DNA Minor Groove Binding of Cross-Linked Lexitropsins: Experimental Conditions Required to Observe the Covalently Linked WPPW (Groove Wall-Peptide-Peptide-Groove Wall) Motif

Yong-Huang Chen and J. William Lown
Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2 Canada

ABSTRACT A theoretical analysis of binding interactions between covalently cross-linked lexitropsins and DNA is undertaken, in which a novel cyclic symmetric 2:2 dimeric lexitropsin-DNA-binding model is proposed. Applicability of commonly used techniques including NMR, quantitative footprinting, CD, and ethidium fluorometry to differentiate the covalently linked WPPW (groove Wall-Peptide-Peptide-groove Wall) from a 2:2 cross-linked lexitropsin-DNA duplex structure is examined.

INTRODUCTION

The naturally occurring oligopeptide antibiotics netropsin 1 and distamycin 2 (Fig. 1) represent important paradigms of sequence-selective DNA minor groove binders (Zimmer and Wahnert, 1986). Generalized information-reading molecules analogous to these natural products are defined as lexitropsins, and their DNA recognition capacity has been exploited extensively (Kopka et al., 1985; Lown, 1988, 1990, 1992, 1993). Structural elucidation of 2:1 antiparallel side-by-side binding between distamycin and oligonucleotides further contributes to our understanding of interactions between minor groove binders and DNA (Pelton and Wemmer, 1989, 1990, Fig. 2). This minor groove wall-peptide-peptide-minor groove wall (WPPW) antiparallel side-by-side motif was recently found to be extendable to interactions between imidazole-containing lexitropsins and corresponding matched base sequences with much higher cooperativity (Dwyer et al., 1992; Mrksich et al., 1992). More interestingly, one imidazole-containing lexitropsin molecule and one distamycin molecule can occupy the same site, constituting a heterodimeric WPPW motif, with greater binding strength than either of the corresponding homodimeric motifs at the same site (Geierstanger et al., 1993; Mrksich and Dervan, 1993a). Particular interest in the latter heterodimeric complex attaches to its unique discriminatory strand-specific recognition of sequences containing GC base pairs. To exploit the latter property and, therefore, to provide potentially higher sequence-specific recognition, novel covalently linked dimeric lexitropsins were designed, synthesized, and their DNA-binding characteristics examined (Dwyer et al., 1993; Mrksich and Dervan, 1993b, 1994; Chen and Lown, 1994) (Figs. 2 and 3).

In one study of polymethylene-linked bis-oligopeptides 3a–3d (Fig. 3), it was concluded that the linker length has little effect on the binding mode and binding strength of these dimeric ligands (Dwyer et al., 1993; Mrksich and Dervan, 1993b). From the NMR studies, it was indicated that all four dimeric lexitropsins bind with nearly identical geometry to the oligonucleotide-possessing 5′-TGACT-3′ core in a uniform WPPW 1:1 bidentate mode. Structures of the covalently unconnected WPPW motif and the covalently connected WPPW motif were suggested to be nearly identical in the 5′-TGACT-3′ site (Dwyer et al., 1993). From the quantitative DNase I footprinting experiments, binding constants for lexitropsin dimers 3a–3d showed little difference (Mrksich and Dervan, 1993). There is an apparent consistency between NMR spectroscopy and quantitative footprinting results.

For a similar series of structures 4a–4d (Fig. 3), we concluded: the tetrakis(methylene) chain linker is too short to allow any bidentate binding of dimer 4a to the alternating AT polymer and certainly the tris(methylene) chain linker is far too short by extrapolation; the pentakis(methylene) and hexakis(methylene) linkers permit dominance of the WPPW 1:1 bidentate binding of dimers 4b and 4c but with major local structural distortions; the heptakis(methylene) is a much better linker to facilitate the bidentate binding of 4d but is still far from ideal; the linker length clearly has a profound effect on the binding mode and binding strength (Chen and Lown, 1994). This conclusion is based on interpretation of CD titration curves and ethidium fluorometry experiments with homo- and alternating AT polymers, by applying the McGhee-von Hippel equation. There is a strict correspondence among CD, ethidium fluorescence quenching, and ethidium displacement data. Despite substantial differences in terms of lexitropsin-DNA duplex-interacting systems and the experimental methods used to establish the binding characteristics in these two independent studies, we believe that it is worth searching for reasons other than these apparent differences, because the effect of linker length has to be understood to further the design of cross-linked or bridged lexitropsins. Moreover, these novel and unprecedented dimeric lexitropsin structures and their interactions with DNA oligomers and polymers require wide ranging examination

Received for publication 7 June 1994 and in final form 1 November 1994.
Address reprint requests to Dr. J. William Lown, Department of Chemistry, University of Alberta, E3-44 Chemistry Building, Edmonton, Alberta T6G 2G2, Canada. Tel.: 403-492-3646; Fax: 403-492-8231.

Abbreviations used: WPW, groove wall-peptide-groove wall; WPPW, groove wall-peptide-peptide-groove wall antiparallel side-by-side motifs.

0006-3495/95/05/2041/08 $2.00

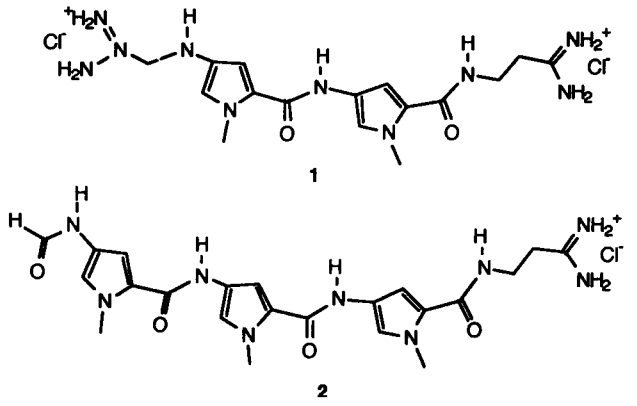

FIGURE 1 DNA minor groove-binding oligopeptides netropsin and distamycin.

using an array of physical techniques in case unanticipated modes of binding are encountered, and possible limitations of various techniques have to be appreciated to derive meaningful conclusions. The present report summarizes our theoretical exploration to resolve disagreements between two independent studies, by analyzing the various binding phenomena these ligands may exhibit under different conditions and, particularly, to delineate the utility and limitations of different analytical techniques that have been used.

