimates the stoichiometry as 1 dye per 10 base pairs, which is lower than 1 per 6 for
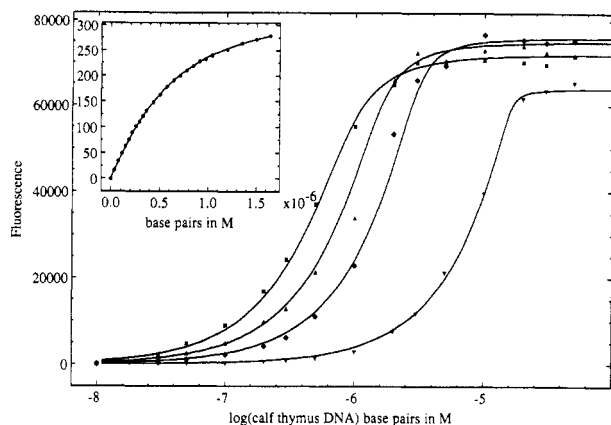
FIGURE 6: Titrations of different but constant concentrations of Hoechst 33258 with calf thymus DNA at pH 7. Each titration is simulated with the independent-site model (see Appendix) to give the number of base pairs per site, N. From left to right: (*) 10 nM dye, $N = 80$; (△) 30 nM dye, $N = 50$ (◆) 100 nM dye, $N = 30$; (▽) 1000 nM dye, $N = 16$. All simulations are done with $K_{diss} = 10^{-7}$ (M base pairs)$^{-1}$, since this is the K_{diss} for 10 nM dye and the simulations for [dye] > 10 nM are independent of the exact value of K_{diss}. The characteristics of a tight binding curve (narrowing of the width and a sharper turning at the upper right corner) increase from left to right as the titrated concentration of the dye increases; simultaneously, the effective number of base pairs per site N decreases. 1.5 nM dye is shown in the insert with a linear scale of added sonicated DNA. N is not reliably determined for 1.5 nM dye (see Appendix); if we assume $N = 80$ base pairs, $K_{diss} = 1.7 \times 10^{-9}$ M^{-1}.

poly[d(A-T)]. This difference between poly[d(A-T)] and calf thymus DNA is also evident in the difference absorption plots in Figure 3. It arises from the lower frequency of potential binding sites in calf thymus DNA (see the fluorescence titrations below).

Structure-Mediated Binding with Poly[d(G-C)]. The fluorescence increase for binding of Hoechst 33258 to poly-[d(G-C)] is very small but is not mediated by a 1:1 dye-phosphate complex since the corresponding Job plot (bottom right plot in Figure 3, 1.1 μM in phosphate) shows a distinct interaction with χ_e between $\chi_t = 0.091$ and $\chi_t = 0.167$ for 1 or possibly 2 dyes per 5 G-C base pairs. The slopes cross at $\chi = 0.08$ (close to $\chi_t = 0.091$ or 1 dye per 5 base pairs), comparable to the sequence- or structure-mediated binding observed with poly[d(A-T)]. Independent evidence is provided by the difference absorption Job plot in Figure 3 (top right) with a sharp inflection point, now with $\chi_e = 0.18$ near to $\chi_t = 0.167$ or 2 dyes per 5 G-C pairs. This inflection point and the other two at $\chi_e = 0.5$ and $\chi_e = 0.75$ (representing charge- and dye-mediated binding) were always observed with poly-[d(G-C)] as wavelength-independent and unique inflection points between 340 and 420 nm; this contrasts with poly[d-(A-T)] (Figure 4a).

Fluorescence Quenching and Dye-Mediated Binding as a Consequence of Sequence-Mediated Binding with Poly[d(A-T)]. Upon increasing of the total concentrations of Hoechst 33258 and poly[d(A-T)] from 1.4 to 15 μM in the fluorescence Job plots in Figure 4c, two additional (Table I) inflection points at $\chi_e = 0.36$ (shoulder) and $\chi_e = 0.52$ (strongly quenched fluorescence) appear, suggesting that 3 ($\chi_t = 0.375$) and especially 6 ($\chi_t = 0.545$) dyes per 5 base pairs cause the quenching. These two values exclude that quenching of dye fluorescence results primarily from charge-mediated binding to the phosphates, which would require a stoichiometry at $\chi_e = 0.67$. The data by Stokke and Steen (1985) for 50% quenching and 50% light scattering confirm this; indeed, when expressed in terms of base pairs and χ, their equivalents are
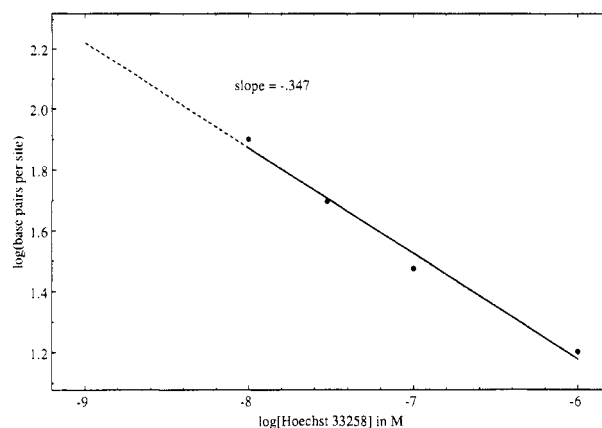

FIGURE 7: log N (the number of base pairs per site in the independent-site model) for calf thymus DNA (Figure 6) vs log [Hoechst 33258]. The linear regression is extrapolated to 1 nM Hoechst 33258, where N cannot be reliably determined. This linear log–log relation describes the data well and allows for an extrapolation to the very low [dye] region.

