Total Synthesis of Distamycin A and 2640 Analogues: A Solution-Phase Combinatorial Approach to the Discovery of New, Bioactive DNA Binding Agents and Development of a Rapid, High-Throughput Screen for Determining Relative DNA Binding Affinity or DNA Binding Sequence Selectivity

Dale L. Boger,* Brian E. Fink, and Michael P. Hedrick

Contribution from the Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received November 30, 1999. Revised Manuscript Received March 2, 2000

Abstract: The development of a solution-phase synthesis of distamycin A and its extension to the preparation of 2640 analogues are described. Thus, solution-phase synthesis techniques with reaction workup and purification employing acid/base liquid−liquid extractions were used in the multistep preparation of distamycin A (8 steps, 40% overall yield) and a prototypical library of 2640 analogues providing intermediates and final products that are ≥95% pure on conventional reaction scales. The complementary development of a simple, rapid, and high-throughput screen for DNA binding affinity based on the loss of fluorescence derived from displacement of prebound ethidium bromide is disclosed which is applicable for assessing relative or absolute binding affinity to DNA homopolymers or specific sequences (hairpin oligonucleotides). Using hairpin oligonucleotides, this method permits the screening of a library of compounds against a single predefined sequence to identify high affinity binders, or the screening of a single compound against a full library of individual hairpin oligionucleotides to define its sequence selectivity. The combination permits the establishment of the complete DNA binding profile of each member of a library of compounds. Screening the prototypical library provided compounds that are 1000 times more potent than distamycin A in cytotoxic assays (**67**, $IC_{50} = 29$ nM, L1210), that bind to poly[dA]−poly[dT] with comparable affinity, and that exhibit an altered DNA binding sequence selectivity. Several candidates were identified which bound the five-base-pair AT-rich site of the PSA-ARE-3 sequence, and one (**128**, $K = 3.2 \times 10^6$ M^{-1}) maintained the high affinity binding ($K = 4.5 \times 10^6$ M^{-1}) to the ARE-consensus sequence containing a GC base-pair interrupted five-base-pair AT-rich site suitable for inhibition of gene transcription initiated by hormone insensitive androgen receptor dimerization and DNA binding characteristic of therapeutic resistant prostate cancer.

The regulation of gene expression is based on the sequence selective recognition of nucleic acids by a host of repressor, activator, and enhancer proteins. Selective external disruption or control of such processes has been a long standing goal of molecular biology, chemistry and medicine. Small molecules that are able to selectively bind DNA and activate (block a repressor) or inhibit (block an activator) gene expression hold great promise as therapeutics. However, the discovery of such agents has been slow due in part to the biophysical complexity associated with understanding small molecule-DNA interactions, the iterative process of designing and synthesizing individual molecules toward specific DNA sequences, and the technically demanding techniques involved in determination of their selectivity and binding affinity for a given sequence. Herein we report the combinatorial synthesis of a prototypical library of 2640 analogues of distamycin A[1,2] and the development of a rapid, technically nondemanding screen for the determination of DNA binding affinity and sequence selectivity applicable to such libraries which should help accelerate the discovery process.

To date, few small molecule libraries have been disclosed that are designed to identify new DNA binding agents.[3,4] Using distamycin A as a lead structure (Figure 1), we have developed solution-phase combinatorial strategies based on those we previously reported[5] to systematically replace the *N*-methylpyrrole subunit with other heterocyclic amino acids to produce a prototypical first generation library in a small mixture format. This first generation library was further functionalized with a basic side chain to mimic the amidine group of distamycin A.

A variety of techniques are commonly used to investigate the DNA binding properties of small molecules.[6] However, most approaches are technically challenging and time-consuming, rendering them inapplicable to high-throughput screening. The

(1) Arcamone, F.; Penco, P.; Orezzi, P.; Nicolella, V.; Pirelli, A. *Nature* **1964**, *203*, 1064.

(2) Johnson, D. S.; Boger, D. L. DNA Binding Agents. In *Comprehensive Supramolecular Chemistry*; Vol. 4, J.-M. Lehn, Series Ed., Y. Murakami, Vol. Ed., Pergamon Press: Oxford, England, 1996; pp 73−176 and references therein.

(3) Shipps, G. W.; Pryor, K. E.; Xian, J.; Skyler, D. A.; Davidson, E. H.; Rebek, J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11833. Lescrinier, T.; Hendrix, C.; Kerremans, L.; Rozenski, J.; Link, A.; Samyn, B.; Van Aerschot, A.; Lescrinier, E.; Eritja, R.; Van Beeumen, J.; Herdewijn, P. *Chem. Eur. J.* **1998**, *4*, 425. Klekota, B.; Hammond, M. H.; Miller, B. L. *Tetrahedron Lett.* **1997**, *38*, 8639.

(4) Behrens, C.; Nielsen, P. E. *Comb. Chem. High Throughput Screening* **1998**, *1*, 127. Guelev, V. M.; Harting, M. T.; Lokey, R. S.; Iverson, B. L. *Chem. Biol.* **2000**, *7*, 1.

(5) Cheng, S.; Comer, D. D.; Williams, J. P.; Myers, P. L.; Boger, D. L. *J. Am. Chem. Soc.* **1996**, *118*, 2567. Boger, D. L.; Tarby, C. M.; Myers, P. L. Caporale, L. H. *J. Am. Chem. Soc.* **1996**, *118*, 2109. Cheng, S.; Tarby, C. M.; Comer, D. D.; Williams, J. P.; Caporale, L. H.; Myers, P. L.; Boger, D. L. *Bioorg. Med. Chem.* **1996**, *4*, 727.

10.1021/ja994192d CCC: $19.00 © 2000 American Chemical Society
Published on Web 06/20/2000

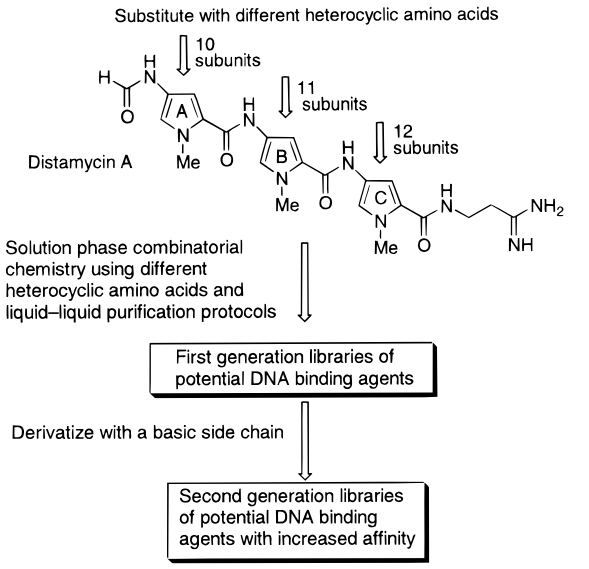


Figure 1. Solution-phase strategy for DNA binding agent libraries.

commonly used DNAse I footprinting is able to give base pair resolution of preferred DNA binding sequences, but suffers from the sequence specificity of DNAse I itself.[7] Methods such as Dervan's MPE footprinting and affinity cleavage overcome this problem but require precisely controlled reaction conditions or the derivatization of the DNA binding agents themselves.[8] The recently reported RESPC technique also alleviates some of the commonly encountered problems in footprinting but require the preparation of multiple DNA constructs.[9] Most techniques require the knowledge of specialized biochemical procedures and assay reproducibility comes only with experience. None are applicable to high-throughput screening required for assaying a library of compounds against a library of sequences. We have developed a technically nondemanding procedure for rapidly determining DNA binding selectivity entailing the competitive displacement of prebound ethidium bromide[6,10−12] from defined hairpin oligonucleotides (Figure 2). DNA of interest (homopolymers, heteropolymers, or predefined hairpin oligonucleotides) in 96-well plates is treated with ethidium bromide, yielding a large fluorescence increase upon DNA intercalation. Addition of a nonfluorescent DNA binding agent results in a decrease in fluorescence due to displacement of bound ethidium bromide. The decrease in % fluorescence is directly related to the extent of DNA binding providing relative DNA binding affinities and, through subsequent quantitative titration, is capable of providing accurate absolute binding constants.

As detailed herein, this technique may be used to screen a library of compounds for DNA binding to a single DNA sequence or for the complementary screening of a single compound against a full library of DNA sequences which results in the definition of the sequence specificity of a given agent. Combining these in the assay of a library of compounds against a library of DNA, provides qualitative and/or quantitative information on the binding of all library members against a library of available sequences, allowing complete characterization of the DNA binding profiles of each agent in a single experiment.

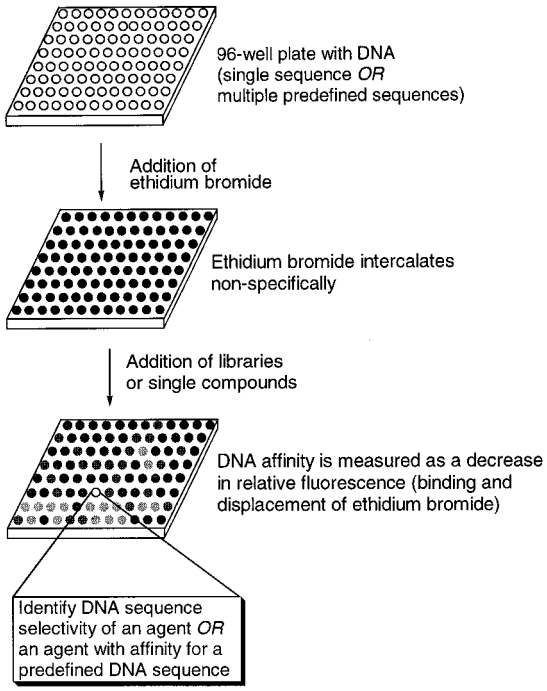


Figure 2. General procedure for the rapid DNA binding screen adaptable for determination of the sequence selectivity of a library of DNA binding agents.

Solution-Phase Total Synthesis of Distamycin A. As an initial demonstration of the approach, we first conducted a total synthesis of distamycin A utilizing solution-phase synthesis techniques that require only acid/base liquid−liquid extraction purification protocols. Previous total syntheses of distamycin A generally have relied on the coupling of *N*-methyl-4-nitropyrrole-2-carboxylic acid chlorides followed by nitro group reduction and further coupling steps.[1,13] In developing a general set of reaction conditions suitable for the preparation of libraries, several requirements not intrinsic to the natural product synthesis needed to be addressed. In all cases, unreacted starting materials, coupling agents, and their reaction byproducts needed to be removed by simple acid/base extraction. Although acid chlorides could be used, the preparation and long-term storage of numerous heterocyclic acid chlorides would be difficult, requiring the implementation of coupling protocols that use the carboxylic acids directly. In addition, a nitro group reduction step introduces a reaction of variable generality (reaction time, catalyst poisoning), precludes the inclusion of subunits sensitive to reduction conditions, and requires the resultant free amines to be relatively stable. Consequently, we adopted the more direct use of BOC-protected amines. A general set of coupling conditions that gives high yields of coupled product was developed enlisting the amines and carboxylic acids directly and the water soluble 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDCI) with (dimethylamino)pyridine (DMAP) as an additive. The unreacted starting materials,

(6) (a) *Drug-DNA Interactions Protocols*; Fox, K. R., Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90. (b) Jenkins, T. C. Optical Absorbance and Fluorescence Techniques for Measuring DNA-Drug Interactions. In *Drug-DNA Interactions Protocols*; Fox, K. R. Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90, p 195.