EFFECT OF CROSS-LINKAGE ON BINDING EQUILIBRIUM

Let us start with a dimeric lexitropsin and an oligonucleotide capable of accommodating a WPPW motif. First of all, it must be recognized that a dimeric ligand with a flexible linker may bind to two DNA duplexes simultaneously, in addition to the desired 1:1 bidentate binding, as illustrated in Fig. 2. From model inspection, cross-linked lexitropsin dimers with linkers longer than a bis(methylene) chain can avoid serious steric clash between two bound duplexes. Therefore, the formation of polymeric aggregation in the solution of the DNA and the dimeric ligand has to be taken into account. This aggregation factor should be more significant

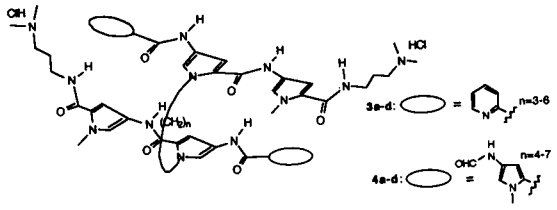

FIGURE 3 Design of cross-linked lexitropsins.

under NMR titration conditions where the DNA concentration is relatively high. Second, structural differences between the 1:1 bidentate WPPW motif and those WPPW motifs in symmetric cyclic n:n complexes (e.g., 2:2, 3:3, Fig. 4) cannot be readily differentiated from standard one-dimensional and two-dimensional NMR (e.g., NOESY) spectra. Similar sets of signals are expected for all of these motifs. Therefore, the observation of a single DNA-ligand complex with a stoichiometric ratio of 1:1 is not sufficient in itself to prove which structure is really present in solution. For these two reasons, symmetric higher polymeric complexes have to be considered as alternative possibilities.

With these considerations in mind, we may gain better insight by analyzing the binding equilibrium that includes a putative 2:2 symmetric complex, as illustrated in Fig. 5.

The binding isotherm equation is as follows:

$$\gamma = \frac{K_{11}[L] + 2K_{11}K_{21}[L]^2 + K_{11}K_{12}[L][D] + 2K_{11}K_{21}K_{22A}[L]^2[D]}{1 + K_{11}[L] + K_{11}K_{21}[L]^2 + 2K_{11}K_{12}[L][D] + 2K_{11}K_{21}K_{22A}[L]^2[D]}.$$

γ is the binding density, the number of bound ligands per oligonucleotide. $[D]$ and $[L]$ are concentrations of free DNA and the bidentate ligand. K_{11}, K_{21}, K_{12}, K_{22A}, and K_{22B} are composite binding constants: $K_{11} = [DL]/[D][L]$, $K_{21} = [DL_2]/[DL][L]$, $K_{12} = [D_2L]/[DL][D]$, $K_{22A} = [D_2L_2]/[DL_2][D]$, $K_{22B} = [D_2L_2]/[D_2L][L]$. $[DL]$ and $[D_2L_2]$ are total concentrations of DL and D_2L_2 species, respectively. Pathway A ($DL \rightarrow DL_2 \rightarrow D_2L_2$) and pathway B ($DL \rightarrow DL_2 \rightarrow D_2L_2$) have the same start and end points. Therefore, $K_{21}K_{22A} = K_{12}K_{22B}$. $\alpha = [DL]_{closed}/[DL]_{open}$, which characterizes the tendency to form the closed bidentate complex versus the open monodentate complex or the 1:1 bidentate-binding enhancement over the 1:1 monodentate binding. Similarly, β =

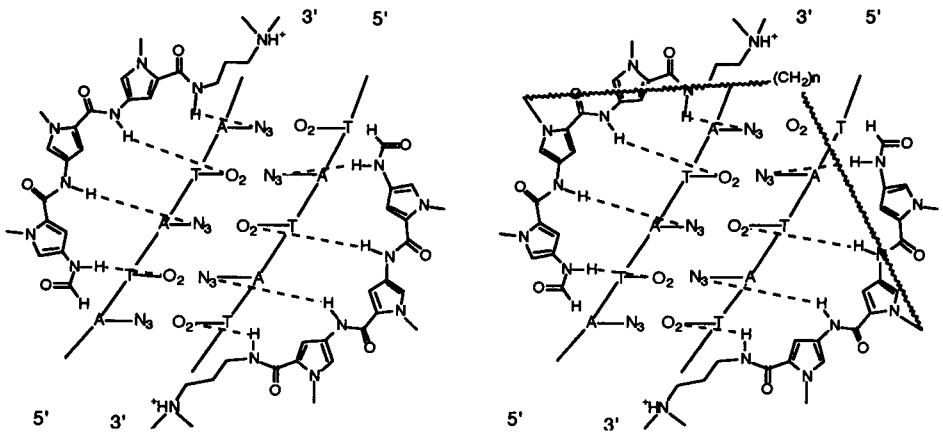

FIGURE 2 The 2:1 WPPW binding motif and the desired bidentate binding of a cross-linked lexitropsin (only hydrogen-bonding interactions are shown).

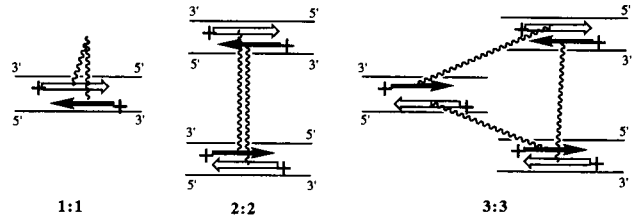

FIGURE 4 Possible structures of 1:1 stoichiometric ratio.

$[D_2L_2]_{closed}/[D_2L_2]_{open}$, which expresses the tendency to form the closed 2:2 complex versus the open 2:2 complex or the 1:1 bidentate-binding enhancement of the "bidentate ligand" DL_2 over the 1:1 monodentate binding. It is noted here that all concentrations are in M.