Table II: Comparison of K_{assoc} (25 °C) with DNA Concentrations Expressed in Base Pairs and Per Site

| | K_{assoc} | |
	molar base pairs	molar sites
calf thymus DNA	7×10^6	$6 \times 10^{8\,a}$
poly[d(A-T)]	3.8×10^7	$3.8 \times 10^{8\,b}$
d(CCGGAATTCCGG)	2.6×10^7	$3.2 \times 10^{8\,c}$
d(CGCGAATTCGCG)	3×10^7	$4 \times 10^{8\,c}$

[a] Independent-site model, $N = 80$. [b] Independent-site model, $N = 10$ (for $n = 5$–6). [c] Independent-site model, $N = 12$ (or per duplex).

$\chi_e = 0.5$, close to our $\chi_e = 0.54$, and $\chi_e = 0.67$, respectively (Table I). The stoichiometries of 3 and 6 are confirmed by the same two values of χ_e (see Table I) for the shoulder and the pronounced minimum at the shorter wavelengths in the difference absorption Job plot at 30 μM (Figure 4b). The most complex Job plots were obtained by difference absorption at the longer wavelengths (Figure 4a). Within the 375–420-nm interval, there are at least four inflection points with χ_e smaller than 0.6: these correspond to 1, 2, 4, and 6 dyes per 5 base pairs. In addition, the right-hand portion of Figure 4a shows two inflection points corresponding to 1 and 2 dyes per phosphate.

Fluorescence Titrations at Very Low Concentrations: Two Binding Models. To interpret the binding isotherms, we apply two models: (1) the independent-site model and (2) the excluded-site model. These models and the different meaning of the number of base pairs, N and n, per site are given in the Appendix. It is assumed that the dye–DNA complex has constant fluorescence properties throughout the titration and that the concentrations and the dye/DNA ratios are sufficiently low so as not to quench the fluorescent species.

Fluorescence Titrations with Calf Thymus DNA and Independent-Site Model for Natural-Sequence DNA. Titrations of different dye concentrations (1 nM to 1 μM) with calf thymus DNA are shown in Figure 6. The curves are simulated with the independent-site model with N elements per site (the elements are the base pairs; see Figure 7). The determination of N for 1.5 nM dye titration (in the insert) is not accurate due to the low dye concentration (see Appendix), but the effective K_{diss} is well determined. Titrations of Hoechst 33258 with calf thymus DNA were also carried out in the presence of 25% EtOH; the results are given in Table III. It

Table III: K_{diss} and Number of Base Pairs, N, per Site for Hoechst 33258 Binding to Calf Thymus DNA[a]

polymer	Hoechst (nM)	EtOH (%)	NaCl (mM)	$10^9 K_{diss}$ [M (site)]	N
CT DNA	1.5	none	100	*1.7*	ND
CT DNA	10	none	1	*1.1*	70
CT DNA	10	none	10	*0.7*	60
CT DNA	10	none	100	*1.3*	80
CT DNA	10	none	500	*4*	80
CT DNA	30	none	100	*2*	50
CT DNA	100	none	100	*3.3*	30
CT DNA	1000	none	100	*6.3*	16
CT DNA	100	25	1	8.0	35
CT DNA	100	25	10	36	35
CT DNA	100	25	100	230	35
CT DNA	100	25	500	1140	35
CE DNA	10	none	10	1.1	12
CE DNA	10	none	20	3.1	13
CE DNA	10	none	30	3.8	13
CE DNA	10	none	50	5.5	13
poly[d(A-T)]	10	none	10	4.0	10
poly[d(A-T)]	10	none	10	5.0	10
poly[d(A-T)]	10	none	50	7.9	10
poly[d(A-T)]	10	none	100	14	10
poly[d(A-T)]	10	none	200	25	10
poly[d(A-T)]	10	none	500	40	10
poly[d(A-T)]	100	none	100	6.3	10
poly[d(A-T)]	1000	none	100	6.3	10
poly[d(A-T)]	100	25	1	13	10
poly[d(A-T)]	100	25	10	50	10
poly[d(A-T)]	100	25	100	68	10
poly[d(A-T)]	100	25	500	320	10
poly[d(A-T)]	10	25	100	68	10
poly[d(A-T)]	1000	25	100	59	10

[a] K_{diss} values in italics were used in the analysis to determine N; in these cases the determination of N is insensitive to the exact value of K_{diss} used. However, in these cases N was accurately determined. CT, calf thymus; CE, chicken erythrocyte; ND, not determined.

FIGURE 9: Ionic strength dependence of K_{diss} determined with calf thymus DNA, chicken erythrocyte DNA, and poly[d(A-T)], at pH 7 with and without EtOH. (a) Titration curves for poly[d(A-T)] with EtOH. The concentrations of NaCl from left to right are 1, 10, 100, 100, 100, and 500 mM. For all of the data [Hoechst 33258] = 100 nM, except for curve X for which [Hoechst] = 10 nM and curve O for which [Hoechst] = 1 μM. The smooth curves represent fits from the excluded-site model with n = 6. The K_{diss} values are given in Table III. (b) log K_{diss} versus log [NaCl] for poly[d(A-T)] and chicken erythrocyte DNA without EtOH. At the bottom, the regression line is for poly[d(A-T)] only, but can also describe the chicken erythrocyte data and the data for calf thymus DNA and poly[d(A-T)] in the presence of 25% EtOH. At the top, the solid line is the linear regression for calf thymus DNA, and the dashed line is for poly[d(A-T)].
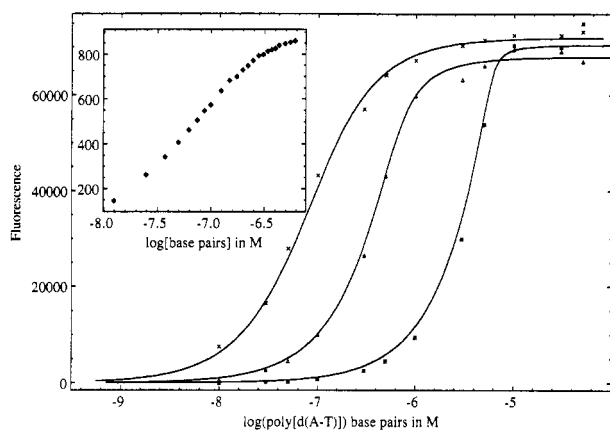

