Year 2 Report: 1508070

Our NSF application had two important objectives: (1) synthesize a series of *C,O*-aryl glycoside derivatives based on the nogalamycin core structure and (2) to evaluate the nucleic acid binding affinity and sequence of these *C*-glycoside derivatives using UV and fluorescence spectroscopy, as well as isothermal calorimetry. In addition we wished to assess the ability of our ligands to inhibit the binding of the transcription factor AP-1 to its consensus sequence.

Cell-permeable small molecules that bind to specific regions of DNA in the promoters or coding regions of genes hold the promise of being able to regulate gene expression/readout. Essential characteristics of such molecules are both high affinity and the ability to bind locations on DNA in which regulatory proteins bind. Since the major groove is the site of binding for most DNA-associating proteins, we have been exploring the development of major-groove binding ligands. Based on our previous results, we envisioned that ligands capable of recognizing the major groove will contain a major-groove targeting moiety and a suitably functionalized carbohydrate moiety capable of engaging in base-specific hydrogen-bonding contacts.

While our original ligand scaffold included morindone, an aromatic moiety reminiscent of the nogalamycin chromophore, we have subsequently changed our aromatic moiety to the azinomycin B chromophore (Figure 1). The reasons behind this modification are twofold: first, the azinomycin chromophore (a substituted naphthalene derivative) is much more easily prepared and stable than the anthraquinone morinodone. The synthesis of derivatives of substituted naphathlenes is relatively straightforward. Second, the azinomycin chromophore has been shown to intercalate the backbone of duplex DNA, directing attached moieties into the major groove. We have now verified this propensity in our own studies of derivatives of this chromophore.

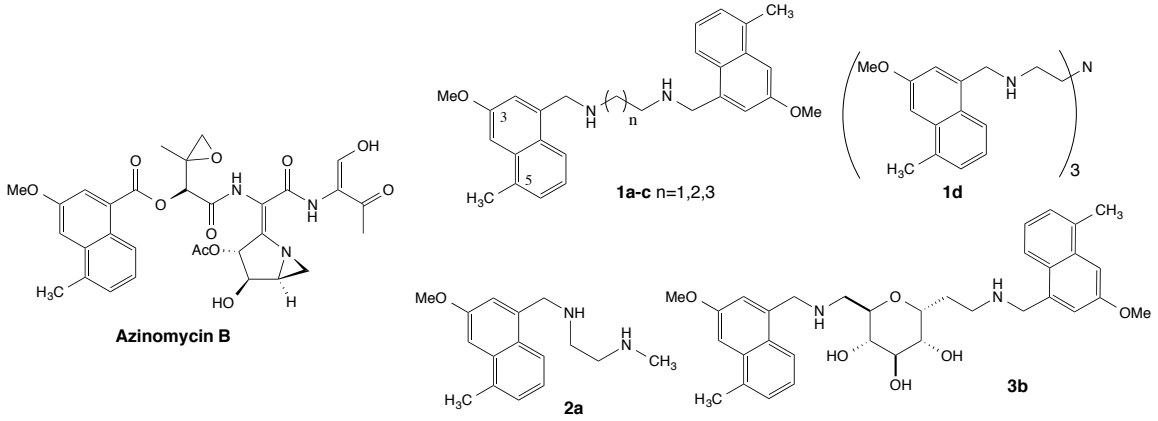


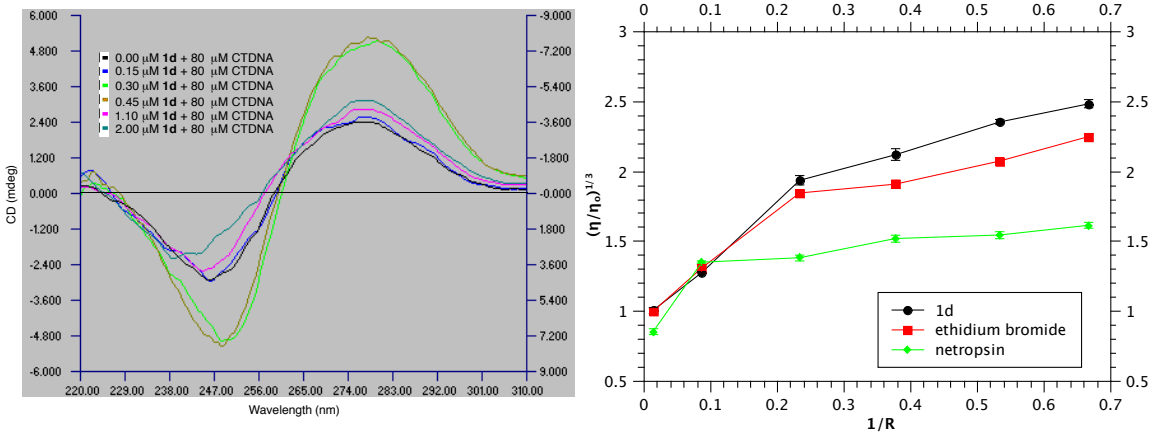
Figure 1*. Azinomycin Chromophore Derivatives.*