We may define $k_{11} = [DL]_{open}/[D][L]$, $k_{21} = [DL_2]/[DL]_{open} \times [L]$, $k_{12} = [D_2L]/[DL]$open $\times [D]$ and $k_{22a} = [D_2L_2]_{open}/[DL_2] \times [D]$. As a result, $K_{11} = (1 + \alpha)k_{11}$, $K_{21} = k_{21}/(1 + \alpha)$, $K_{12} = k_{12}/(1 + \alpha)$ and $K_{22A} = (1 + \beta)k_{22a}$. $2K_{11}K_{21}[L]^2$ and $K_{11}K_{12}[L][D]$ terms in the numerator of the binding isotherm equation, representing contributions from DL_2 and D_2L species, respectively, are equal to $2k_{11}k_{21}[L]^2$ and $k_{11}k_{12}[L][D]$. If nonbinding appendages of bidentate ligands do not influence 1:1 WPW and 2:1 WPPW motifs significantly, which is supported by our previous binding constant measurements (Chen and Lown, 1994), k_{11} and k_{21a} will be more or less unchanged in comparison with those of the monomer, and k_{11}, k_{12}, and k_{22a} will be approximately equal. Therefore, the DL_2 species as well as the D_2L species have more or less constant coefficients for their corresponding terms, among all dimeric lexitropsins. The same situation applies to the open DL species as well as the open D_2L_2 species. Logically, the interesting species to investigate will be the closed 1:1 bidentate DL complex and the closed symmetric 2:2 D_2L_2 complex, both having changeable coefficients (due to changeable α and β) for their corresponding terms among all dimeric ligands. More importantly, cases of interest to us have either the bidentate 1:1 complex or the symmetric 2:2 complex assume dominance (i.e., $\alpha \gg 1$ or $\beta \gg 1$), thus contributing to difficulties of differentiation, for example, by NMR spectroscopy. Correspondingly, the DL_2, D_2L, the monodentate-binding DL, and the open D_2L_2 contributions should be negligible in such cases.

If the bidentate-binding DL predominates under the experimental conditions, the following relationship holds: $[DL]_{closed} \gg [D_2L_2]_{closed}$ or $[DL] \gg [D_2L_2]_{closed}$.

Therefore, $(1 + \alpha)k_{11}[L] > 2K_{11}K_{21}(\beta k_{22a})[L]^2[D] = 2\beta k_{11}k_{21}k_{22a}[L]^2[D]$,

or

$$[D][L] < (1 + \alpha)/2\beta k_{21}k_{22a}. \quad (1)$$

When the added ligand is $1/N$ times the total DNA concentration (i.e., $[L]_{total} = [D]_{total}/N$, $[DL] \approx [L]_{total} = [D]_{total}/N$, because very little free ligand exists in the solution during titration because of the large binding constants generally observed (over 10^6 M^{-1}). Because $[D]_{total} \approx [D] + [DL]$, $N \times [DL] = [D] + [DL]$ or $[DL]/[D] = 1/(N - 1)$.

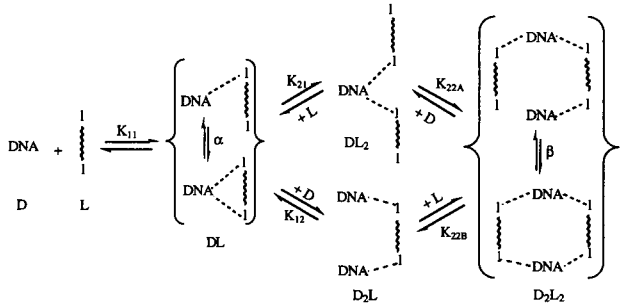

FIGURE 5 A binding model including a symmetric 2:2 complex. D represents the DNA with a binding site; L is a symmetric dimeric lexitropsin.

Because $K_{11} = (1 + \alpha)k_{11} = [DL]/[D][L]$, $[L] = 1/(N - 1) \times (1 + \alpha)k_{11})$.

Substitution of this expression into Eq. 1 leads to to the following inequality:

$$[D] < (N - 1) \times (1 + \alpha)^2k_{11}/2\beta k_{21}k_{22a}.$$

Also,

$$[D]_{total} \approx [D] + [DL] = (1 + 1/(N - 1) \times [D].$$

Therefore,

$$[D]_{total} < \frac{N \times (1 + \alpha)^2k_{11}}{2\beta k_{21}k_{22a}} \quad (2)$$

Equation 2 is approximately the condition to be satisfied if the closed bidentate DL complex is dominant.

In contrast, if the symmetric tetraplex D_2L_2 is predominant, (1) will be reversed:

$$[D][L > (1 + \alpha)/2\beta k_{21}k_{22a}. \quad (3)$$

When $[L]_{total} = [D]_{total}/N$, $2[D_2L_2]_{closed} \approx 2[D_2L_2] \approx [L]_{total} = [D]_{total}/N$. Because $[D]_{total} \approx [D] + 2[D_2L_2]_{closed}$, $2N \times [D_2L_2]_{closed} \approx [D] + 2[D_2L_2]_{closed}$ or $[D_2L_2]_{closed} \approx [D]/2(N - 1)$.

Because $K_{11}K_{21} \times (\beta k_{22a}) = k_{11}k_{21} \times (\beta k_{22a}) = [D_2L_2]_{closed}/[D]^2[L]^2 = 1/(2(N - 1) \times [D][L]^2)$, $[L] = (2(N - 1) \times [D]k_{11}k_{21}\beta k_{22a})^{-0.5}$.