FIGURE 8: Titrations of different but constant concentrations of Hoechst 33258 with poly[d(A-T)] at pH 7. Simulation of the titrations with either the excluded-site model using n = 6 or the independent-site model with N = 10 (see Appendix) results in the same fit. The shape of the curves becomes increasingly characteristic of very tight binding as the concentration of the titrated dye is increased; from left to right: (X) 10 nM; ($\triangle$) 100 nM; (*) 1000 nM. N = 10 or n = 6 (depending upon the model), independent of the [dye].

is known that with 25% ethanol no quenching of the fluorescence is seen at a high dye/DNA ratio (Stokke & Steen, 1985; our observations). The same conclusions as in the absence of EtOH are applicable; however, 25% EtOH decreases the affinity by 2 orders of magnitude and increases the number of the highest affinity sites since, in the presence of 25% EtOH, the largest value of N is 30 instead of 80.

Fluorescence Titrations with Poly[d(A-T)]. Figure 8 presents titrations with poly[d(A-T)] for different dye concentrations. In contrast to natural-sequence DNA, we apply
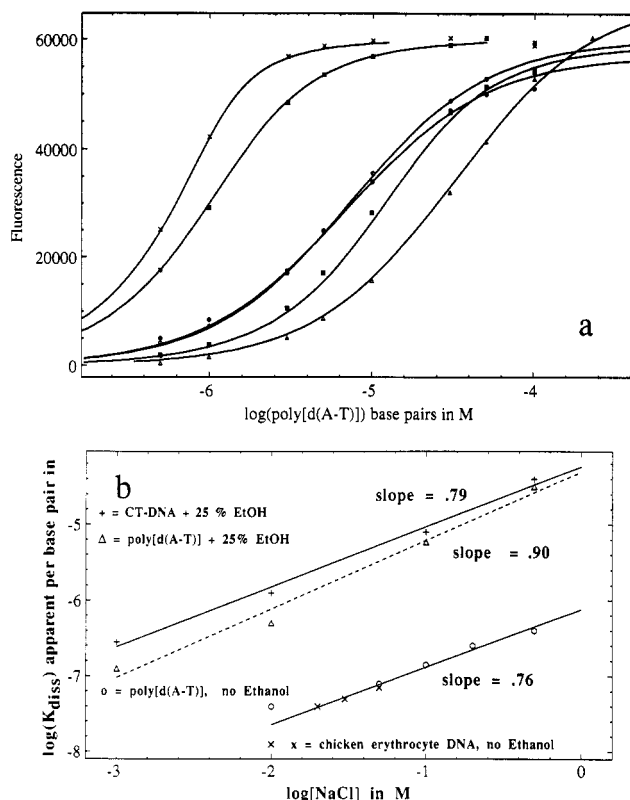
the excluded-site model where n is the number of base pairs per site (see Appendix); this is valid for a uniformly repetitive sequence. Because the ratio of bound dye to base pairs is relatively low throughout all but the very beginning of the titration, the data for poly[d(A-T)] can also be analyzed according to the independent-site model (see above); the corresponding number of base pairs per site, N, is 10 (and is constant), and the K_{diss} is larger than for the most affine sites of calf thymus DNA. As for calf thymus DNA, the presence of 25% EtOH lowers the affinity of the dye for poly[d(A-T)]; the number of base pairs per site, n, remains the same with and without EtOH.

Ionic Strength Dependence of K_{diss}. We have measured the ionic strength dependence of K_{diss} for calf thymus DNA, chicken erythrocyte DNA, and poly[d(A-T)]. Figure 9 and Table III summarize the results. The affinity decreases with ionic strength: $\partial \ln K_{diss} / \partial \ln [Na^+]$ for Hoechst 33258 binding to DNA equals 0.75–0.9.