(7) Drew, H. R.; Travers, A. A. *Cell* **1984**, 37, 491.

(8) Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5470.

(9) Hardenbol, P.; Wang, J. C.; Van Dyke, M. W. *Bioconjugate Chem.* **1997**, *8*, 616.

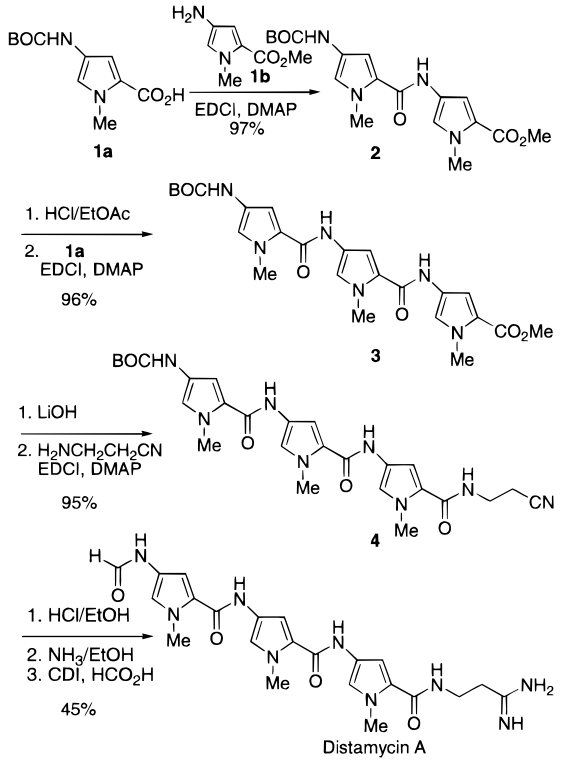
(10) Morgan, A. R.; Lee, J. S.; Pulleyblank, D. E.; Murray, N. L.; Evans, D. H. *Nucleic Acids Res.* **1979**, *7*, 547.

(11) Baguley, B. C.; Falkenhaug, E.-M. *Nucleic Acids Res*. **1978**, *5*, 161.

(12) Boger, D. L.; Invergo, B. J.; Coleman, R. S.; Zarrinmayeh, H.; Kitos, P. A.; Thompson, S. C.; Leong, T.; McLaughlin, L. W. *Chem.−Biol. Interact*. **1990**, *73*, 29. Boger, D. L. Sakya, S. M. *J. Org. Chem*. **1992**, *57*, 1277.

(13) Lown, J. W.; Krowicki, K. *J. Org. Chem*. **1985**, *50*, 3774. Grehn, L.; Ragnarsson, U. *J. Org. Chem*. **1981**, *46*, 3492. Bialer, M.; Yagen, B.; Mechoulam, R. *Tetrahedron* **1978**, *34*, 2389.

Scheme 1



reagents, and reaction and reagent byproducts all may be removed by acid/base extractions.

Starting with the pyrrole carboxylic acid **1a**,[14] coupling with aminopyrrole **1b**[14] using EDCI/DMAP afforded **2** in high yield (97%), Scheme 1. Removal of the BOC-protecting group with HCl/EtOAc followed by coupling to pyrrole **1a** afforded the tripeptide **3** in good yield (96%). Saponification of **3** followed by coupling with β-aminopropionitrile afforded nitrile **4** in excellent yield (95%). Treatment of nitrile **4** with HCl/EtOH followed by NH_3/EtOH afforded the desired amidine with concomitant removal of the BOC group (Scheme 1). Due to the intrinsic instability of this free amine, it was immediately treated with *N*-formyl imidazole to afford distamycin A. This provided distamycin A in 40% overall yield for eight steps without deliberate optimization and required only acid/base liquid−liquid extraction to afford all intermediates and the final product with >95% purity as demonstrated by their ^{1}H NMR spectra (Supporting Information Figure 1).

Library Design. Two prototypical libraries of potential DNA binding agents were prepared in a small mixture format, Figure 3. Using eleven *N*-BOC heterocyclic amino acids and twelve amino esters, the individual subunits were coupled using EDCI/DMAP to provide all possible 132 individual dipeptides in parallel. The use of EDCI and DMAP allows for the removal of excess coupling agents and their reaction byproducts along with unreacted starting materials by acid/base liquid−liquid extraction. These individual dimers were deprotected and coupled to a mixture of 10 *N*-BOC carboxylic acids to give 132 mixtures of 10 *N*-BOC-trimers where only the last position (subunit A) is undefined (1320 compounds). Removal of the BOC group and coupling to the basic side chain, *N*,*N*-dimethylaminobutyric acid (DMABA), affords an analogous DMABA-trimer library (1320 compounds). The amidine group found in distamycin A was replaced with a dimethylamino group in order to avoid the variable yielding Pinner reaction and for

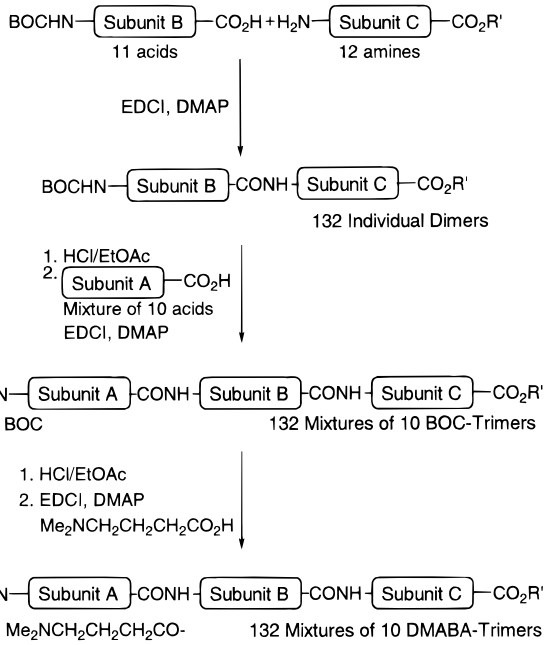


Figure 3. Reaction sequence for preparation of DNA binding agent libraries.

overall ease of synthesis. The decision to place the basic side chain at the *N*-terminus rather than the *C*-terminus resulted from observation of inefficiencies during hydrolysis of certain *C*-terminal monomer subunits (see below). Comparison studies detailed in a following section determined that these changes had little effect on the DNA binding affinities of the resultant agents.

This strategy of preparing all individual dimers offered the advantage of examining the reactivity of all possible combinations of acids and amines to ensure that the coupling conditions and purification protocols were general and provided the levels of purity desired. Since the reactions are carried out in solution, each of 132 dimers were characterized by ^{1}H NMR (see Supporting Information), purity was established by conventional techniques, and 50−100 mg quantities were accessible. This allows for the preparation of numerous second generation libraries from stocks of dimers and for the deconvolution of libraries from stored and archived samples of all library components.

The heterocyclic amino acids selected for the first prototypical libraries are shown in Figure 4. Included in this set are the pyrrole, imidazole, and thiazole amino acids, studied by Dervan and Lown, and the indole and CDPI amino acids studied in our laboratories. While not intended to be a survey of optimal heterocyclic amino acids, this set provides built-in known DNA binding agents and could be expected to address issues of identification, practicality, and viability which proved to be especially useful in comparisons with positional scanning libraries.[15]

The subunits **1**, **5**, **6**, **9**, **10**, and **13** were prepared according to known procedures.[14−17] The preparation of the remaining subunits is detailed in Supporting Information Schemes S1−S4 from readily available materials following established procedures and proceeding through intermediates **16**−**29**.[18−26]

(14) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

(15) Boger, D. L.; Dechantsreiter, M. A.; Ishii, T.; Fink, B. E.; Hedrick, M. P. *Bioorg. Med. Chem*. In press.

(16) Nishiwaki, E.; Tanaka, S.; Lee, H. *Heterocycles* **1988**, *27*, 1945.

(17) Boger, D. L.; Coleman, R. S.; Invergo, B. J. *J. Org. Chem*. **1987**, *52*, 1521. Boger, D. L.; Yun, W.; Han, N. *Bioorg. Med. Chem.* **1995**, *3*, 1429.

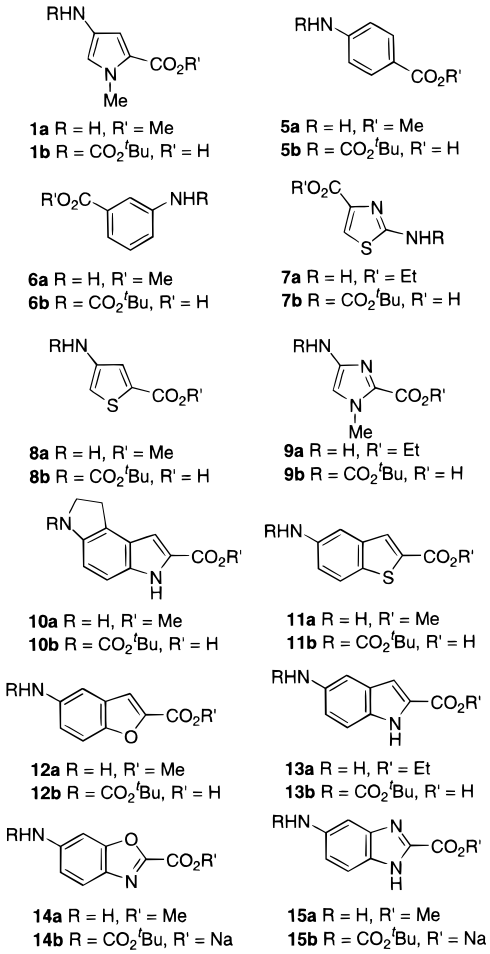


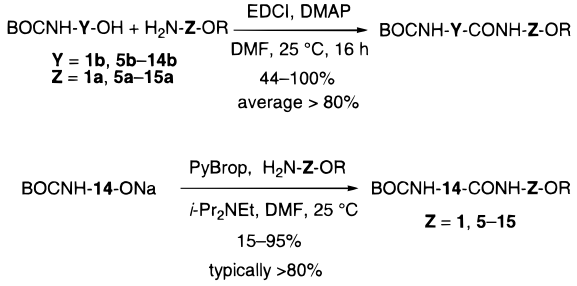
Figure 4. Amino acid subunits used in the preparation of libraries.