Substitution of the latter expression into Eq. 3 leads to the following inequality:

$$([D]/2(N - 1) \times k_{11}k_{21}\beta k_{22a})^{0.5} > (1 + \alpha)/2\beta k_{21}k_{22a},$$

i.e.

$$[D] > (N - 1) \times (1 + \alpha)^2k_{11}/2\beta k_{21}k_{22a}.$$

$$[D]_{total} \approx [D] + 2[D_2L_2] = (1 + 1/(N - 1))\,[D].$$

i.e.

$$[D]_{total} > \frac{N \times (1 + \alpha)^2k_{11}}{2\beta k_{21}k_{22a}} \quad (4)$$

Equation 4 specifies the required condition to allow the symmetric 2:2 complex to prevail. Equations 2 and 4 are different only in the direction of the inequality symbol. They

are merged to give Eq. 5:

$$[D]_{total}\ Vs\ \frac{N \times (1+\alpha)^2 k_{11}}{2\beta k_{21} k_{22a}} = Z. \quad (5)$$

In other words, the relative magnitude of Z and the total DNA concentration decides whether 1:1 or 2:2 complexes will predominate. If $[D]_{total} > Z$, the 2:2 symmetric complex is predominant and, in contrast, the 1:1 bidentate complex prevails when $[D]_{total} < Z$.

Equation 5 has some important physical implications. First, for a given bidentate ligand, the experimental conditions (e.g., $[D]_{total}$) have a significant role to play in shifting the DL and D_2L_2 equilibrium, and sensitive techniques facilitate the observation of the 1:1 bidentate complex. Second, the binding constant k_{21} is more directly relevant to the binding equilibrium because k_{11} and k_{22a} are similar to each other from the early analysis and cancel each other. Moreover, under the same conditions, bidentate ligands of various linker length will behave differently in terms of the position of equilibrium, depending on the significance of the bidentate binding (α) and the ease of formation of the symmetric 2:2 complex (β). Finally, the order of α is 2, whereas the order of β is 1 in Eq. 5, which is a direct manifestation of different stoichiometries of the bidentate DL and symmetric 2:2 complexes. Clearly, the stoichiometry factor will always favor the formation of the 1:1 bidentate complex over the 2:2 complex, even for cases with $\alpha < 1$.

To assess further the binding equilibrium under various conditions using Eq. 5, we need to know theoretical limits of α and β values. When two monomeric ligands of a "macromolecule" are covalently linked to provide a dimeric ligand, its bidentate-binding constant ideally equals 55.6 times the product of two individual monomeric binding constants, provided that there is neither net enthalpy change nor other entropy gain or loss except the mixing entropy gain due to differential stoichiometry of binding equilibria between the monomeric ligands-macromolecule system and the dimeric ligand-macromlecule system (LePecq and Roques, 1986; Chipman and Sharon, 1969). Specifically, the ideal condition means: 1) there is no interaction between two monomeric moieties in the dimeric ligand in solution; 2) there is no conformational enthalpy and entropy gain or loss from the covalent linkage during binding of the dimeric ligand, the covalent linkage does not interact with the macromolecule, and it does not make the monodentate binding of the dimeric ligand significantly different from the binding of monomers structurally and energetically; 3) structures of the dimeric ligand-macromolecule binary complex and monomeric ligands-macromolecule ternary complex are essentially the same except the former has a linker; and 4) there is no additional translational, rotational, and vibrational entropy gain or loss in the step from the monodentate binary complex to the bidentate binary complex, in comparison with the corresponding step from the 1:1 monomer-"macromolecule" complex to the 2:2 complex (Page and Jencks, 1971). The relationship between the ideal bidentate-binding constant and two monomeric binding constants appears to serve well as a reference point to assess the covalent linkage effect (LePecq and Roques, 1986; Chipman and Sharon, 1969). Here it is noted that the hypothetical 1 M standard state for solutes has to be adopted for this relationship to be valid. If the hypothetical mole fraction unity state is adopted instead, the coefficient has to change from 55.6 into 1 (Chipman and Sharon, 1969).

Therefore, the dimeric lexitropsin ligand L in Fig. 5 has an ideal bidentate-binding constant $k_{11} = [DL]_{close}/[D][L] = \alpha k_{11} = 55.6\ k_1 \times k_2$, under the ideal condition just stated. Here k_1 and k_2 are stepwise binding constants of the monomeric ligand molecule when bound to the oligonucleotide to form a 2:1 WPPW motif. $\alpha = 55.6\ k_1 \times k_2/k_{11} \approx 55.6\ k_2$, because $k_1 \approx k_{11}$. If we take the intermediate DL_2 as a bidentate ligand of the oligonucleotide, its ideal bidentate-binding constant $k_{22a} = [D_2L_2]_{close}/[DL_2][D] = \beta k_{22a} = 55.6\ k_1 \times k_2$. $\beta = 55.6\ k_1 \times k_2/k_{22a} \approx 55.6\ k_2$, because $k_1 \approx k_{22a}$.

When $N = 2$, which is the halfway point in a titration experiment,

$$Z = \frac{N \times \alpha^2 k_{11}}{2\beta k_{21} k_{22a}} = \frac{\alpha^2 k_{11}}{\beta k_{21} k_{22a}} \quad (6)$$

If both α and β are optimized to the ideal value $55.6\ k_2$, $Z = 55.6\ k_2\ k_{11} / k_{21}k_{22a}$. $Z \approx 55.6$ M, because $k_2 \approx k_{21}$ and $k_{11} \approx k_{22a}$. According to Eq. 5, the total concentration of the oligonucleotide has to achieve a physically impossible high value of 55.6 M to have the symmetric D_2L_2 dominant. Therefore, under normal conditions, the bidentate DL 1:1 complex will predominate. This shows that the symmetric D_2L_2 tetraplex cannot compete with the covalently connected WPPW motif when both motifs can be formed in the ideal binding affinity. It follows naturally that the tetraplex will compete even less well, if the binding affinity is not optimized to the ideal value for the tetraplex but the covalently connected 1:1 WPPW motif.