Fluorescence Titrations with d(CCGGAATTCCGG) and Poly[d(A-T)]: Determination of ΔH and ΔS for Binding of Hoechst 33258. To directly measure the affinity of a site to Hoechst 33258 in the absence of the complications with polymeric DNA, we have titrated low concentrations (1 nM) of this dye at different temperatures with the dodecamer d-(CCGGAATTCCGG). This molecule is very similar to the dodecamer d(CGCGAATTCGCG) complexed with Hoechst
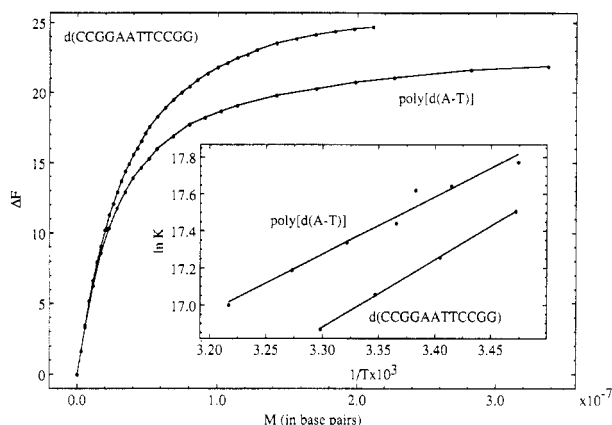
FIGURE 10: Determination of the binding parameters of d-(CCGGAATTCCGG) or poly[d(A-T)] and Hoechst 33258 at low concentration. The curves through the data (ΔF = increase in fluorescence relative to the starting fluorescence) are simulations according to the independent-site model with the following parameters: $K_{assoc} = 3.8 \times 10^{-8}$ M and $\Delta F_{max} = 30.4$ for a titration of 2.288 mL of 1.1 nM Hoechst 33258 adding a total of 33.3 μL of 8.0 μM d(CCGGAATTCCGG) at 25.8 °C with $\lambda_{ex} = 365$ nm and $\lambda_{em} > 450$ nm (upper curve); $K_{assoc} = 2.9 \times 10^{-8}$ M and $\Delta F_{max} = 24.3$ for a titration of 2.271 mL of 1.5 nM Hoechst 33258 adding a total of 25 μL of 15.7 μM poly[d(A-T)] at 27.9 °C with $\lambda_{ex} = 350$ nm and $\lambda_{em} = 460$ nm. The inset is a van't Hoff plot used to determine the thermodynamic parameters in Table II.

Table IV: Thermodynamic Parameters for Hoechst Binding to d(CCGGAATTCCGG) and Poly[d(A-T)]

	$10^7 K_{assoc}$ (BP)$^{-1}$	kcal mol^{-1}		ΔS (eu)
		ΔG	ΔH	
d(CCGGAATTCCGG)	2.63 ± 0.02	−10.12 ± 0.01	−7.3 ± 0.1	+9.6 ± 0.3
poly[d(A-T)]	3.8 ± 0.2	−10.33 ± 0.03	−6.2 ± 0.25	+13.9 ± 1.6

33258 in X-ray diffraction studies (Pjura et al., 1987; Teng et al., 1988). We have also measured the binding of this dye to d(CGCGAATTCGCG) at one temperature (Table II). Since only 1 nM dye is present in these titrations, only the very tight 1:1 complex (dye/dodecamer) for sequence-mediated binding at the AATT region is measured. Titrations of 1.5 nM dye with poly[d(A-T)] at different temperatures were also performed. Examples of such titrations, analyzed according to the independent-site model (simple binding) are shown in Figure 10, which includes the linear van't Hoff plots for determination of the thermodynamic binding parameters listed in Table IV.

DISCUSSION

At Hoechst 33258/base pair ratios less than 2, we find five well-defined stoichiometric complexes with poly[d(A-T)]; the two lowest ratio complexes are also observed with poly[d(G-C)]. The higher stoichiometric complexes are probably mediated by the structure of the grooves (Figure 5) and by aggregation of dye molecules onto those dye molecules already bound in the grooves. This is reminiscent of the aggregation of Hoechst 33258 in neutral solution at concentrations of 5–20 μM. The difference absorption spectra upon binding of the dye to DNA at dye/base pair < 0.15 are similar to those for the dye free in solution, suggesting that specifically DNA-bound dye can induce aggregation on the DNA. At dye/base pair ≥ 2, binding of Hoechst 33258 is largely mediated by the phosphate charges; the dye also aggregates stoichiometrically onto the phosphate-bound dye.

The above conclusions are based upon our Job plots with the following properties that render the deduced stoichiometries

reliable: (1) the data cover the entire mole fraction range; (2) the slopes of the signal change pass through each origin; (3) a common equivalence point, $\chi_e = 0.5$ (charge-mediated binding), for Hoechst 33258 binding to calf thymus DNA, poly[d(A-T)], and poly[d(G-C)] attests to the reliability of the concentrations used; (4) the stoichiometries are the same when observed with difference absorption or fluorescence measurements; (5) the total concentrations of 1, 15, and 30 μM produce no shifts in χ_e values, indicating that the concentrations are sufficiently high to maintain tight binding (an exception is the deviation from linearity in the top plot in Figure 4c, consistent with the minimum in the two bottom plots for a 15-fold increase in concentration); (6) χ_e at an inflection point is independent of the wavelengths used, except for a minor wavelength shift of sequence-mediated equivalence from $\chi_e = 0.14$ to $\chi_e = 0.17$ in Figure 4a (this can represent two binding possibilities as discussed in the following paragraph).

Our data with poly[d(A-T)] and Hoechst 33258 at 30 μM suggest a total of at least seven possible complexes in solution. The basic stoichiometry for sequence-mediated binding to poly[d(A-T)] is 1 dye per 5 or 6 A-T pairs, corresponding to closely spaced inflection points ($\chi_e = 0.17$ and $\chi_e = 0.14$) depending on the detection mode and the wavelength used rather than on experimental uncertainty. This stoichiometry is consistent with the footprinting data (Martin & Holmes, 1983; Harshman & Dervan, 1985; Portugal & Waring, 1988) and with the crystallographic data of Teng et al. (1988), where the dye covers AATT in d(CCGGAATTCCGG) and touches the two neighboring G-C pairs. It is possible that on an extended alternating lattice like poly[d(A-T)] the aligned dye molecules are not juxtaposed but interspaced by one A-T base pair; then, instead of 1 to 5, the complementarity would be 1 to 6.