Ester hydrolysis of the precursors to both **14** and **15** proceeded efficiently with NaOH in MeOH, but attempts to isolate the free carboxylic acids led to decarboxylation. However, the sodium salts **14b** and **15b** were isolated in quantitative yield without detectable decarboxylation and used effectively in our efforts.

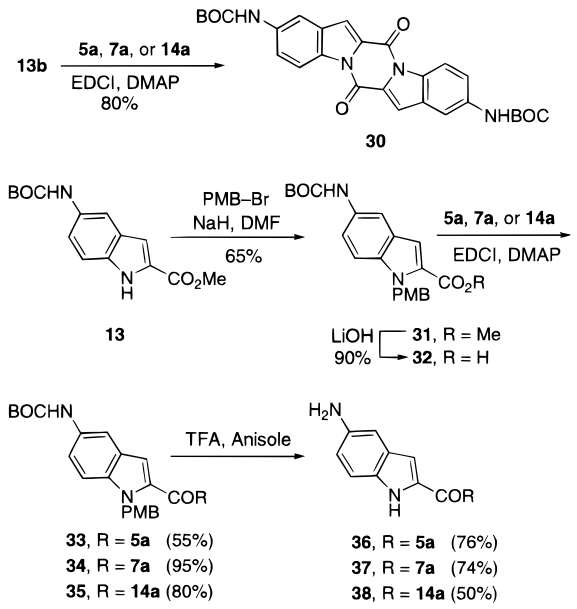
Parallel Synthesis of Dimers. Using **1b** and **5b**−**13b** as the acid component and **1a** and **5a**−**15a** as the amine component, 120 individual dimers were prepared (Scheme 2). Each dimer was prepared in 70−80 mg quantities in parallel, using only acid/base liquid−liquid extraction purification to afford products in typically >80% yield and >95% purity.

Incorporating the benzoxazole **14** into the dimers presented a unique problem as attempts to isolate the free acid (**14**, R′ = H) led to decarboxylation. It was possible to isolate the sodium salt (**14b**, R = Na), but the reaction conditions used to prepare

Scheme 2



Scheme 3

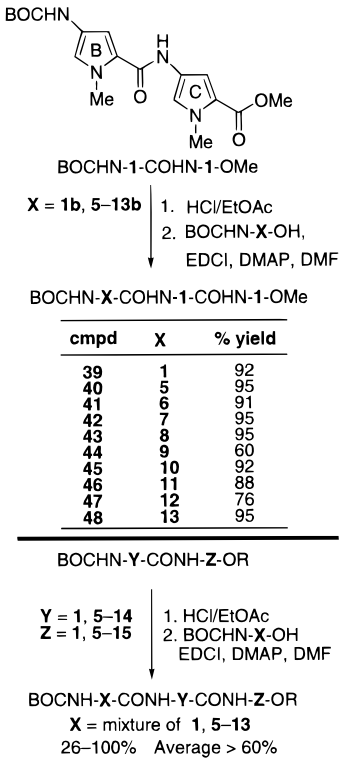


the individual dimers in Scheme 2 (EDCI, DMAP) led to complex reaction mixtures and significant amounts of the decarboxylated benzoxazole. However, PyBrop was found to work well in their preparation and was used in lieu of the standard conditions (Scheme 2). In three instances, when the indole subunit **13b** was used as the acid component in couplings with the three unreactive amines (**5a**, **7a**, and **14a**), the diketopiperazine **30** was isolated due to indole dimerization (Scheme 3). Although not the topic of the present work, the properties of the indole diketopiperazines have proved extraordinary, providing potent cytotoxic agents displaying effective DNA binding properties in their own right.[27] To circumvent this problem, the indole nitrogen was protected with a *p*-methoxybenzyl group to afford indole **31**. Hydrolysis to afford the free acid **32** and coupling to the three individual amines afforded the desired dipeptides in moderate yield. Simultaneous deprotection of both the *p*-methoxybenzyl and BOC-protecting groups (TFA/anisole, 60 °C) afforded the desired amines.[28]

Synthesis of BOC-trimer Libraries. The preparation of the trimer libraries was investigated initially by preparing several sets of individual trimers to ensure the reaction conditions were appropriate. For example, the preparation of all 10 individual trimers from the BOCNH-**1**−CONH-**1**−OMe dimer is given in Scheme 4. Deprotection of the dimer with HCl/EtOAc followed by coupling with **1b**, **5b**−**13b** afforded the 10 BOC-

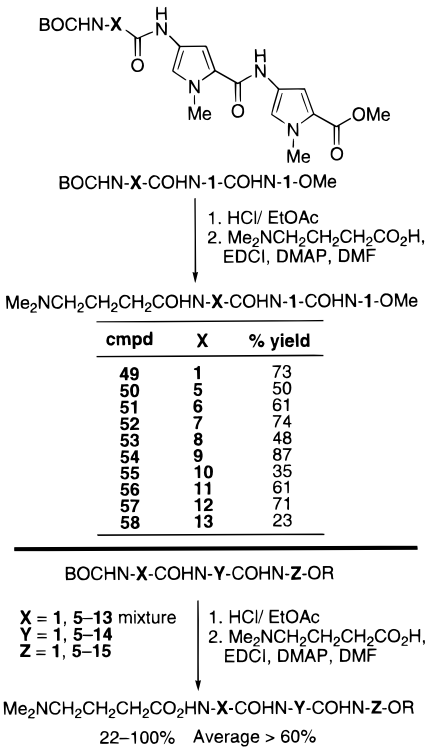
(18) Sprague, J. M.; Lincoln, R. M.; Ziegler, C. *J. Am. Chem. Soc*. **1946**, *68*, 266.
(19) Foye, W. O.; Hefferren, J. J.; Feldman, E. G. *J. Am. Chem. Soc*. **1954**, *76*, 1378. Shih, C.; Gossett, J. S.; Worzalla, J. F.; Rinzel, S. M.; Grindey, G. B.; Harrington, P. M.; Taylor, E. C. *J. Med. Chem*. **1992**, *35*, 1109.
(20) Osuga, H.; Suzuki, H.; Tanaka, K. *Bull. Chem. Soc. Jpn*. **1997**, *70*, 891.
(21) Van Wijngaarden, I.; Kruse, C. G.; van der Heyden, J. A. M.; Tulp, M. T. M. *J. Med. Chem*. **1988**, *31*, 1934.
(22) Moller, H. *Liebigs Ann. Chem*. **1971**, *749*, 1.
(23) Crivello, J. V. *J. Org. Chem*. **1981**, *46*, 3056.
(24) Bistrzycki, A.; Przeworski, G. *Chem. Ber*. **1912**, *45*, 3483.
(25) Boger, D. L.; Yun, W.; Cai, H.; Han, N. *Bioorg. Med. Chem.* **1995**, *3*, 761.
(26) Rastogi, R.; Sharma, S.; Iyer, R. N. *Ind. J. Chem., Sect. B* **1979**, 464.
(27) Boger, D. L.; Fink, B. E.; Hedrick, M. P. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 4269.
(28) We were unable to prepare dimers where benzimidazole **15b** was the acid component. Therefore, this monomer was only used in the third position (C) of the trimers.

Scheme 4



cmpd	X	% yield
39	1	92
40	5	95
41	6	91
42	7	95
43	8	95
44	9	60
45	10	92
46	11	88
47	12	76
48	13	95

Scheme 5



cmpd	X	% yield
49	1	73
50	5	50
51	6	61
52	7	74
53	8	48
54	9	87
55	10	35
56	11	61
57	12	71
58	13	23

trimers in high yield and with >90% purity using only acid/base liquid−liquid extraction purification. This set based on the dipyrrole dimer is of special interest because it contains a close analogue of distamycin, the tripyrrole **39** (BOCNH-**1**−CONH-**1**−CONH-**1**−OMe).

Having established the conditions for coupling and workup, each of the individual dipeptides was converted to a mixture of 10 tripeptides (Scheme 4). Removal of the BOC group with anhydrous HCl, followed by coupling to an equimolar mixture of 10 acids afforded the BOC-trimer mixtures on a 50 μmol scale. An excess of the amine component was used to ensure complete consumption of the 10 acids in the reaction mixture. Full matrix mixtures analyzed by mass spectrometry ensured all expected components were present (see Supporting Information). Since the benzoxazole **14b** did not couple efficiently under the standard conditions and required an additional purification step, it was omitted from the first position (subunit A) to ensure library fidelity and purity.

Dimethylaminobutyric Acid Libraries. In a similar manner, optimization of reaction conditions to incorporate a dimethylaminobutyric acid side chain (DMABA) was carried out on individual trimers (Scheme 5). The yields were found to vary more widely than that of the previous steps since some of the derivatives showed appreciable water solubility. Thus, the typical acid/base liquid−liquid purification protocol was modified. Simply removing the solvent from the reactions, followed by suspension of the products in water and extraction with EtOAc gave the desired products.

By using the conditions described for the individual compounds, each of the BOC-trimer mixtures was converted to the corresponding DMABA-trimer mixture as shown in Scheme 5. Full matrix mixtures analyzed by mass spectrometry ensured all components were present (see Supporting Information). With the DMABA-trimer library, 2640 distamycin analogues were available in the format of two small mixture libraries.

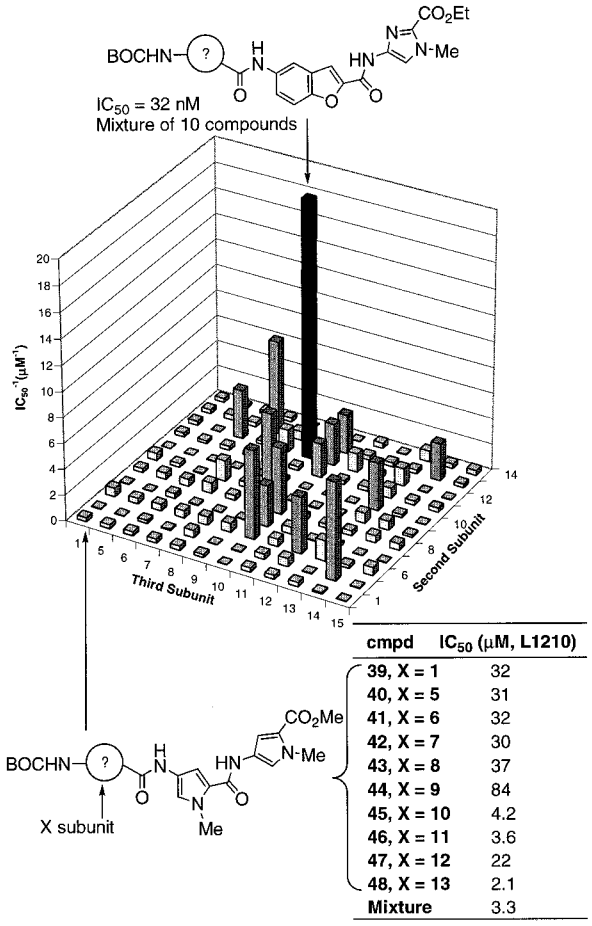
Cytotoxic Activity. In addition to our interest in the DNA binding properties of the compounds, we were also interested in comparing the behavior of the libraries in functional assays. In particular, we were interested in comparing the performance of the library format of 10 compound mixtures versus larger mixture testing required of positional scanning.[15] Consequently, the two libraries were examined in a functional assay for cytotoxic activity (L1210).[29] While all of the BOC-trimer mixtures showed some activity in the cell-based assay (IC_{50} < 10 μM), thirteen showed activity at less than 1 μM, and one showed activity at less than 100 nM (Figure 5). The most active library contained the benzofuran subunit (**12**) at the central position (subunit B) and the imidazole subunit (**9**) at the final position (subunit C).