If we assume that β is optimized to the ideal value of 55.6 k_2, a perfectly maximized value, the critical α value, which decides whether a 1:1 bidentate or a 2:2 symmetric complex is going to be predominant, can be obtained from the following procedure. From Eq. 6, $Z = \alpha^2 k_{11}/\beta k_{21} k_{22a} = \alpha^2 k_{11}/55.6 k_2 k_{21} k_{22a} \approx \alpha^2/55.6(k_2)^2$. Substitute this into Eq. 5 and rearrange the equation to give:

$$\alpha\ Vs\ k_2 \times (55.6[DNA]_{total})^{0.5}. \quad (7)$$

In a typical NMR titration, $[DNA]_{total} \approx 1$ mM. Therefore, Eq. 7 can be then converted into: $\alpha\ Vs\ 0.23\ k_2$. That is to say, if the bidentate-binding enhancement α factor exceeds 0.23 k_2, which is only 0.41% of the ideal value 55.6 k_2, the bidentate binding is going to remain dominant. Therefore, some considerable strain or entropy loss can be accommodated without compromising the binding mode.

However, because the value of k_2 is in general rather large, the critical value 0.23 k_2 is not something easily achieved experimentally. For example, k_2 in our dimeric distamycin cases is $\sim 20 \times 10^6$ M^{-1} (Chen and Lown, 1994) (the k_2 constant for imidazole- and pyridine-containing lexitropsins

with the corresponding matched oligonucleotide is likely higher than this value because of the high cooperativity generally observed), and the critical value is calculated to be 4.7×10^6. The bidentate-binding enhancement of this magnitude has not been accomplished experimentally yet (Mrksich and Dervan, 1993b, 1994; Chen and Lown, 1994) (our unpublished data). In short, if the bidentate-binding enhancement does not reach a still rather large value, the symmetric D_2L_2 is going to become dominant. Larger k_2 values will make the dominance more complete. This conclusion is based on the formation of an ideally optimized symmetric tetraplex. Let us assume that β is one million times less than the ideal value ($\sim 1.1 \times 10^9$), the critical α value is derived as 4700 from Eq. 6. This threshold value is still substantial.

Our CD titration conditions have DNA concentration at ~ 8 μM (a ligand is assumed to occupy a 5-base-pair site), whereas the fluorescence experiments have it at 0.1 μM. Assuming β is optimized to the ideal value, critical α values 4.2×10^5 and 4.7×10^4 can be derived, respectively. If β is one million times less than the ideal value, critical α values as low as 4.2×10^2 and 4.7×10 can be derived, respectively. Clearly, highly sensitive techniques enhance the possibility of detecting the closed 1:1 *DL* complex. Other important factors, unique to CD and fluorometry experimental conditions, also facilitate the observation of the 1:1 *DL* complex, which are described further in the following "Reevaluation of CD and Ethidium Fluorometry" section.

If the 1:1 bidentate binding is not available at all (i.e., $\alpha = 0$), the above analysis indicates the extremely favorable formation of a 2:2 symmetric complex. We may also gain some insights into this problem in a different way. The binding isotherm equation can be simplified as:

$$\gamma = \frac{k_{11}[L] + 2k_{11}k_{21}[L]^2 + k_{11}K_{12}[L][D] + 2k_{11}k_{21}K_{22A}[L]^2[D]}{1 + k_{11}[L] + k_{11}k_{21}[L]^2 + 2k_{11}K_{12}[L][D] + 2k_{11}k_{21}K_{22A}[L][D]}.$$

Previous NMR titration experiment showed that DL_2 is the only species observed when L is a monodentate ligand containing an imidazole or pyridine heterocycle (Dwyer et al., 1992; Mrksich et al., 1992). Pyrrolecarboxamide lexitropsins also interact cooperatively with alternating AT sequences in the 2:1 WPPW mode (Chen and Lown, 1994).

Therefore, $k_{11} \ll k_{21}$, $k_{11}[L] \ll 2k_{11}k_{21}[L]^2$, $k_{11}K_{12}[L][D] \ll 2k_{11}k_{21}[L]^2$. We may simply compare the $2k_{11}k_{21}[L]$ and $2k_{11}k_{21}K_{22A}[L]^2[D]$ to understand the binding equilibrium, especially the tendency to form the symmetric 2:2 complex.

$2k_{11}k_{21}[L]^2$ *Vs* $2k_{11}k_{21}K_{22A}[L]^2[D]$ is equivalent to 1 *Vs* $K_{22A}[D]$. Thus, it follows that the relative magnitude of $[D]$ and $1/K_{22A}$ will decide the position of the binding equilibrium.

Because K_{22A} may reach an ideal value of $55.6\ k_1 \times k_2$ under optimized conditions ($\sim 10^{15}$), the concentration of the free DNA can easily exceed the reciprocal of K_{22A}. For example, at the halfway point in a titration again, $[D] = [D]_{total}/2$. $[D]$ *Vs* $1/K_{22A}$ is equivalent to $[D]_{total}$ *Vs* $2/K_{22A}$. Clearly, the tendency to form the symmetric 2:2 complex is very high under normal NMR titration conditions.

REINTERPRETATION OF NMR TITRATION AND QUANTITATIVE FOOTPRINTING EXPERIMENTS

Upon close inspection, one may notice that those structures characterized by Dwyer et al. (1993) under NMR titration conditions cannot account for the small enhancement of apparent binding constants from the quantitative footprinting measurements. If those structures, so close to the "strainless" ideal geometry of the unconnected WPPW motif, were true representations of the bidentate binding motif, binding constants would approximate the theoretical value, $55.6\ k_1 \times k_2$. The binding enhancement would be very large, not a factor of 10.