Dimeric and trimeric derivatives of the azinomycin chromophore (**1a-1d**, **2a, 3b** Figure 1) have been prepared from the corresponding naphthaldehyde and appropriate diamines and triamines by reductive amination chemistry. These compounds display binding affinities in the $10^6\,M^{-1}$ – $10^7\,M^{-1}$ range for CT DNA (Table 1) and also intercalate the backbone of DNA, as verified by viscosity and circular dichroism experiments (Scheme 1). The tightest binding derivative is trimer **1d** (K_d=2.42x10^7 M^{-1}), an observation that may be rationalized based on the strong electrostatic interaction between the DNA phosphate backbone and the 4 protonated amines of **1d** at physiological pH. Indeed reduced binding to CT DNA was observed for all azinomycin derivatives in the presence of high salt concentrations (1 M NaCl).

***Table 1**. Apparent association constants and binding ratios of ligands **1a-1d**, **2a**, and **3b** for CT DNA.*

Ligand	K_{app} (x10^6 M^{-1})[a]	r_{bd}[b]
1a	7.19±0.92	4.2
1b	3.15±0.60	6.2
1c	5.00±0.46	7.0
1d	**24.2±1.5**	**9.1**
2a	0.47±0.03	4.0
3b	6.47±0.65	10.0

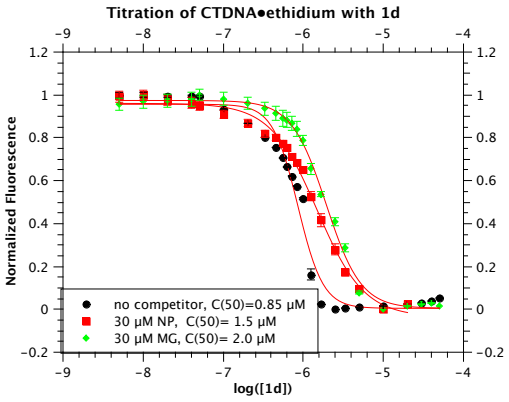
Scheme 1.** Evidence for intercalation (CD, viscosity) by tightest binding ligand **1d



That the azinomycin chromophore projects attached moieties into the major groove was established by competition binding studies with methylated lambda phage DNA and with major and minor groove binding ligands. Displacement of ethidium bromide from CT DNA was studied in the presence and absence of methyl green (MG, a major groove binder) and netropsin (NP, a minor grove binder). In the presence of MG, a 2.5-fold drop in the binding affinity of **1d** for DNA was observed; in the presence of NP, a more modest 1.7-fold drop was observed (Scheme 2). Furthermore, the binding affinity of **1d** for lambda phage DNA was 3-fold lower than that for CT DNA; in contrast, the binding affinity of NP was higher for lambda phage DNA than for CT DNA (Table 2). These observations support the notion that the azinomycin chromophore positions attached moieties in the major groove of DNA.

***Table 2.** Comparison of the binding of **1d** with various polynucleotides*

DNA	K_{app} (x10^6 M^{-1})
polydG•polydC	18.4±0.6
Calf thymus	24.2±1.5
poly(dA•dT)$_2$	17.7±1.5
polydA•polydT	20.0±2.5
Lambda phage	8.02±0.7



***Scheme 2.** Competition DNA binding studies of **1d** with MG and NP*

As can be seen in Table 2, these ligands display poor DNA sequence discrimination. Most notably, ligand **3b**, containing a carbohydrate linker, does not show enhanced DNA binding relative to ligands with aliphatic linkers (**1a**-**c**). This may indicate that any stabilization due hydrogen-bond interactions between the glucosyl moiety linking the azinomycin chromophores and the DNA bases is offset either by steric crowding or entropic restriction of bond rotation in the intercalated complex. Clearly, the nature of the linker region connecting the chromophore moieties is crucial for obtaining sequence selectivity. Molecular docking studies of ligand **1d** with duplex DNA (5'-ATGCAT-3') show an energy-minimized bis-intercalation complex with the chromophore tethering moiety occupying the major groove, and ligand amino groups engaging in hydrogen bonding contacts with the edges of the sandwiched GC base pairs (Figure 2). We are currently replacing the aliphatic tethering moieties in these ligands with arginine and glutamine amino acids to assess if anticipated side chain interactions with the edges of the G and A bases lead to sequence selective binding (Scheme 3).