This mixture was deconvoluted by resynthesis of the 10 components, beginning with the stored BOCNH-**12**−CONH-**9**−OEt dimer (Table 1). This resynthesis from the immediate precursor required a single day and provided 5 mg samples of each component. A second round of testing revealed that the most active components contained either the benzofuran (**12**) or the benzothiophene (**11**) at the first (A) position, with IC_{50}'s of 29 nM and 68 nM for **66** and **67**, respectively (Table 1). When compared to distamycin A (IC_{50} = 42 μM), both are 1000 times more potent. For comparison purposes, **39**−**48**, which contain the dipyrrole subunits of distamycin and its close tripyrrole analogue **39**, were also examined in the L1210 assay (Figure 5). Consistent with the behavior of distamycin A, **39** was essentially inactive (IC_{50} = 32 μM) and approximately 1000 times less potent than **66** and **67**.

In contrast, the IC_{50} values for the DMABA-trimer mixtures were found to be on the order of 10−100 fold higher than the BOC-trimer libraries (Figure 6).[30] The most active mixture contains the CDPI subunit (**10**) in the final position (subunit C), and the thiophene subunit (**8**) at the central position (subunit B). This mixture was deconvoluted by synthesis of the 10

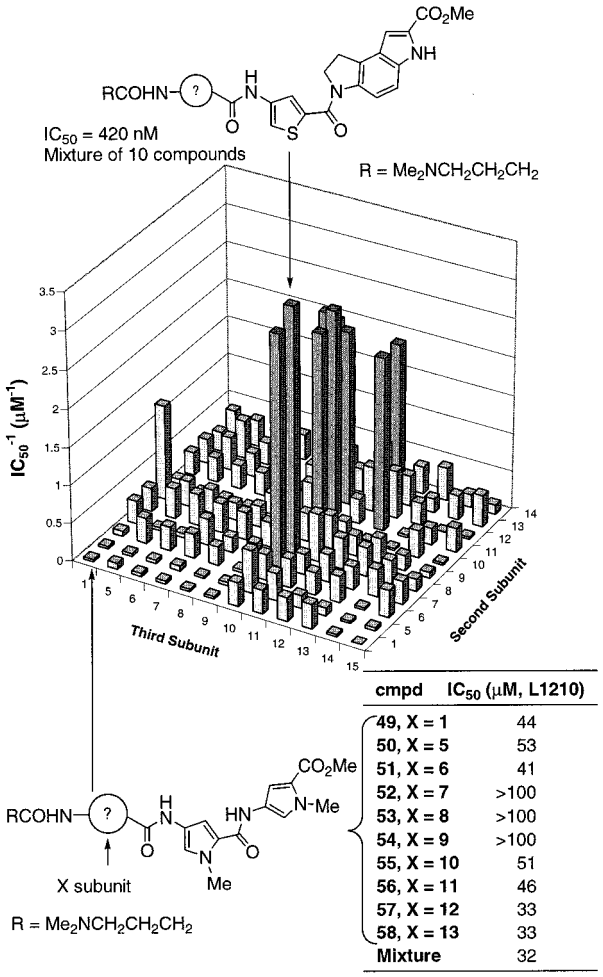
(29) Boger, D. L.; Mullican, M. D.; Mitscher, L. A.; Drake, S.; Kitos, P. A. *J. Med. Chem.* **1985**, *28*, 1543.

(30) This may be the result of decreased cell penetration of a charged species.



cmpd	IC_{50} (μM, L1210)
39, X = 1	32
40, X = 5	31
41, X = 6	32
42, X = 7	30
43, X = 8	37
44, X = 9	84
45, X = 10	4.2
46, X = 11	3.6
47, X = 12	22
48, X = 13	2.1
Mixture	3.3

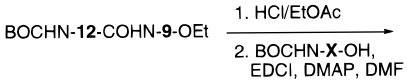
Figure 5. Cytotoxicity (L1210) for BOC-trimer libraries expressed as I/IC_{50} for clarity.



cmpd	IC_{50} (μM, L1210)
49, X = 1	44
50, X = 5	53
51, X = 6	41
52, X = 7	>100
53, X = 8	>100
54, X = 9	>100
55, X = 10	51
56, X = 11	46
57, X = 12	33
58, X = 13	33
Mixture	32

Figure 6. Cytotoxicity (L1210) for DMABA-trimer libraries expressed as I/IC_{50} for clarity.

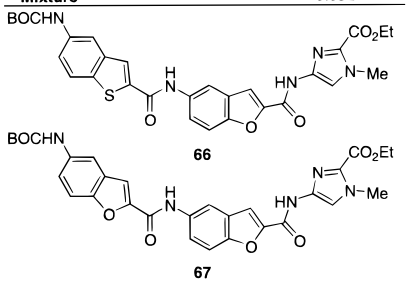
Table 1



BOCHN-**X**-COHN-**12**-COHN-**9**-OEt

Table 1

cmpd	X	% yield	IC_{50} (μM, L1210)
59	**1**	42	0.30
60	**5**	52	2.7
61	**6**	68	>100
62	**7**	54	0.34
63	**8**	89	1.5
64	**9**	65	1.8
65	**10**	55	1.3
66	**11**	63	0.069
67	**12**	69	0.029
68	**13**	72	1.4
Mixture			0.032



components, beginning from the BOCNH-**8**−CONH-**10**−OMe dimer (Table 2). A second round of screening revealed that the most active component of this library contained the benzothiophene (**11**) at the first position (subunit A), with an IC_{50} of 0.46 μM for **86** and it was >10 times more active than any other compound in the mixture and 100 times more potent than any of the individual components of the **49**−**58** mixture based on and including the close distamycin analogues (Figure 6). The tripyrrole analogue **49** exhibited an IC_{50} of 44 μM indistinguishable from that of distamyin A (42 μM) and 100 times less potent than **86**.

In the instances examined, including additional mixtures that were deconvoluted while examining the DNA binding properties of the agents (see Table 3), the activity of the mixtures in the cell-based assay approximated that of the individual components and established the reliability of testing in the small mixture format for the libraries.

The individual BOC-dimers were also tested in the L1210 functional assay (Supporting Information Figure S2). Nearly all the members were essentially inactive (IC_{50} > 1 μM) with the exception of those containing the B subunit **6** and the bicyclic heterocycles **10**−**15** in the final position (subunit C), five of which exhibited IC_{50}'s < 1 μM. The most potent, BOCNH-**6**−CONH-**10**-OMe, exhibited superb activity with an IC_{50} = 28 nM being >1000 times more active than distamycin or its dipyrrole analogue BOCNH-**1**−CONH-**1**−OMe.

DNA Binding Studies. The most interesting opportunities for us lie with the establishment of the DNA binding properties of the library members. Although a variety of techniques are commonly used to investigate the DNA binding properties of small molecules, most are technically challenging and time-consuming, making them inapplicable to high-throughput screening. However, one technique entailing the competitive displacement of prebound ethidium bromide[6,10−12] does represent

Table 2.

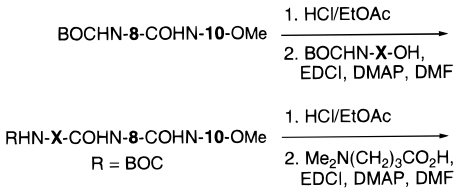
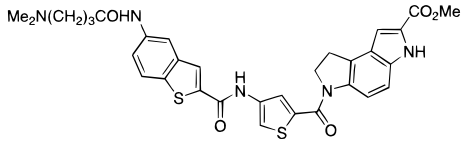


Table 2

X	yield (R = BOC)	IC_{50} (μM, L1210)	yield (R = DMABA)	IC_{50} (μM, L1210)
1	**69**, 48%	0.82	**79**, 98%	220
5	**70**, 48%	25	**80**, 80%	230
6	**71**, 34%	2.3	**81**, 99%	420
7	**72**, 39%	1.9	**82**, 99%	200
8	**73**, 40%	7.5	**83**, 49%	20
9	**74**, 40%	45	**84**, 79%	64
10	**75**, 51%	0.80	**85**, 95%	>1000
11	**76**, 49%	35	**86**, 95%	0.46
12	**77**, 54%	67	**87**, 90%	7.6
13	**78**, 24%	27	**88**, 88%	>1000
Mixture		2.7		0.42



86

a potentially useful high-throughput assay when used in conjunction with a 96-well fluorescent plate reader (Figure 2). Ethidium bromide yields a large fluorescence increase upon DNA intercalation. Addition of another nonfluorescent DNA binding agent results in a decrease in fluorescence due to displacement of bound ethidium bromide. The procedure provides a rapid, flexible, and reliable indication of the relative binding affinities of a wide variety of DNA binding ligands. This technique was first examined as a rapid screen for binding to poly[dA]−poly[dT] and poly[dG]−poly[dC] and subsequently extended to hairpin oligonucleotides containing unique sequences. Using a single agent concentration, the relative decrease in % fluorescence is proportional to the affinity of a given mixture or individual compound for a particular DNA sequence. In addition, we have found that the well-defined linear reduction in fluorescence upon titration with agents related to distamycin can be used to establish absolute binding constants.[6,12] This ability to provide both relative and absolute binding constants enlisting a technically nondemanding assay provides a powerful screening complement to the library synthesis.

Poly[dA]−Poly[dT] and Poly[dG]−Poly[dC]. The binding results for the DMABA-trimer library with poly[dA]−poly[dT] (Figure 7A) showed several general trends: (1) all of the DMABA-trimers induce some decrease in fluorescence, indicating the libraries have an overall AT affinity; (2) high affinity libraries contain one of the larger subunits at the second (B) position (monomers **10**−**14**); and (3) the smaller subunits at the third (C) position (monomers **1, 5**−**9**) appear to be more active. Notably, four of the 10 compound mixtures showed a higher affinity than the pyrrole sublibrary containing **49**, the tripyrrole analogue of distamycin. The highest affinity mixture contains the CDPI subunit (**10**) at the second position and the imidazole subunit (**9**) at the third position. The second most effective mixture contains the benzothiophene (**11**) at the second position and the pyrrole (**1**) at the third position.