In addition to 1 dye per 5 or 6 A-T base pairs, four additional complexes composed of 2, 3, 4, and 6 dyes per 5 A-T pairs are formed by binding either directly to a structural element of the helix or to dyes already bound to the helix by sequence- or structure-mediated effects. A Hoechst 33258 molecule that is bound snugly in the minor groove has only its outer edge available for additional interactions; the nitrogen atoms of the benzimidazole ring closest to the DNA (Figure 5) are hydrogen donors to O2 of T and N3 of A (Pjura et al., 1987; Teng et al., 1988), and the outer nitrogen atoms could function as H-bond acceptors for a second dye molecule, bound edge-to-edge, probably with a head-to-tail orientation to minimize charge repulsion of the piperazines in the dimer. Aggregation could involve a repetition of this process, in combination with planar stacks similar to "Scheibeaggregates" (Scheibe, 1939; Yuzhakov, 1975). Such a tendency is prevalent with benzimidazole itself (Hofman, 1953). Furthermore, there are also at least two sequence-independent complexes of 1 dye (charge mediated) and also 2 dyes (dye mediated) per phosphate group.

Hoechst 33258 can interact with G-C rich sequences in ways other than via the phosphate charge (Figure 3); the corresponding fluorescence increase of the dye is much less intense than with A-T sequences. With poly [d(G-C)] prepared from a pure primer sequence this corresponds to 1–2 dyes per 5 G-C pairs, (in fluorescence) or 2 dyes per 5 G-C pairs (in difference absorption), comparable to the two lower stoichiometries for binding to poly[d(A-T)]. We tentatively suggest that this previously undetected binding to poly[d(G-C)] is structurally mediated with a stacked dye dimer in the major groove (Figure 5). However, simultaneous binding with comparable affinity in the minor and major grooves is also possible since Jorgenson

et al. (1988) performed Hoechst 33258 fluorescence competition experiments with poly[d(A-T)] and poly[d(G-C)].

The five complexes with poly[d(A-T)] at χ values lower than for binding to backbone phosphate probably occur as ribbons in the grooves without participation of dyes bound directly to the phosphates (Figure 5). The complex with 1 dye per 5 A-T pairs in the minor groove is sequence (or structure) mediated. Three possible structures for a repeating unit of 2 dyes per 5 A-T pairs in poly[d(A-T)] are as follows: (1) a hydrogen-bonded edge-to-edge dimer, formed by dye-mediated binding to a dye initially bound in the minor groove; (2) a structure formed by one of the two dyes bound in the minor groove and the other bound in the major groove; (3) a structure, similar to that with poly[d(G-C)], with two dyes stacked as a dimer in the major groove. Any of these structures with 2 dyes per 5 A-T pairs would then provide the nucleus for dye-mediated binding with 3, 4, and 6 dyes per 5 A-T pairs. A complex with 6 dyes per 5 A-T pairs, which corresponds to the nearly total fluorescence quenching, could be formed in several ways: (1) an edge-to-edge dye dimer in the minor groove could be the sole nucleus for an aggregate with four additional dyes by a combination of edge-on and stacking interactions; (2) edge-to-edge dimers could form in both the minor and major groove, with a third dimer sandwiched in between; (3) one edge-to-edge dimer in the minor groove and two such dimers in the major groove could form. Possibilities 2 and 3 are presented in Figure 5. The stoichiometric complexes of 3, 4, and 6 dyes per 5 base pairs were not observed with poly[d(G-C)], possibly due to a lack of dye bound in its minor groove as a nucleus for dye-mediated binding.

The fluorescence of Hoechst 33258 molecules bound to DNA at high concentrations of the dye is quenched by the formation of an aggregate of 6 dyes per 5 A-T base pairs. At a dye/DNA base pair even higher than 6/5, Hoechst 33258 binds to backbone phosphates and forms stoichiometric aggregates involving 1, 2, and possibly more dyes per phosphate. The propensity of the dye to form distinct stoichiometric aggregates which cover the entire outside of the DNA could provide the driving force to unwind supercoiled plasmids. The activities of DNase I and micrococcal nuclease are apparently only slightly affected by dye aggregation since at high dye/DNA ratios and high dye concentrations (Portugal & Waring, 1988; Jorgenson et al., 1988) the results of footprinting experiments are similar to those obtained at low ratios of dye/DNA and to those using methidiumpropyl-EDTA–Fe(II) to cleave the DNA (Harshman & Dervan, 1985).

The titrations at low concentrations reveal the specific or sequence-mediated binding to natural-sequence DNA. The K_{diss} values of Table III range from 1 nM to 1 μM (Tables II and III). The number of base pairs per site, N (independent-site model, see Appendix), depends upon the titrated dye concentration (Figure 7). $1/N$ is an estimate of the average frequency of occurrence per base pair for a strongly binding sequence on the calf thymus DNA. The largest N for the strongest binding of Hoechst 33258 to calf thymus DNA is estimated to be 100 from Figure 7, much greater than the 4–6 base pairs physically covered by a tightly bound dye molecule. As the data in Figure 7 clearly indicate, N is not constant for all titrations, and a more complex model than simply binding to one type of site must be considered.