Moreover, the best-fit binding equation for a true bidentate binding should be first order with respect to the concentration of any free dimeric ligand, not second order as actually obtained (Mrksich and Dervan, 1993b). The second order equation can only be interpreted as the domination of 2:1 stoichiometry under the footprinting conditions, as these authors also indicated. The dominance of 2:1 stoichiometry is not by accident. Imidazole- or pyridine-containing lexitropsin monomers have been observed to interact highly cooperatively in the 2:1 WPPW mode (Dwyer et al., 1992; Mrksich et al., 1992). The corresponding term in the binding isotherm equation is $2K_{11} \times K_{21}[L]^2 = 2k_{11} \times k_{21}[L]^2$. $2k_{11} \times k_{21}$ and, therefore, $2K_{11} \times K_{21}$ are relatively unchanged if the appendage effect on the 1:1 WPW and 2:1 WPPW binding modes is small and similar. Clearly, the insignificance of the bidentate binding, i.e., the small α value and, thus, small contribution from the $K_{11}[L] = (1 + \alpha)k_{11}$ term, for the pentakis(methylene) and hexakis(methylene) linked dimers 3c and 3d and its total absence for the tris(methylene) and tetrakis(methylene) linked dimers 3a and 3b (i.e., $\alpha = 0$ and $K_{11}[L] = (1 + \alpha)K_{11}$), as supported by our experiments (Chen and Lown, 1994), results in the dominance of the 2:1 WPPW term and, therefore, the second order fitting equation experimentally obtained from quantitative footprinting. The highly cooperative 2:1 WPPW interaction in effect obscures the experimentally observable signal from the 1:1 bidentate interaction completely. As a result, the apparent binding constants, obtained as the square root of the coefficients of the $[L]^2$ term, are approximately the square root of $2K_{11} \times K_{21}$ and should be relatively unchanged. Not surprisingly, the apparent binding constants obtained as such are not related to the bidentate-binding mode.

The dominance of the 2:1 WPPW binding mode for these imidazole-containing dimeric lexitropsins under quantitative footprinting conditions logically leads to further aggregation under NMR titration conditions where, as we know, concentrations are much higher. Therefore, we can conclude that the covalently connected WPPW motif characterized by the NMR techniques are actually the unconnected WPPW motif in symmetric cyclic *n*:*n* complexes instead of the true WPPW 1:1 bidentate-binding motif (Dwyer et al., 1993). We suggest further that these symmetric complexes are likely of 2:2 stoichiometry. Model inspection suggests that other higher

symmetric polymers are more rigid geometrically and, accordingly, less capable of escaping from charge repulsion of outer phosphate groups. Therefore, the stoichiometry factor would strongly favor formation of the lower aggregation state, the symmetric 2:2 complex. These considerations may be substantiated in detail by undertaking a binding equilibrium analysis again. Fig. 6 shows the inclusion of a symmetric 3:3 closed complex in a binding model.

The binding isotherm equation is as follows:

$$\gamma = (K_{11}[L] + 2K_{11}K_{21}[L]^2 + K_{11}K_{12}[L][D] + 2K_{11}K_{21}K_{22A}[L]^2[D] + 3K_{11}K_{21}K_{22A}K_{32}[L]^3[D] + 2K_{11}K_{21}K_{22A}K_{23}[L]^2[D]^2, + 3K_{11}K_{21}K_{22A}K_{32}K_{33C}[L]^3[D]^2)/(1 + K_{11}[L] + K_{11}K_{21}[L]^2 + 2K_{11}K_{12}[L][D] + 2K_{11}K_{21}K_{22A}[L]^2[D] + 2K_{11}K_{21}K_{22A}K_{32}[L]^3[D] + 3K_{11}K_{21}K_{22A}K_{23}[L]^2[D]^2, + 3K_{11}K_{21}K_{22A}K_{32}K_{33C}[L]^3[D]^2).$$

To understand the competition between the symmetric 2:2 and 3:3 closed complexes, we only need to compare the corresponding terms $2K_{11}K_{21}K_{22A}[L]^2[D]$ and $3K_{11}K_{21}K_{22A}K_{32}K_{33A}[L]^3[D]^2$ in the numerator of the binding isotherm equation. Following the similar procedure used earlier, we can derive Eq. 8:

$$[D]_{total}\ Vs\ 8N \times (\beta + 1)^3/9\delta^2k_2 \qquad (8)$$

where $[D]_{total} > 8N \times (\beta + 1)^3/9\delta^2k_2$ if the symmetric 3:3 complex is dominant; $[D]_{total} < 8N \times (\beta + 1)^3/9\delta^2k_2$ if the symmetric 2:2 complex is major. Here $N = [D]_{total}/[L]_{total}$, the same as previously defined. Because of geometric rigidity and consequent charge repulsion for the symmetric 3:3 complex, β is likely to be larger than δ. Assuming $N = 2$, which represents the halfway point of a titration, $k_2 = 20 \times 10^6\ M^{-1}$ and $[D]_{total} = 1$ mM, typical of a NMR titration concentration, β should be around 10^4 to allow the symmetric 2:2 complex to predominate according to Eq. 8. If one imagines DL_2 species of tris(methylene) to hexakis(methylene)-linked dimeric lexitropsins as "bidentate ligands," the effective cross-linker length for such "bidentate ligands" at least equals that of two tris(methylene) chains, plus the distance between N-methyl groups of two central pyrrole rings in a 2:1 WPPW antiparallel side-by-side binding motif. This effective length should be similar to that of the octakis(methylene) chain (Chen and Lown, 1994). Our recent cross-linkage optimization studies suggest that dimeric lexitropsins with cross-linkers of the approximate length of the octakis(methylene) chain can achieve a binding constant enhancement over 10^4 (Y.-H. Chen and J. W. Lown, unpublished data). Therefore, it is not unreasonable to assume that the binding constant enhancement β is around 10^4, and the symmetric 2:2 complex should prevail over the symmetric 3:3 complex.

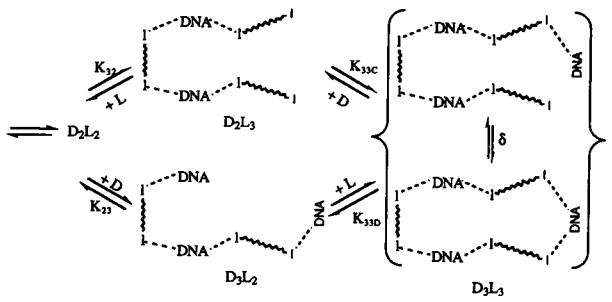

FIGURE 6 A binding model that includes a symmetric 3:3 complex. See Fig. 5 for the early portion of the complete binding equilibrium.