Table 3

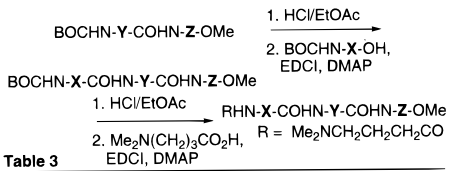
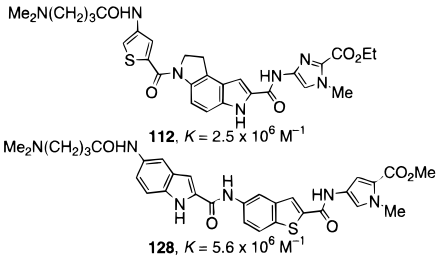


Table 3

X	% yield (R = BOC)		% yield (R = $Me_2NCH_2CH_2CH_2CO$)	
	Y = 10, Z = 9	Y = 11, Z = 1	Y = 10, Z = 9	Y = 11, Z = 1
1	**89**, 60	**99**, 74	**109**, 78	**119**, 46
5	**90**, 66	**100**, 81	**110**, 67	**120**, 22
6	**91**, 63	**101**, 77	**111**, 99	**121**, 79
7	**92**, 82	**102**, 96	**112**, 61	**122**, 45
8	**93**, 50	**103**, 81	**113**, 83	**123**, 38
9	**94**, 46	**104**, 78	**114**, 78	**124**, 12
10	**95**, 83	**105**, 95	**115**, 50	**125**, 64
11	**96**, 90	**106**, 80	**116**, 74	**126**, 18
12	**97**, 82	**107**, 78	**117**, 68	**127**, 36
13	**98**, 94	**108**, 68	**118**, 42	**128**, 78
	IC_{50} (L1210, μM)			
1	**89**, 26	**99**, >100	**109**, >100	**119**, 32
5	**90**, 30	**100**, 1.7	**110**, 32	**120**, 4.7
6	**91**, 35	**101**, 0.13	**111**, 32	**121**, 32
7	**92**, 33	**102**, 0.24	**112**, 34	**122**, 3.3
8	**93**, 17	**103**, 0.65	**113**, 25	**123**, 32
9	**94**, 9.5	**104**, 6.3	**114**, 48	**124**, 32
10	**95**, 0.44	**105**, 0.55	**115**, 32	**125**, >100
11	**96**, 17	**106**, 32	**116**, 50	**126**, 33
12	**97**, 1	**107**, 1.3	**117**, 7.8	**127**, 3.2
13	**98**, 37	**108**, 2.6	**118**, 33	**128**, 32
Mixture	5.0	3.1	32	3.3
	% fluorescence at 10 μM (Poly[dA]−Poly[dT])			
	49 29	**109** 74	**119** 61	
	50 100	**110** 79	**120** 78	
	51 70	**111** 73	**121** 34	
	52 71	**112** 33	**122** 66	
	53 25	**113** 67	**123** 83	
	54 78	**114** 46	**124** 49	
	55 72	**115** 74	**125** 56	
	56 53	**116** 51	**126** 91	
	57 67	**117** 37	**127** 72	
	58 54	**118** 35	**128** 9	



112, $K = 2.5 \times 10^6$ M^{-1}

128, $K = 5.6 \times 10^6$ M^{-1}

Both were deconvoluted through parallel synthesis of the 10 individual components (Table 3) and their assay revealed that **112** and **128** showed the highest affinity for poly[dA]−poly[dT]. Quantitative ethidium bromide titration of **112** and **128** afforded binding constants of 2.5×10^6 M^{-1} and 5.6×10^6 M^{-1}, respectively. The latter agent proved essentially indistinguishable from the tripyrrole analogue of distamycin A (**49**, $K = 5.9 \times 10^6$ M^{-1}). Similarly, **79**−**88** which constitute the individual compounds of the mixture that exhibited the most potent cytotoxic activity of the DMABA-trimers were also examined and the results are recorded in Figure 7.[31] Consistent with its cytotoxic activity, **86** exhibited effective binding to poly[dA]−poly[dT], Table 3

The DMABA-trimers were also screened for binding to poly[dG]−poly[dC] (Figure 7B). As expected, the affinity is much lower than for poly[dA]−poly[dT]. The BOC-trimer libraries were also screened for DNA binding to both poly[dA]−poly[dT] and poly[dG]−poly[dC] (data not shown), and they showed substantially lower affinity as expected.

(31) In those instances where the monomer **10** was present in the mixture, longer incubation times were required due to insolubility (see Supporting Information).

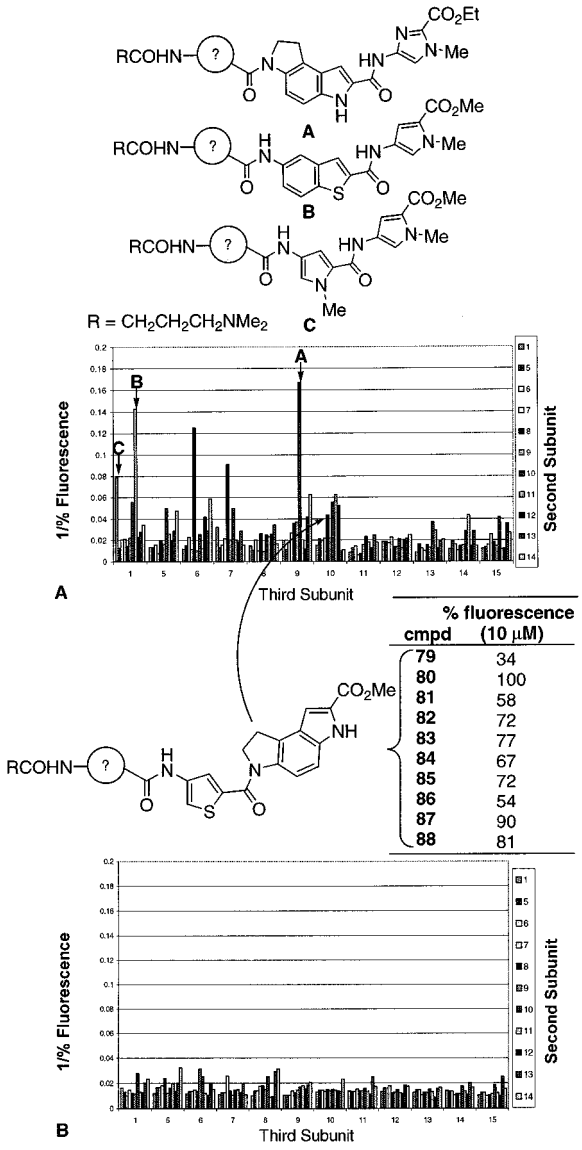


Figure 7. Ethidium bromide displacement assay for DMABA-trimer libraries (100 μM) expressed as 1/% Fluorescence. **A** poly[dA]−poly-[dT]. **B** poly[dG]−poly[dC]. DNA at 0.88×10^{-5} M, ethidium bromide at 0.44×10^{-5} M. Larger numbers indicate higher affinity for DNA.[31]

Defined Sequence Within a Hairpin Oligonucleotide. Although general trends may be detected by examining the binding characteristics with homopolymer DNAs, the most useful information is derived by examining their binding at defined sequences. The extension of the rapid screening to such individual sequences is illustrated with two hairpin oligonucleotides containing two related sequences of the androgen response element, the 14-base pair ARE-consensus[32] and PSA-ARE-3[33] sequences (Figure 8). The emergence of hormone-independent, constituently active androgen receptor dimer, unresponsive to competitive antagonist treatment, is responsible for prostate cancer relapse resistant to chemotherapeutic treatment.[34] A potentially effective treatment for such resistant prostate cancer could entail administration of a DNA binding agent selective for both the PSA-ARE-3 and ARE-consensus sequences that would competitively inhibit the androgen receptor DNA binding and its transcription activation.

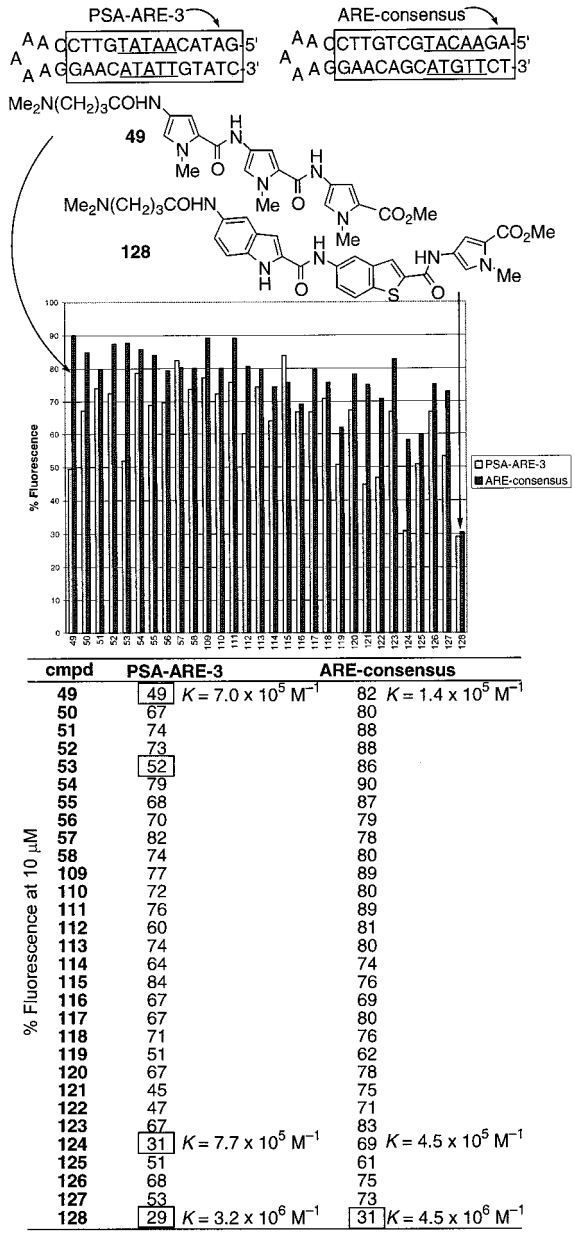


Figure 8. Ethidium bromide displacement assay for selected DMABA-trimers (100 μM) with hairpin oligonucleotides: PSA-ARE-3 oligomer and ARE-consensus (0.88×10^{-5} M DNA, 0.44×10^{-5} M ethidium bromide).