If the titration curve displays the characteristics of tightly bound stoichiometric complexes, $1/N$ is interpreted as the fraction of base pairs containing a strong site with a K_{diss} small relative to the dye concentration. At the stoichiometric ratio of dye/DNA base pairs of $1/N$, the sites with a $K_{diss} \ll$ total

[dye] will be almost fully occupied, and at lower ratios of dye/DNA the concentration of free dye will be unmeasurably low. This situation appears as a stoichiometric binding of the dye to a single type of very affine site. Apparently a large range of K_{diss} values exists corresponding to different sequences of the sites. As the concentration of the dye is increased, the apparent number of base pairs per site, N decreases due to an increase in the total spectrum of sites with a $K_{diss} <$ total [dye] (Table III; Figures 6 and 7). A sequence corresponding to the strongest site of an A-T specific ligand will occur once in 32 base pairs if (1) the dye occupies 5 consecutive A-T pairs, (2) the nucleotide sequence at the site requires either A or T, so that the probability of correct nucleotides is $1/2$ for every position, and (3) the DNA has a random sequence. This estimate of $1/2^5 = 1/32$ is larger than the $1/100$ found for calf thymus DNA at the lowest dye concentrations (Table III and Figure 7), indicating that the site requirements for strong binding are more stringent than just the presence of A-T base pairs and that specific sequence permutations are necessary.

This hitherto unrecognized high affinity of Hoechst 33258 (Table III) could have important implications for the design of antiviral and anticancer drugs such as netropsin analogues (lexitropins) or isohelical sequence-reading polymers (isolexins) (Goodsell & Dickerson, 1986). Drug design often strives to increase the complementarity with an appropriate base pair sequence; however, to avoid undesirable interactions, there must be a way to stop the drug from binding to closely related but unwanted sites. This will be difficult to achieve with simple ligand molecules. Unless the concentration of the ligand can be accurately controlled (this is especially difficult in the nanomolar range), binding to similar sequences with less affinity will occur.

We find, contrary to the earlier literature, that the tight specific binding decreases with ionic strength. The values of $\partial \ln K_{diss}/\partial[Na^+]$ are close to the theoretical prediction of 0.88 (Record et al., 1976, 1978) as expected for binding of a monocationic dye. In the earlier literature the "true" ionic strength dependence of the specific binding remained hidden since titrations of such high dye concentrations ($\sim 1 \mu$M) were actually stoichiometric determinations of the total [dye] (see for instance Figure 6). The dependence of K_{diss} upon the ionic strength is identical for all of the DNAs tested and is not affected by the presence of 25% ethanol.

The results obtained with d(CGCGAATTCGCG) and d-(CCGGAATTCCGG) in Table II clearly show that both these sequences have a K_{diss} (referring to one site per dodecamer) of ~ 3 nM at 25 °C, which is only slightly larger than ~ 2.5 nM found for poly[d(A-T)] (per site with the excluded-site model) or 1–2 nM for the tightest binding sites of calf thymus (per site with the independent-site model). The ΔH's for binding of Hoechst 33258 to poly[d(A-T)], $\Delta H = -7.3 \pm 0.1$ kcal mol^{-1}, and to d(CCGGAATTCCGG), $\Delta H = -6.2 \pm 0.5$ kcal mol^{-1}, are nearly equal, indicating a similarity in the hydrogen bonding and or van der Waals contacts in the dye–DNA structures. The similarities between ΔG and ΔH for binding of Hoechst 33258 to the dodecamer and to poly-[d(A-T)] and ΔG for the strongest binding to calf thymus DNA are indirect proof that our independent-site analysis for calf thymus DNA as well as our estimate of the length of the strongest sites $N = 100$ are correct. It is likely that the dodecamer site is very close to the optimal sequence. These facts strengthen our conjecture that a large spectrum of Hoechst 33258 binding affinities exist on natural-sequence DNA.

The thermodynamic parameters for binding of Hoechst 33258 with poly[dA-T)] and d(CCGGAATTCCGG) in Table

IV allow a comparison with other A-T-preferential and peripherally binding dyes and drugs. The enthalpy change correlates well with the number of hydrogen bonds to the bases in poly[d(A-T)] and with the number of van der Waals contacts. Distamycin A forms four NH hydrogen bonds, ΔH = -18.5 kcal mol^{-1} (Breslauer et al., 1987); netropsin has three NH bonds, ΔH = -9.2 kcal mol^{-1} (Marky et al., 1983) or -11.2 kcal mol^{-1} (Breslauer & Marky, 1987; Breslauer et al., 1987); Hoechst 33258 has two NH bonds (Pjura et al., 1987; Teng et al., 1988), ΔH = -6.2 kcal mol^{-1} (Figure 10); and DAPI has only one NH hydrogen bond, ΔH = -4.5 kcal mol^{-1} at very low D/P values, with the ΔH becoming more positive but constant (-3 kcal mol^{-1}) for D/P values between 0.1 and 0.3 (Manzini et al., 1983). These calorimetrically determined values for DAPI agree with our preliminary determination of ΔH from fluorescence titrations for the binding of this dye to d(CCGGAATTCCGG) (Loontiens et al., 1989b).

ACKNOWLEDGMENTS

APPENDIX

Expectations of a Simple Independent-Site Model. Consider a simple binding reaction between two chemical species A and B that have only one independent site per molecule and form a complex C with a dissociation constant K_{diss} = [A]·[B]/[C]. Assume that neither A nor B fluoresces but C does; A is the titrated species (constant input concentration [A$_0$]), and B is the titrant (varying input concentration [B$_0$]). If [A$_0$] $\ll K_{diss}$, a plot of the fluorescence versus $\log_{10}$ [B$_0$] appears symmetric, and its width, the difference on the $\log_{10}$ [B$_0$] scale between the 0.1 and the 0.9 fraction of bound titrated species [C]/[A$_0$], is about 1.9. If [A$_0$] $\gg K_{diss}$, essentially all of B is quantitatively bound unless [B$_0$] > [A$_0$]; this is "high affinity stoichiometric binding", and the 0.1–0.9 width of the titration curve extends over only a $\log_{10}$ [B$_0$] scale of 0.95, and the curve is nonsymmetrical (Figure 6). At intermediate concentrations, the shape of the titration curve will more or less resemble either of these extremes. At low [A$_0$], K_{diss} can be accurately determined if [B$_0$] is known. At high [A$_0$], [B$_0$] can be accurately determined if [A$_0$] is known, but the value of K_{diss} cannot be determined. For intermediate [A$_0$], both [B$_0$] and K_{diss} can be determined if [A$_0$] is known.