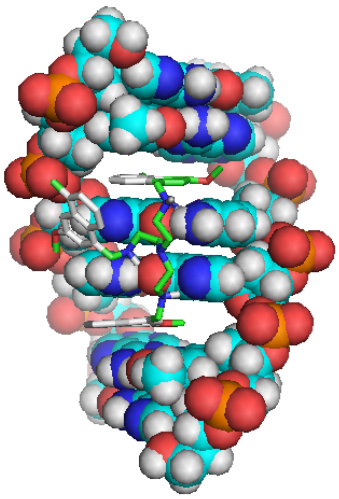
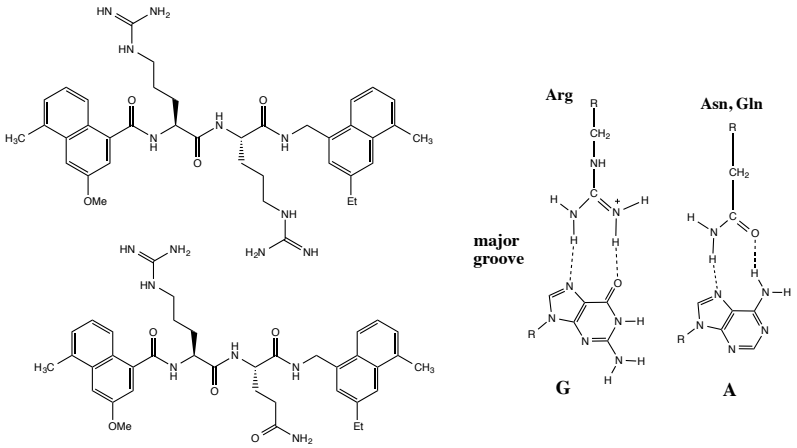
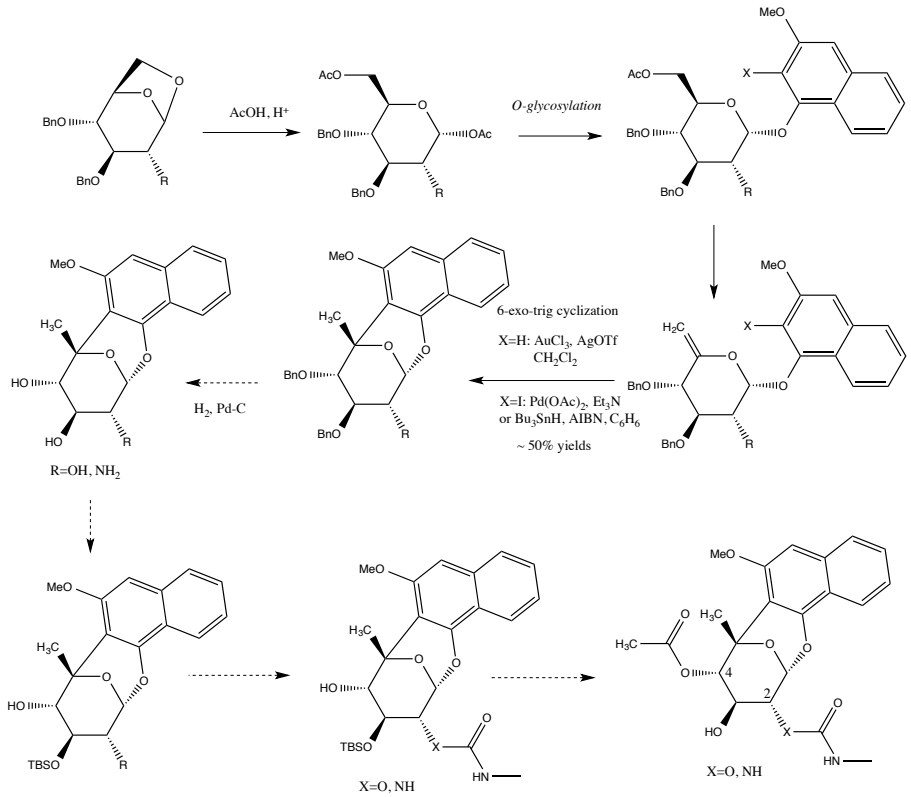


*Figure 2. A **1d**/DNA complex*



***Scheme 3**. Current synthetic ligands based on the azinomycin chromophore*.

The unconstrained sugar of ligand **3b** may adopt multiple conformations when bound to DNA, leading to low sequence selectivity and binding affinity. To resolve this entropic problem, the construction of a *C,O*-glycoside would provide a rigid, wedge-shaped scaffold that can present carbohydrate functional groups directly to the edges of the G/C and A/T bases in the major groove. Some synthetic difficulties are attendant in the synthesis of C,O-glycosides, however. Our current approach involves the preparation of an α-*O*-glycoside, elimination of the C.6 hydroxyl group to form the methylene carbohydrate, and then 6-exo-trig cyclization promoted by either transition metal catalysts or by the generation of an aryl carbon radical. Yields for this cyclization are typically in the 50% range. Subsequent manipulations are then required to remove protecting groups and introduce the requisite functionality at C.2 and C.4 of the carbohydrate for molecular recognition of the DNA bases. The overall route is quite lenghthy, and regioselective protection and deprotection of the carbohydrate moiety after *C,O*-glycoside formation has been difficult to achieve in high yields thus far.