Screening the entire DMABA-trimer library for binding to the two hairpin oligonucleotides,[35] enlisting the ethidium bromide displacement assay, revealed that the mixture containing the pyrrole subunit (**1**) at both the second (B) and third (C) position gave the largest decrease in fluorescence with the PSA-ARE-3 hairpin oligonucleotide, which contains a five-base-pair AT-rich site. Substitution of a single AT base pair with a GC base pair (ARE-consensus) at the center of this sequence results in loss of affinity for this mixture. Screening the individual components of this mixture afforded the direct distamycin A

(32) Cato, A. C. B.; Hernerson, D.; Ponta, H. *EMBO J.* **1987**, *33*, 545.
(33) Cleutjens, K. B. J. M.; van der Dorput, H. A. G. M.; van Eekelen, C. C. E. M.; van Rooij, H. C. J.; Faber, P. W.; Trapman, J. *Mol. Endocrinol.* **1993**, *7*, 23.
(34) Chang, C. C.; Satzman, A.; Yeh, S.; Young, W.; Keller, E.; Lee, H.-J.; Wang, C.; Mizokama, A. *Crit. Rev. Eucaryotic Gene Expression* **1995**, *5*, 97. Bentel, J. M.; Tilley, W. D. *J. Endocrinol.* **1996**, *151*, 1. Veldscholte, J.; Ris-Stalper, C.; Kuiper, G. G. J. M.; Jenster, G.; Berrovoets, C.; Claasen, E.; van Rooij, H. C. J.; Trapman, J.; Brinkman, A. O.; Mulder, E. *Biochem. Biophys. Res. Commun.* **1990**, *173*, 534. Galbraigth, S. M.; Duchesne, G. M. *Eur. J. Cancer* **1997**, *33*, 545.
(35) Complete experimental details and results are provided in Supporting Information.

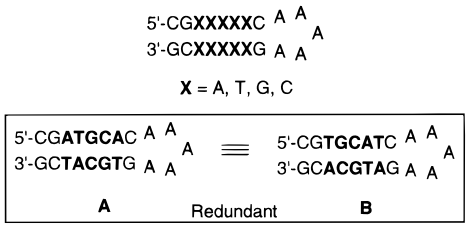


Figure 9. Structure of hairpin oligonucleotides representing all possible combinations of five base pairs. The position within the hairpin was not considered, making **A** and **B** redundant.

analogue **49** as having the highest affinity, followed closely by **53** containing the thiophene subunit at the first position (A). The same overall pattern was observed with poly[dA]−poly-[dT] where **49** and **53** also showed the highest affinity (Table 3).

Both of these agents exhibited diminished affinity for the ARE-consensus sequence presumably resulting from the intervening GC base pair. Two additional mixtures, **109**−**118** and **119**−**128**, also bound the PSA-ARE-3 sequence effectively with the general trend **119**−**128** > **49**-**58** > **109**−**118**, the same general trend seen with poly[dA]−poly[dT] (Table 3). The individual trimers **124** and **128** displayed tight binding to the PSA-ARE-3 sequence analogous to **49** and **53**. Importantly, **124** showed a loss of affinity to the ARE-consensus analogous to **49** and **53**, but **128** retained equal affinity making this agent ideal in maintaining high affinity for both the PSA-ARE-3 and ARE-consensus sequences.

This raised the issue of the DNA binding selectivity of **128** and its distinctions from **49** or distamycin A that are responsible for the PSA-ARE-3 and ARE-consensus binding. Consequently, the ethidium bromide displacement assay was enlisted to define the complete sequence selectivity of both distamycin A and **128**.

Complete Sequence Selectivity of a Prototypical DNA Binding Agent Distamycin A: Rank Order Binding to a Library of Hairpin Oligonucleotides. Distamycin A is among the best characterized DNA binding compounds. Its DNA binding properties have been studied in depth through footprinting,[8,36] calorimetry,[37] NMR,[38] and X-ray crystallography.[39] However, even for distamycin A, a detailed study of its rank order binding to all possible DNA sequences has not been described, and its affinity for nonoptimal binding sites is not easily assessed using common techniques. Consequently, it represents an ideal example with which the ethidium bromide technique could be examined in efforts to assess its use for establishing DNA binding selectivity. Thus, a survey of distamycin A binding to all possible five-base-pair DNA sequences was conducted using a library of 512 hairpin DNA oligonucleotides containing all possible five-base-pair sequences of the general format 5′-GCXXXXXC-3′ with a 5-A loop (Figure 9). Although there are 1024 possible sequences containing five base pairs, two complementary sequences are contained in each hairpin differing only in their location relative to the position of the adenine loop making, for example, the sequence 5′-ATGCA equivalent to the sequence 5′-TGCAT (Figure 9).

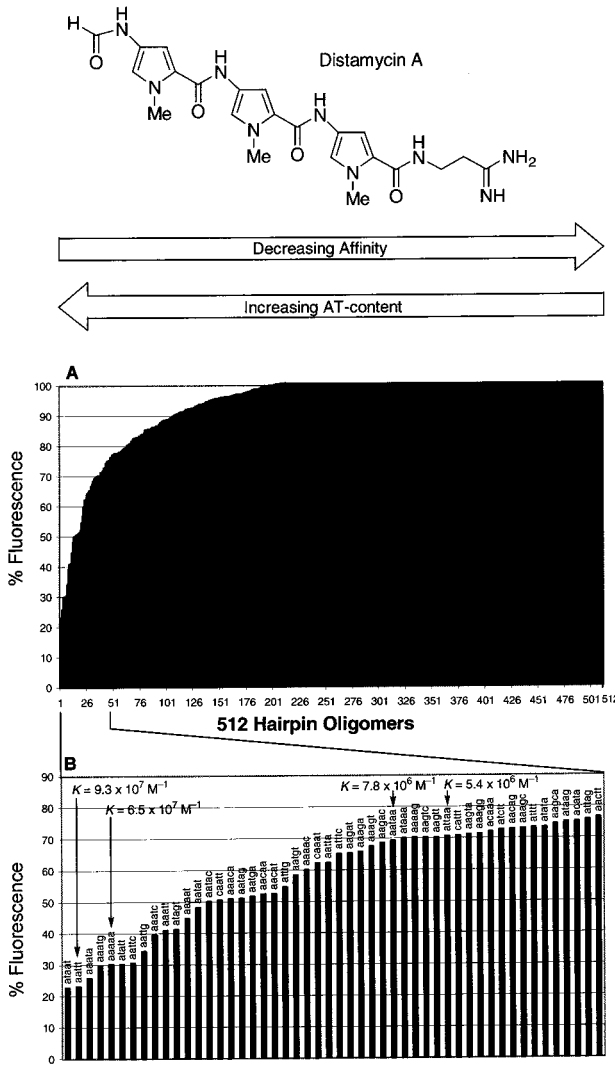


Figure 10. Screen of distamycin A against a library of DNA hairpin oligonucleotides, (**A**) All 512 sequences, (**B**) 50 sequences showing highest affinity.

Table 4

DNA sequence	K (M^{-1})	K (M^{-1})$^{lit.}$
5′-AATTT-3′	9.4×10^7	[a] 3.1×10^7
5′-AAAAA-3′	6.5×10^7	[b] 2.6×10^7
5′-AATAA-3′	7.8×10^6	[b] 1.4×10^7
5′-ATTAA-3′	5.4×10^6	[b] 1.9×10^6

[a] Calorimetry, ref 37. [b] Footprinting on a close analogue of distamycin A, ref 40.

The results of screening the 512-membered library of hairpin oligonucleotides using distamycin A is given in Figure 10. As expected, affinity increases with increasing AT content. The top sequences include the sites 5′-ATAA, 5′-AATT, 5′-AAAT, and 5′-AAAA and among the 20 hairpins showing the greatest decrease in % fluorescence, three four-base-pair sequences occur most often: 5′-AATT, 5′-AAAT, and 5′-AATA.

Although distamycin A has been studied in extensive detail, suprisingly few absolute binding constants for short AT-rich sequences have been published. The comparison of all those disclosed show the relative trend 5′-AATTT > AAAAA > AATAA > ATTAA (Table 4).[37,40] The ethidium bromide displacement assay revealed the same general trend and a

(36) Portugal, J.; Waring, M. J. *Eur. J. Biochem.* **1987**, *167*, 281. Portugal, J.; Waring, M. J. *FEBS Lett.* **1987**, *225*, 195. Abu-Daya, A.; Brown, P. M.; Fox, K. R. *Nucleic Acids Res.* **1995**, *23*, 3385. Abu-Daya, A.; Fox, K. R. *Nucleic Acids Res.* **1997**, *25*, 4962.

(37) Rentzeperis, D.; Marky, L. A.; Dwyer, T. J.; Geierstanger, B. H.; Pelton, J. G.; Wemmer, D. E. *Biochemistry* **1995**, *34*, 2937.

(38) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723. Klevit, R. E.; Wemmer, D. E.; Reid, B. R. *Biochemistry* **1986**, *25*, 3296. Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393.

(39) Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8385.

(40) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11385.

Table 5. Ethidium Bromide Binding Constants[a]

polynucleotide	K_{EB} ($\times 10^6$ m^{-1})
poly[dT]−poly[dAT]	9.5
poly[dA]−poly[dT]	0.65
poly[dGC]−poly[dGC]	9.9
poly[dG]−poly[dC]	4.5
poly[dAC]−poly[dGT]	9.8
poly[dAG]−poly[dCT]	1.3
calf thymus	10

[a] Reference 11.

quantitative titration measurement of binding constants with the hairpin oligonucleotides containing these sequences afforded binding constants that are not only consistent with the relative trend (Figure 10) but also within a factor of 2−3 of *all* of the absolute binding constants previously determined through calorimetry and footprinting (Table 4). Given that the DNA upon which the measurements were made is different, that the buffer conditions are not identical, and that entries 2−4 in Table 4 were derived from a close analogue of distamycin A, all of which may contribute to small discrepancies in the absolute binding constants, the ethidium bromide displacement titration assay appears to be remarkably accurate at reproducing absolute binding constants.