In short, under NMR titration conditions, the bidentate binding is not strong enough to overcome the cooperativity factor, because of the shortness of the cross-linker. The initially formed 2:1 dimeric lexitropsin-oligonucleotide complex renders the effective linkage of two unbound moieties much longer, leading to a genuine bidentate binding to a second oligonucleotide. The experimental proof of such a complex may come from molecular weight determination such as gel permeation chromatography, light scattering, and ultracentrifugation, or by fitting the dilution curve numerically using a sensitive monitoring parameter. By including consideration of other possible structures accounting for NMR spectra, correcting the interpretation of quantitative footprinting experiments and, thus, establishing the consistency between NMR and quantitative footprinting data, we can resolve all apparent differences between two independent sets of conclusions.

REEVALUATION OF CD AND ETHIDIUM FLUOROMETRY EXPERIMENTS

In light of the facile formation of higher aggregates, we have to reexamine our CD and fluorometry experiments. We have excluded the possibility of complexation between a dimeric lexitropsin ligand and two polymeric DNA duplexes, by analogy with the behavior of dimeric intercalators (Chen and Lown, 1994). Reexamination of previous experimental data shows that the polymeric aggregation effect does not interfere in our CD and fluorescence experiments. For example, the tetrakis(methylene) chain-connected dimer 4a should be the most likely candidate for cross-complexation because the short linker disfavors the bidentate binding and provides an opportunity for cross-complexation without competition. Its alternating AT polymer titration curve is remarkably similar to that of the monomer, indicating no sign of cross-complexation.

Two factors may account for this lack of observable cross-complexation. First, CD and fluorescence experiments are carried out under much lower concentrations in terms of actual binding sites in comparison with the NMR titration, which shifts the binding equilibrium into the low aggregation state. Second, the DNAs used in CD and fluorescence experiments are much longer and contain many identical bind-

ing sites, in contrast to single-sited DNAs in the NMR titration experiments. This has a series of consequences. There is a rather strict distance constraint for paired sites suitable for cross-complexation, which renders the majority of sites available only for other binding modes without such a constraint, even if two polymers are perfectly aligned. If a cross-complexation takes place at one site, two connected polymers would adopt a crossover orientation rather than a parallel one to avoid charge repulsion between outer phosphate groups. This crossover orientation should reduce the chance of further cross-complexation. Even if the parallel orientation is adopted because of another cross-complexation, the imperfect match of two polymers would also reduce the possibility of cross-complexation simply because of statistical reasons.

It is evident that both CD titration curves and ethidium fluorometry with the alternating AT DNA polymer are very responsive to the linker change and relatively free of the aggregation complication. This responsiveness also stems from the weak binding cooperativity to the polymer by tricarboxamide lexitropsins in the 2:1 WPPW mode, as demonstrated from NMR studies (Fagan and Wemmer, 1992) and from intrinsic binding constant measurements (Chen and Lown, 1994). This weak cooperativity facilitates the manifestation of both 1:1 WPW and bidentate-binding modes under both sets of experimental conditions.

A corollary of the above analysis is that the cross-complexation of two single-sited DNA duplexes by a dimeric lexitropsin with a linker ranging from tris(methylene) to heptakis(methylene) is highly favored under NMR titration conditions, if only the WPW mode is available for the binding interaction. In contrast, our recent CD titration experiments showed no sign of cross-complexation essentially when the homo AT DNA polymer was used, which allows only the WPW mode for binding interaction (Y.-H. Chen and J. W. Lown, unpublished data). Thus, the application of the homo AT polymer for CD experiments is free of aggregation complication.

In summary, using weakly cooperative lexitropsin-DNA interacting systems and experimental conditions close to those of CD and fluorometry experiments has a decisive advantage over highly cooperative interacting systems and experimental conditions requiring higher concentrations and single-sited oligonucleotides as in the NMR studies.

THE SIGNIFICANCE OF 2:2 OLIGONUCLEOTIDE-PEPTIDE COMPLEXES AND A CAUTIONARY NOTE ABOUT BIOPHYSICAL AND BIOCHEMICAL CHARACTERIZATION OF DESIGNED SYSTEMS

The formation of dimeric ligand-oligonucleotide 2:2 complexes is interesting in its own right, and this may actually represent, to our knowledge, the first example. The crossover of two DNA molecules mediated by peptides is structurally reminiscent of some DNA recombination intermediates, D-loop structures, and transient intermediates in the three-dimensionally nonrandom search for a specific site on a large DNA molecule by a protein. One remarkable feature of these now inferred 2:2 complexes is that the linkage can be quite short. How do DNA duplexes position relative to each other? A parallel orientation would introduce some charge repulsion when the linkage is relatively short and therefore a perpendicular orientation should be preferred in such a case. There are some other interesting questions to be answered. Perhaps the further cross-linkage of these dimeric ligands provides a novel class of tetrameric lexitropsins with very strong cross-complexation ability, which serve as useful structural probes of DNA as well as simple models of DNA crossover structures. There was no significant 2:2 complex formation under quantitative footprinting conditions where a large DNA fragment and lower concentrations of DNA and ligands were used, as suggested from the experimentally derived second order binding equations with correlation coefficients over 97% (Mrksich and Dervan, 1993b). This means that a stronger cross-complexation interaction is required for the formation of a DNA crossover structure under such conditions. It seems very likely that dimeric intercalators and major groove binders can be designed and elaborated to cross-complex with two DNA duplexes.

In summary, sensitive techniques like CD and fluorescence have their roles to play in elucidating molecular interactions despite the fact that they are indirect and nonvisual. High dilution conditions eliminate some structural "artifacts" that would not appear under biological conditions. After all, in vivo processes often take place under relatively low concentration conditions. A powerful technique like NMR spectroscopy has its limitations in distinguishing symmetric structures, and the insensitivity of this technique will continue to affect binding studies wherever equilibria exist. Therefore, the combination with other sensitive hydrodynamic and spectroscopic techniques is highly desirable in certain situations. Structural studies should become more meaningful when correlated with careful thermodynamic analysis. The apparent binding constant formalism should be treated with due caution because the constant might be a sort of averaging over intrinsic binding constants, and this average value is meaningless if one is not aware of the possibility of mutual cancellation of the intrinsic binding constants in certain situations. This cautionary note also demonstrates a possibly general situation in molecular engineering where the molecular system has to be optimized to a high threshold point observe the intended function with some conventional techniques, even though the design feature is essentially correct. Therefore, investigation of marginally optimized systems, although informative, can be challenging, and one should proceed with caution.