The titration curve's position on the $\log_{10}$ [B$_0$] axis depends upon the concentration of the titrated species. If [A$_0$] $\ll K_{diss}$, the titration curve will be positioned as far to the left as is possible, and the position of the 0.5 bound fraction on the $\log_{10}$ [B$_0$] axis is equal to $\log_{10}$ [K_{diss}], independent of [A$_0$]. As [A$_0$] increases, the position of the titration curve will eventually shift to the right; then, the [B$_0$] for a 0.5 bound fraction of [A$_0$] no longer equals K_{diss}. When [A$_0$] $\gg K_{diss}$, the shift will be linear with $\log_{10}$ [A$_0$]. However, at any concentration, the titration curve will be predictable, and the curve can always be described with a constant stoichiometry of 1:1.

All of the above arguments remain valid if A has a single binding site per molecule, as above, and the titrant molecule (now defined as B$_N$) has N identical elements (or subunits) but only one binding site in the molecule, that is, one binding site per N identical elements of titrant (the concentration of elements in the B$_N$ molecule is defined as [B$_E$], where [B$_E$] = N[B$_N$] and K_{diss} = [A][B$_N$]/[C]). If the input concentration of B is known, and expressed as [B$_{0,E}$] (elements) rather than [B$_{0,N}$] (sites), the titration curves will simply be offset to the

right by the amount $\log_{10}$ N. If [A$_0$] is known, we can determine N by carrying out titrations with [A$_0$] $\gg K_{diss}$ and varying [B$_{E,0}$]. An accurate determination of N can be made. However, when [A$_0$] $\ll K_{diss}$, the binding curve can be simulated equally well by assuming any concentration unit for the B species; for such low concentrations of A$_0$ the binding equations depend only upon K_{diss}/N. For intermediate [A$_0$] both K_{diss} and N can be determined. For such simple systems the value of N will be constant, regardless of the interactions.

Excluded-Site Model for the Repetitive Sequence Poly[d-(A-T)]. For a repetitive sequence like poly[d(A-T)] the potential binding sites overlap according to the excluded-site binding model (McGhee & von Hippel, 1974; Zasedatelev et al., 1971); the number of base pairs per site, n, refers directly to the extension of the site on a fully occupied lattice as defined according to the excluded-site model. The only variables are n and K_{diss}. If the ratio of dye/DNA is sufficiently low, the data can be fitted by either the independent-site model (yielding N) or the excluded-site model (yielding n), and it can be shown that $N = 2n - 1$ (Schellman, 1974; Macgregor et al., 1985). In a concave Scatchard plot observed with excluded-site binding, $1/n$ is the extrapolated intercept on the ν axis, whereas $1/N$ is the intercept on the ν axis for the slope of the lowest ν values.

REFERENCES

Baguley, B. C. (1982) *Mol. Cell. Biochem. 43*, 167–181.

Breslauer, K. J., Remeta, D. P., Chou, W.-Y., Ferrante, R., Curry, J., Zaunczkowski, D., Synder, J. G., & Marky, L. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8922–8926.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Part III, pp 1135–1139, W. H. Freeman, New York.

Chou, W. Y., Marky, L. A., Zaunczkowski, D., & Breslauer, K. J. (1988) *J. Biomol. Struct. Dyn. 5*, 345–359.

Commings, D. E. (1975) *Chromosoma 52*, 229–243.

Goodsell, D., & Dickerson, R. E. (1986) *J. Med. Chem. 29*, 727–733.

Harshman, K. D., & Dervan, P. B. (1985) *Nucleic Acids Res. 13*, 4825–4835.

Hofman, K. (1953) *The Chemistry of Heterocyclic Compounds, Part I, Imidazole and Its Derivatives* (Weissberger, A., Ed.) Interscience Publishers, New York.

Huang, C. Y. (1982) *Methods Enzymol. 87*, 509–525.

Job, P. (1928) *Ann. Chim. (Paris) 9*, 113–203.

Jorgenson, K. F., Varshney, U., & van de Sande, J. H. (1988) *J. Biomol. Struct. Dyn. 5*, 1005–1023.

Krey, A. K., & Hahn, F. E. (1975) *Biochemistry 14*, 5061–5067.

Langlois, R. G., & Jensen, R. H. (1979) *J. Histochem. Cytochem. 27*, 72–79.

Latt, S. A., & Wohlleb, J. C. (1975) *Chromosoma 52*, 297–316.

Latt, S. A., & Stetten, G. (1976) *J. Histochem. Cytochem. 24*, 24–33.

Loontiens, F. G., Regenfuss, P., Dumortier, L., & Clegg, R. M. (1989a) *Arch. Int. Physiol. Biochem. 97*, B104.

Loontiens, F. G., Regenfuss, P., Zechel, A., & Clegg, R. M. (1989b) *Arch. Int. Physiol. Biochem. 97*, B105.

Macgregor, R. B., Jr., Clegg, R. M., & Jovin, T. M. (1985) *Biochemistry 24*, 5503–5510.