Because the fluorescence derived from ethidium bromide binding varies from sequence to sequence, it is the % fluorescence decrease and not the final absolute fluorescence that is proportional to the extent of DNA binding. For tight binding agents such as distamycin A which display a well-defined linear loss of fluorescence, the absolute binding constants can be established by quantitative titration independent of a consideration of the ethidium bromide binding constants using a noncompetitive model of $K = 1/[\text{agent}] - 0.5[\text{DNA}]r$ where K = binding constant, [agent] = concentration at 50% reduction in fluorescence, [DNA] = DNA concentration, and r = 1/binding site size.[12] The latter can be determined experimentally from the extrapolated *x*-intercept (0% fluorescence) of a % fluorescence vs [agent]/[base pair] plot. In the present study, this was established to be 0.125 which corresponds to 1 bound agent/8 base pairs or 1 agent bound per hairpin oligonucleotide. The alternative use of a competitive binding model to calculate absolute binding constants requires a knowledge of the ethidium bromide binding constants for each sequence and follows from $K = K_{EB}[EB]/[\text{agent}]$ where K = binding constant, K_{EB} = binding constant for ethidium bromide, [EB] = ethidium bromide concentration, [agent] = agent concentration at 50% fluorescence.[6,10] The ethidium bromide binding constant varies considerably (Table 5)[11] and the displacement does not follow a 1:1 stoichiometry, both of which complicate the use of a competitive binding model for establishing binding constants. The most accurate usage likely would employ the calf thymus K_{app} (10×10^6 M^{-1}) which represents an average binding constant. Neglecting the stoichiometry of the displacement as recommended,[6] this provides a $K = 7.9 \times 10^7$ M^{-1} for distamycin A binding to AATTT, essentially indistinguishable from the value established using the noncompetitive model (9.4×10^7 M^{-1}). The advantage of this method is that it does not require establishment of binding site size for the compound, but this necessarily introduces inaccuracies due to this assumption. Thus, while the competitive model provides reasonable constants for high affinity sequences, it overestimates the binding constants for the weaker sequences.[6] Therefore, we prefer and recommend the use of the noncompetitive binding model for establishing absolute binding constants with hairpin oligonucleotides and caution that they should be further validated by other techniques.

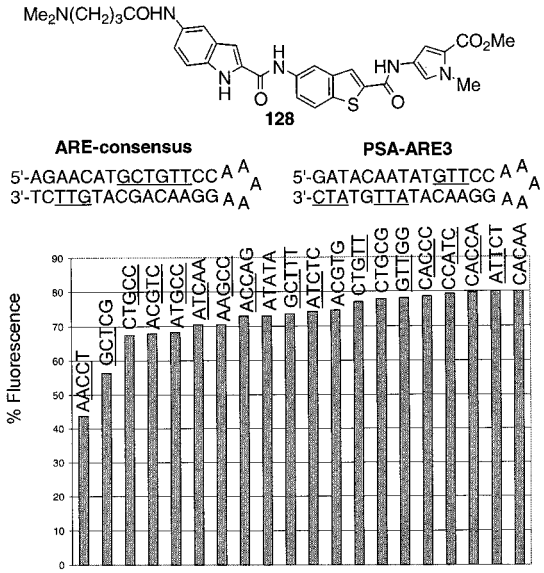


Figure 11. Screen of compared **128** with a library of 512 DNA hairpin oligonucleotides (top 20 sequences shown). Underlined sequences represent purine/pyrimidine/pyrimidine segments.

Notably, there is more information, or rather a higher resolution of information, regarding the selectivity of distamycins binding to DNA in this single experiment than may be found in all past work combined, which typically is limited to identification of the highest affinity sites. The establishment of rank order affinity for all possible binding sites including the ability to determine relative *or* absolute binding constants for even modest or low affinity sites provide a new opportunity to qualitatively or quantitatively compare the DNA binding selectivity of various compounds. For example, simply the shape of the merged bar graph shown in Figure 10A, the curve and slope of the resulting graph, and the area under the curve provide means by which to qualitatively and quantitatively compare selectivities of DNA binding. These are presently under investigation and our assessments of their value will be disclosed in due course.

Sequence Selectivity Determination for 128: A Novel DNA Binding Agent. As detailed in an earlier section, **128** bound poly[dA]−poly[dT] with an affinity equal to the distamycin analogue **49**. However, it also bound poly[dG]−poly[dC] with only a slightly reduced affinity being 25−30 times more effective than **49** and, unlike **49**, it bound tightly to both the PSA-ARE-3 and ARE consensus sequences. The sequence selectivity of **128** was established by screening it against the library of 512 hairpin oligonucleotides (Figure 11). Compound **128** was found to clearly bind with a selectivity distinct from that of distamycin A and it appears to exhibit a significant preference for PuPyPy sequences. Of the 20 highest affinity sequences, 16 contain the PuPyPy motif (80%), where statistically 37.5% of the sequences would be expected to contain this motif in a random sample. One of the four exceptions contained a five-base-pair AT-rich site. Within both of the androgen response elements used to identify **128**, the PuPyPy motif is repeated three times. It appears this may be the reason for the equally high affinity binding of **128** with both sequences. Further studies on **128** are required to establish the structural basis for its DNA binding selectivity and these will be disclosed in due course.

Table 6

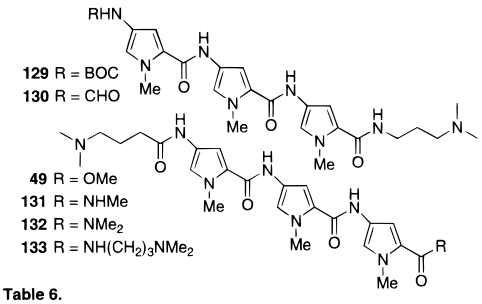
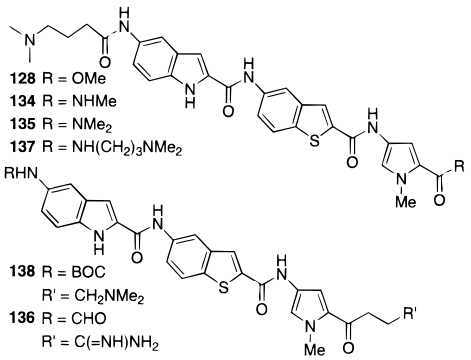


Table 6.

	K (x 10^6 M^{-1})[a] [dA]–[dT]	[dG]–[dC]	ΔG^b [dA]–[dT]	IC_{50} μM (L1210)
Dist	15.0	0.071	–9.78	42
49	5.9	0.073	–9.22	44
129	2.1	0.083	–8.74	32
130	15.9	0.089	–9.80	100
131	11.8	0.110	–9.64	100
132	2.5	0.076	–8.72	100
133	12.7	0.150	–9.69	>10

[a] K calculated from $K = K_e$[EtBr]/[Agent], see ref 6.
[b] $\Delta G = -RT\ln K$ (298 K)



	K (x 10^6 M^{-1})[a] [dA]–[dT]	[dG]–[dC]	ΔG^b [dA]–[dT]	IC_{50} μM (L1210)
128	5.6	1.9	–9.20	17
138	2.5	1.3	–8.72	0.43
134	8.6	1.9	–9.45	19
135	2.5	1.2	–8.72	23
136	7.3	2.3	–9.35	0.72
137	9.5	1.5	–9.51	33

Analysis of Side Chain Position and Nature of the Basic Functionality. Finally, to evaluate the effects of the structural changes that were made relative to distamycin A in the nature and position of the charged functionality, a number of derivatives of the high affinity DNA binding agents including the tripyrrole core of distamycin were prepared (Table 6).

Analysis of these derivatives using the quantitative titration with displacement of prebound ethidium bromide showed that there is very little difference between an amidine as the basic side chain and the dimethylamino group (distamycin vs **130**) or between placing the basic side chain at the *C*- or *N*-terminal end of the trimer (**49** vs **129**), Table 6. Amine **129** shows slightly lower binding affinity to poly[dA]−poly[dT] than **49** which may arise from the incorporation of a bulky *tert*-butyl group present at the *N*-terminus. Interestingly, there is approximately a 3-fold difference in binding between distamycin A or **130** and the tripyrrole **49**. The primary difference between the two molecules is the presence of an additional potential hydrogen bond donor in distamycin and **130** (*N*-terminal formamide) that is not present in either **49** (*C*-terminal ester), **129** (*N*-terminal BOC-group), or **132** (C-terminal dimethylamide). When a potential hydrogen bond donor group is included at the *C*-terminus (**131**), the binding affinity does approximate that of distamycin A. The difference in free energy of binding between those molecules containing an additional donor hydrogen bonding group (distamycin A and **130**−**131**) and those which do not (**49**, **129**, **132**) is approximately 1 kcal/mol, the value of a single hydrogen bond. Interestingly, adding a second substituent containing an additional basic, protonated amine (**133** vs **131**) does not further increase the DNA binding affinity.

Analogous observations were made with derivatives of **128** which bound to poly[dA]−poly[dT] within a factor of 2 of the corresponding distamycin A derivatives. The distinctions between **128** and the distamycin derivatives were that the former bound poly[dG]−poly[dC] 15−30 times more effectively.

Conclusions

An approach to the rapid, parallel solution-phase synthesis of distamycin A analogues was developed enlisting a simple, acid/base liquid−liquid extraction for purification and isolation of each intermediate and final product (≥95% pure). Its utility was demonstrated with the preparation of distamycin A and a prototypical library of 2640 analogues assembled in a small mixture format of two libraries of 132 mixtures of 10 compounds providing each in multimilligram quantities sufficient for screening in multiple assays. Screening of the library in a functional assay for cytotoxic activity (L1210) revealed two uniquely active compounds, **66** and **67**, which were 1000 times more potent than distamycin A, and **86**, which was 100 times more potent than distamycin A. More fundamental, a complementary rapid, high-throughput screen for DNA binding affinity was developed based on the loss of fluorescence derived from displacement of prebound ethidium bromide which is applicable for assessing binding to DNA homopolymers or specific sequences (hairpin oligonucleotides). Using this technique, the distamycin A tripyrrole analogue **49** as well as alternative AT-rich binding agents were identified (**112** and **128**) establishing the validity of the technique and providing two new and effective DNA binding agents. In addition, a comparison of several distamycin analogues established substituent contributions to AT-rich binding that may be safely implemented in future libraries. Extension of these studies to identify effective binders to predefined sequences was conducted in the context of the androgen response elements PSA-ARE-3 and the ARE-consensus sequences, the latter of which contains a GC base pair interrupted AT-rich sequence, and for which effective binders might prove useful in the treatment of hormone antagonist resistant prostate cancer. Three agents, **124**, **128**, and the distamycin analogue **49** were identified as high affinity binders for the PSA-ARE-3 sequence incorporating a five-base-pair AT-rich sequence. Fundamentally more important and unlike the distamycin analogue **49**, **128** retained this high affinity binding to the ARE-consensus sequence, which contains a GC base pair interrupted AT-rich site, suggesting it may serve as an effective inhibitor of androgen receptor DNA binding and its initiated gene transcription. Extension of the DNA binding assay to a powerful technique for establishing the DNA binding selectivity of an individual compound was developed enlisting distamycin A and its comparison with **128** requiring the assay of each compound against a library of 512 hairpin oligonucleotides. This provided their rank of order binding to all possible five-base-pair sequences and completely defined their DNA binding sequence selectivity. The technique, which conservatively requires manual measurement times of 1−2 min/plate (15 min/compound/512 hairpins) on a fluorescent plate reader reproduced the known properties of distamycin A and revealed distinctions with **128** responsible for the differences in binding the two 14-base-pair androgen response elements. Combining

the assay of a library of compounds against a library of DNA hairpin oligonucleotides, with automation of the assay, provides qualitative and/or quantitative information on the binding of all library members against a library of available sequences. Studies on extensions of this work are in progress and will be disclosed in due time.