We thank Professors R. S. Brown and D. L. J Clive for helpful discussions. Professor S. Withers in the Department of Chemistry at the University of British Columbia provided useful clues to some of our analysis.
The Natural Sciences and Engineering Research Council of Canada financially supported this research program via a grant (to J. W. Lown).

REFERENCES

Adamson, A. W. 1954. A proposed approach to the chelate effect. *J. Am. Chem. Soc.* 76:1578–1579.

Chen, Y.-H., and J. W. Lown. 1994. A new DNA minor groove binding motif: cross-linked lexitropsins. *J. Am. Chem. Soc.* 106:6995–7005.

Chipman, D. M., and N. Sharon. 1969. Mechanism of lysozyme action. *Science.* 165:454–465.

Dwyer, T. J., B. H. Geierstanger, Y. Bithini, T. W. Lown, and D. E. Wemmer. 1992. Design and binding of a distamycin A analog to d(CGCAAGTTGGC) · d(GCCAACTTGCG): synthesis NMR studies and implications for the design of sequence-specific minor groove binding oligopeptides *J. Am. Chem. Soc.* 114:5911–5919.

Dwyer, T. J., B. H. Geierstanger, M. Mrksich, P. B. Dervan, and D. E. Wemmer. 1993. Structural analysis of covalent peptide dimers, bis-(pyridine-2-carboxamidonetropsin)$(CH_2)_{3-6}$, in complex with 5′-TGACT-3′ sites by two-dimensional NMR. *J. Am. Chem. Soc.* 115: 9900–9906.

Fagan, P., and D. E. Wemmer. 1992. Cooperative binding of distamycin-A to DNA in the 2:1 mode. *J. Am. Chem. Soc.* 114:1080–1081.

Geierstanger, B. H., T. J. Dwyer, Y. Bathini, J. W. Lown, and D. E. Wemmer. 1993. NMR characterization of a heterocomplex formed by distamycin and its analog 2-ImD with d(CGCAAGTTGGC) · d(GCCAACTTGCG): the 1:1:1 heterocomplex is favored over either 2:1 homocomplex *J. Am. Chem. Soc.* 115:4474–4482.

Kopka, M. L., C. Yoon, D. Goodsell, P. Pjura, and R. E. Dickerson. 1985. Binding of an antitumor drug to DNA: netropsin and CGCGAATT-^{Br}CGCG. *J. Mol. Biol.* 183:553–563.

LePecq, J.-B., and B. P. Roques. 1986. DNA binding and biological properties of bis and tris intercalating molecules. *In* Mechanisms of DNA Damage and Repair: Implication for Carcinogenesis and Risk Assessment. M. G. Simic, L. Grossman, and A. G. Upton, editors. Plenum Press, New York. 219–244. [Note that Eq. 3 of that article should be of second order with respect to the monomeric binding constant to ensure consistency between Eq. 3 and 2. This is likely just a typographical error.]

Lown, J. W. 1988. Lexitropsins: rationale design of DNA sequence reading agents as novel anticancer agents and potential cellular probes. *Anticancer Drug Des.* 3:25–40.

Lown, J. W. 1990. Molecular mechanisms of DNA sequence recognition by groove binding ligands: biochemical and biological consequences. *In* Molecular Basis of Specificity in Nucleic Acid-Drug Interactions. B. Pullman, and J. Jortner, editors. Kluwer Academic Publishers, Dordrecht, The Netherlands. 103–122.

Lown, J. W. 1992. Lexitropsins in antiviral development. *Antiviral Res.* 17:179–196.

Lown, J. W. 1993. Targeting the DNA minor groove for control of biological function: progress, challenges, and prospects. *Chemtracts-Org. Chem.* 6:205–237.

Mrksich, M., and P. B. Dervan. 1993a. Antiparallel side by side heterodimer for sequence-specific recognition of DNA by a distamycin/1-methylimidazole-2-carboxamide-retrospin pair. *J. Am. Chem. Soc.* 115: 2572–2576.

Mrksich, M., and P. B. Dervan. 1993b. Enhanced sequence specific recognition in the minor groove of DNA by covalent peptide dimers: bis(pyridine-2-carboxamidonetropsin)$(CH_2)_{3-6}$. *J. Am. Chem. Soc.* 115:9892–9899.

Mrksich, M., and P. B. Dervan. 1994. Design of a covalent peptide heterodimer for sequence-specific recognition on the minor groove of double helical DNA. *J. Am. Chem. Soc.* 116:3663–3664.

Mrksich, M., W. S. Wade, T. J. Dwyer, B. H. Geierstanger, D. E. Wemmer, and P. B. Dervan. 1992. Antiparallel side by side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc. Natl. Acad. Sci. USA.* 89:7586–7590.

Page, M. I., and W. P. Jencks. 1971. Entropic contribution to rate acceleration in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. USA.* 68:1678–1683. [Earlier discussions on the chelate effect can be found in articles by von G Schwarzenbach (1952) and A. W. Adamson (1954).]

Pelton, J. G., and D. E. Wemmer. 1989. Structural characterization of a 2:1 distamycin A :d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. USA.* 86:5723–5727.

Pelton, J. G., and D. E. Wemmer. 1990. Binding modes of distamycin A with $d(CGCAAATTTGCG)_2$ determined by two-dimensional NMR. *J. Am. Chem. Soc.* 112:1393–1399.

Schwarzenbach, von G. 1952. Der Chelateffekt. *Helv. Chim. Acta.* 291: 2344–2359.

Zimmer, C., and U. Wahnert. 1986. Non-intercalating DNA-binding ligands: specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. *Prog. Biophys. Mol. Biol.* 47:31–112.