Scheme 4. *Progress toward synthesis of azinomycin C, O-glycosides.*

We anticipate that once these protecting group issues are resolved, dimerization of the ligands will be straightforward using iminodiacetic acid.

Another molecular platform for crafting major-groove binding ligands is the chromophore of the well-known DNA stains methyl green and crystal violet. These triaryl methanes associate principally with the major groove because their steric requirements prohibit binding in the narrow minor groove. Furthermore, these dyes do not intercalate the backbone of DNA, as verified by viscosity experiments. Dimers and trimers of these dyes are currently being synthesized containing piperazine (**9**), ethane (**10**), propane (**11**), butane (**12**), and triamine (**13**) linkers. Since these are non-intercalating ligands, the cardiotoxic effects observed with other DNA intercalators could be potentially minimized. Compounds **9**, **11**, and **12** display submicromolar binding affinities for duplex DNA, and piperazine compound **9** has an association constant of $4.5 \times 10^7 M^{-1}$ (K_d=22 nM). Once again, it may be envisioned that the attachment of arginine and glutamine amino acids moieties to this chromophore (**14**) may lead to sequence-specific DNA binding due to hydrogen-bonding interactions with the edges of the DNA base pairs. Utilizing fluorescence and absorption spectroscopy, as well as CD spectroscopy to assess the binding characteristics of our synthetic compounds, we hope to develop a paradigm for the sequence-specific recognition of DNA in the major groove by small organic molecules. Such a paradigm would assist the creation of novel gene therapy agents for a wide range of disease states.

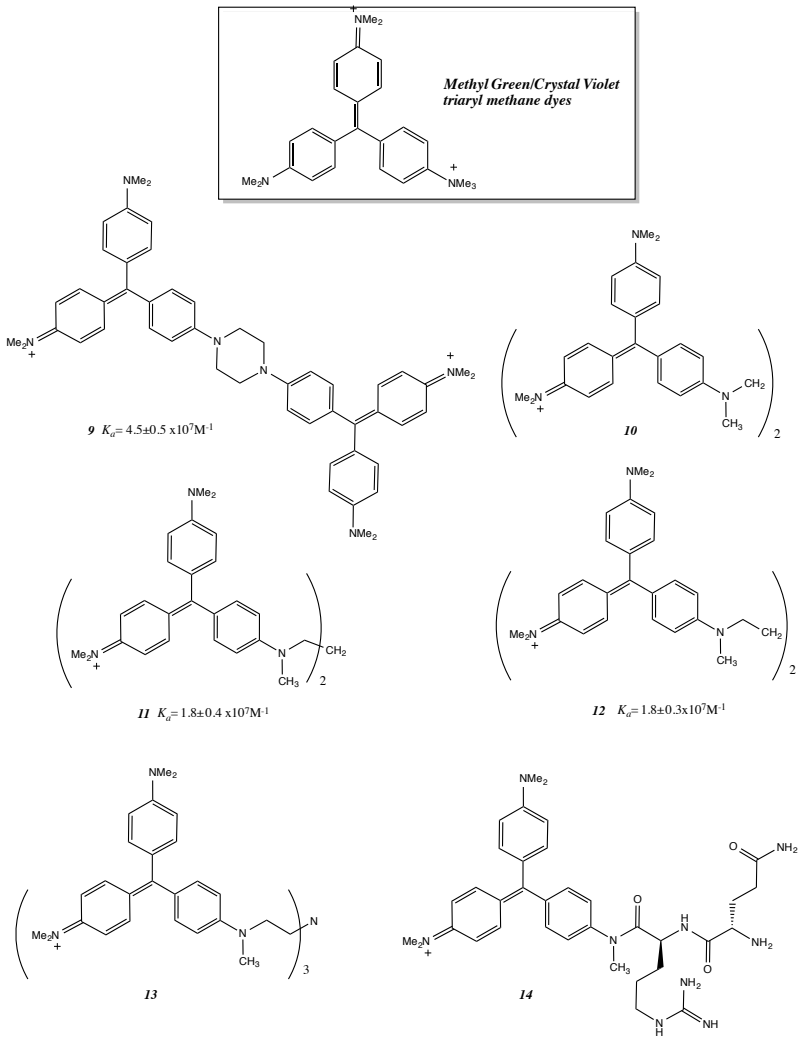


Figure 3*. Derivatives of triarylmethane dyes for molecular recognition in the major groove of DNA.*