Manzini, G., Barcellona, M. L., Avitabile, M., & Quadrifoglio, F. (1983) *Nucleic Acids Res. 11*, 8861–8876.

Marky, L. A., & Breslauer, K. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4359–4363.

Marky, L. A., Blumenfeld, K. S., & Breslauer, K. J. (1983) *Nucleic Acids Res. 11*, 2857–2870.

Martin, R. F., & Holmes, N. (1983) *Nature 302*, 452–454.

McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469–489.

Mikhailov, M. V., Zasedatelev, A. S., Krylov, A. S., & Gurskii, G. V. (1981) *Mol. Biol.* (*Moscow*) *15*, 541–554.

Pjura, P. E., Grzeskowiak, K., & Dickerson, R. E. (1987) *J. Mol. Biol. 197*, 257–271.

Portugal, J., & Waring, M. J. (1988) *Biochim. Biophys. Acta 949*, 148–158.

Pullmann, B. (1985) *Proc. Int. Symp. Biomol. Struct. Interact., Suppl. J. Biosci. 8*, 681–688.

Record, M. T., Lohman, T. M., & DeHaseth, P. (1976) *J. Mol. Biol. 107*, 145–158.

Record, M. T., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103–178.

Regenfuss, P., Loontiens, F. G., & Clegg, R. M. (1989) *Arch. Int. Physiol. Biochem. 97*, B113.

Scheibe, G. (1939) *Angew. Chem. 52*, 631–637.

Schellman, J. A. (1974) *Isr. J. Chem. 12*, 219–238.

Simon, R. H., & Felsenfeld, G. (1979) *Nucleic Acids. Res. 6*, 689–696.

Steiner, R. F., & Sternberg, H. (1979) *Arch. Biochem. Biophys. 197*, 580–588.

Stokke, T., & Steen, H. B. (1985) *J. Histochem. Cytochem. 33*, 333–338.

Teng, M.-K., Usman, N., Frederick, C. A., & Wang, A. H.-J. (1988) *Nucleic Acids Res. 16*, 2671–2690.

Weisblum, B., & Haenssler, E. (1974) *Chromosoma 46*, 255–260.

Yuzhakov, V. I. (1975) *Russ. Chem. Rev. 48*, 1076–1091.

Zasedatelev, A. S., Gurskii, G. V., & Volkenshtein, M. V. (1971) *Mol. Biol. 5*, 245–251.

Zimmer, C., & Waehnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31–112.

Characterization of the Secondary Structure and Melting of a Self-Cleaved RNA Hammerhead Domain by ¹H NMR Spectroscopy[†]

Ann Caviani Pease and David E. Wemmer*

Department of Chemistry and Chemical Biodynamics Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720

ABSTRACT: We have completely assigned the extreme low-field ring–NH nuclear magnetic resonance spectrum of a self-cleaving RNA in the absence of magnesium ions by experiments involving sequential Overhauser enhancements between adjacent base pairs. These assignments substantiate the hammerhead secondary folding model proposed by Symons and co-workers for this class of self-cleaving RNA [Hutchins, C. J., Rathjen, P. D., Forster, A. C., & Symons, R. H. (1986) *Nucleic Acids Res. 14*, 3627–3640; Forster, A. C. & Symons, R. H. (1987) *Cell 49*, 211–220; Kneese, P., & Symons, R. H. (1987) in *Viroids and Viroid-like Pathogens* (Semancick, J. S., Ed.) pp 1–47, CRC Press, Boca Raton, FL]. No resonances due to tertiary base pairs could be identified in the low-field spectrum, and addition of MgCl₂ to the sample did not produce additional resonances in this region of the spectrum.

A site-specific self-cleavage reaction has been observed for the (+) and (−) strands of satellite tobacco ringspot virus RNA (STobRV RNA) which results in 5′-hydroxyl and cyclic 2′,3′-phosphate termini (Prody et al., 1986; Buzayan et al., 1986a,b). Similar reactions have been observed for the (+) and (−) strands of avocado sunblotch virusoid (Hutchins et al., 1986) and lucerne transient streak viroid (Forster & Symons, 1987). Buzayan and co-workers have found that fewer than 100 nucleotides are necessary for efficient cleavage of (+) STobRV RNA (Buzayan et al., 1986). Symons and co-workers have proposed a hammerhead secondary folding model for the self-cleavage domain (Hutchins et al., 1986; Forster & Symons, 1987; Keese et al., 1987). This domain

consists of 13 conserved nucleotides and 3 variable-length double-stranded helices (Figure 1). Cleavage occurs only at a specific phosphodiester linkage and requires no cofactor other than magnesium ions. Although the nucleotides directly at and near the cleavage site may vary, the position of cleavage within the hammerhead domain is conserved. Therefore, it is clear that a specific three-dimensional folding of the hammerhead domain is required for cleavage. The nature of the interactions giving rise to this folding is as yet unknown.

We have chosen to study the three-dimensional structure of such an RNA by nuclear magnetic resonance (NMR). NMR has been used extensively to study the structure of transfer RNA (tRNA) in solution. Particular attention has been paid to the extreme low-field (15–11 ppm) region of the spectrum [for reviews, see Kearns and Shulman (1974), Kearns (1976, 1977), Reid and Hurd (1977), Schimmel and Redfield (1980), and Reid (1981)]. Usually only hydrogen-bonded imino protons involved in secondary and tertiary base pairs resonate in this region, leaving it comparatively well resolved for a large nucleic acid (Reid et al., 1975, 1977, 1979). One-dimensional nuclear Overhauser enhancements (NOEs)