Experimental Section

Methyl 4-[[(4-*tert*-Butyloxycarbonyl)amino-1-methylpyrrol-2-yl]-carbonyl]amino-1-methylpyrrole-2-carboxylate (2). Initial conditions: a solution of **1a** (250 mg, 1.05 mmol, 1 equiv) and **1b** (200 mg, 1.05 mmol, 1 equiv) in DMF (5 mL) was treated with EDCI (403 mg, 2.1 mmol, 2 equiv) and DMAP (320 mg, 2.6 mmol, 2.5 equiv) and the resulting solution was stirred for 14 h at 25 °C. The reaction mixture was poured into EtOAc (50 mL) and washed with 10% aqueous HCl (3 × 50 mL) and saturated aqueous $NaHCO_3$ (3 × 50 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated to provide **2** (350 mg, 89%) as a tan foam. For optimized large scale:[41] A solution of **1a** (3.8 g, 15.8 mmol, 1 equiv) and **1b** (3.0 g, 15.8 mmol, 1 equiv) in DMF (40 mL) and CH_2Cl_2 (10 mL) was treated with EDCI (4.5 g, 23.5 mmol, 1.5 equiv) and DMAP (2.3 g, 18.9 mmol, 1.2 equiv), and the resulting solution was stirred for 18 h at 25 °C. The reaction mixture was poured into EtOAc (60 mL) and washed with 10% aqueous HCl (3 × 50 mL) and saturated aqueous $NaHCO_3$ (3 × 50 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated to provide **2** (5.8 g, 97%): mp 78−79 °C; ^{1}H NMR (DMSO-d_6, 500 MHz) δ 10.07 (s, 1H), 9.87 (s, 1H), 7.46 (s, 1H), 7.11 (s, 1H), 6.96 (s, 1H), 6.90 (s, 1H), 3.88 (s, 3H), 3.84 (s, 3H), 3.73 (s, 3H); ^{13}C NMR (acetone-d_6, 125 MHz) 161.9, 159.5, 153.8, 124.1, 123.9, 123.8, 121.3, 120.0, 117.8, 108.9 (2C), 103.9, 79.2, 51.1, 36.7, 36.5, 28.5; IR (film) ν_{max} 3328, 2954, 1694, 1644, 1586 cm^{-1}; MALDIHRMS (DHB) *m*/*z* 399.165 (M + Na^+, $C_{18}H_{24}N_4O_5$ requires 399.1644).

Methyl 4-[[[4-[[(4-*tert*-Butyloxycarbonyl)amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrole-2-carboxylate (3). Initial conditions: a sample of **2** (50 mg, 0.13 mmol, 1 equiv) was treated with 4.0 N HCl/EtOAc (1 mL). The reaction mixture was stirred at 25 °C for 30 min, then concentrated and dried under reduced pressure for 1 h. EDCI (50 mg, 0.27 mmol, 2 equiv), DMAP (33 mg, 0.27 mmol, 2 equiv), and **1a** (63 mg, 0.27 mmol, 2 equiv) were added to a solution of the crude amine in DMF (1 mL). The reaction mixture was stirred for 3 h at 25 °C, diluted with EtOAc (10 mL) and washed with 10% aqueous HCl (3 × 10 mL) and saturated aqueous $NaHCO_3$ (3 × 10 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated to a yellow solid. The solid material was suspended in 1:1 MeOH/10% NaOH (20 mL) and stirred for 30 min at 25 °C to decompose small amounts of contaminate symmetrical anhydride. The solution was then poured into EtOAc (20 mL) and washed with $NaHCO_3$ (3 × 20 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated to provide **3** (47 mg, 73%) as a yellow foam. For optimized large scale:[41] dipyrrole **2** (2.8 g, 7.43 mmol, 1 equiv) was treated with 4.0 N HCl/EtOAc (20 mL). The reaction mixture was stirred at 25 °C for 30 min and then concentrated and dried under reduced pressure. EDCI (2.1 g, 11.1 mmol, 1.5 equiv), DMAP (1.1 g, 9.0 mmol, 1.2 equiv), and **1a** (2.0 g, 8.2 mmol, 1.1 equiv) were added to a solution of the crude amine in DMF (100 mL). The reaction mixture was stirred for 3 h at 25 °C, diluted with EtOAc (100 mL), and washed with 10% aqueous HCl (3 × 100 mL) and saturated aqueous $NaHCO_3$ (3 × 100 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated to provide **3** (3.6 g, 96%) as a yellow foam: mp 131−133 °C; ^{1}H NMR (DMSO-d_6, 500 MHz) δ 9.26 (s, 1H), 9.23 (s, 1H), 8.13 (s, 1H), 7.48 (d, 1H, J = 1.9 Hz), 7.21 (d, 1H, J = 1.9 Hz), 6.97 (d, 1H, J = 1.5 Hz), 6.92 (s, 1H), 6.90 (d, 1H, J = 2.2 Hz), 6.77 (s, 1H), 3.92 (s, 3H), 3.91 (s, 3H), 3.89 (s, 3H), 3.74 (s, 3H), 1.44 (s, 9H); ^{13}C NMR (acetone-d_6, 125 MHz) 161.9, 159.6, 159.5, 153.8, 124.1, 123.8 (2C), 123.7, 121.3, 120.0, 119.1, 117.7, 109.9 (2C), 104.7, 103.8, 79.2, 51.1, 36.6 (2C), 36.5, 28.5; IR (film) ν_{max} 3314, 2954, 1694, 1644, 1584, 1552 cm^{-1}; MALDIHRMS (DHB) *m*/*z* 521.2123 (M + Na^+, $C_{24}H_{30}N_6O_6$ requires 521.2124).

(41) The large-scale synthesis of **2** and **3** was optimized by Takahiro Ishii.

4-[[[4-[[(4-*tert*-Butyloxycarbonyl)amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methyl-2-(((carbonyl)amino)propio-3-nitrile)pyrrole (4). Tripyrrole **3** (70 mg, 0.14 mmol, 1 equiv) in THF/MeOH (3:1, 2 mL) was treated with a solution of LiOH (24 mg, 0.56 mmol, 4 equiv) in H_2O (0.5 mL). The solution was warmed at 60 °C for 5 h and then diluted with EtOAc (20 mL) and H_2O (20 mL). The layers were separated, and the aqueous layer was brought to pH 3 with 10% aqueous HCl. The resulting slurry was extracted with EtOAc (4 × 20 mL), and the combined organic extracts were dried (Na_2SO_4), filtered, and concentrated. The crude acid (30 mg, 0.062 mmol) in DMF (1 mL) was treated with EDCI (23 mg, 0.12 mmol, 2 equiv) and DMAP (19 mg, 0.16 mmol, 2.5 equiv), followed by 3-aminopropionitrile (13 mg, 0.12 mmol, 2 equiv). The reaction mixture was stirred for 14 h at 25 °C and then diluted with EtOAc (20 mL) and washed with 10% aqueous HCl (3 × 20 mL) and saturated aqueous $NaHCO_3$ (3 × 20 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated to afford **4** (31 mg, 95%) as a yellow solid: mp 170−172 °C; ^{1}H NMR (DMSO-d_6, 500 MHz) δ 9.93 (s, 1H), 9.87 (s, 1H), 9.10 (s, 1H), 8.35 (t, 1H, J = 5.6 Hz), 7.23 (d, 1H, J = 1.5 Hz), 7.22 (d, 1H, J = 1.5 Hz), 7.00 (d, 1H, J = 1.5 Hz), 6.93 (d, 1H, J = 1.9 Hz), 6.89 (s, 1H), 6.84 (s, 1H), 3.84 (s, 3H), 3.81 (s, 6H), 3.40 (q, 2H, J = 6.2 Hz), 2.73 (t, 2H, J = 6.4 Hz), 1.46 (s, 9H);^{13}C NMR (acetone-d_6, 125 MHz) 162.6, 160.0, 159.6, 153.8, 124.5, 124.0, 123.8, 123.6 (2C), 123.5, 119.3, 119.2, 119.1, 117.8, 104.8, 104.7, 103.9, 79.3, 36.6 (2C), 36.5, 36.1, 28.5, 18.6; IR (film) ν_{max} 3318, 2978, 1698, 1644, 1586, 1537 cm^{-1}; MALDIHRMS (DHB) *m*/*z* 559.2387 (M + Na^+, $C_{26}H_{32}N_8O_5$ requires 559.2393).

Distamycin A Hydrochloride. A solution of nitrile **4** (12 mg, 0.022 mmol) in anhydrous EtOH (0.3 mL) was treated with 8.0 N HCl/EtOH (1 mL) at 0 °C for 30 min and then slowly warmed to 25 °C and stirred for 2 h. The solvent was removed under a stream of N_2, and the residue was washed with Et_2O (3 mL) and dried in vacuo for 30 min. The resulting solid was taken up in EtOH (0.3 mL) and treated with 7% NH_3/EtOH (1 mL) at 25 °C. After 1 h the reaction was concentrated to a tan solid and dried under reduced pressure for 1 h. The crude amidine was dissolved in MeOH (0.2 mL) and cooled to −40 °C. The solution was treated with a solution containing *N*-formylimidazole, prepared by treating carbonyldiimidazole (18 mg, 0.11 mmol) in THF (0.4 mL) with a solution of formic acid (4.3 mL, 0.11 mmol) in THF (0.4 mL) at 25 °C for 15 min. The reaction mixture was stirred at −40 °C for 1 h and then concentrated to a volume of 0.2 mL. The product was precipitated with EtOAc (1 mL) and collected by filtration. The crude product was dissolved in cold *i*-PrOH (2 mL) containing decolorizing carbon (100 mg), stirred at 0 °C for 30 min, filtered, and concentrated to a light yellow solid. The solid material was taken up EtOAc/acetone/MeOH/0.01 N HCl (5:3:1:1, 2 mL) and stirred with SiO_2 for 30 min and then filtered through Celite to remove traces of NH_4Cl and afford pure distamycin A (4.9 mg, 45%) identical in all respects with authentic material: mp 186−188 °C; ^{1}H NMR (DMSO-d_6, 500 MHz) δ 10.28 (s, 1H), 9.99 (s, 1H), 9.96 (s, 1H), 9.14 (br s, 2H), 8.93 (br s, 2H), 8.29 (t, 1H, J = 5.6 Hz), 8.11 (s, 1H), 7.25 (s, 1H), 7.21 (d, 1H, J = 1.5 Hz), 7.20 (d, 1H, J = 1.5 Hz), 7.06 (s, 1H), 6.95 (d, 1H, J = 1.5 Hz), 6.94 (d, 1H, J = 1.5 Hz), 3.83 (s, 6H), 3.80 (s, 3H), 3.51 (q, 2H, J = 6.0 Hz), 2.66 (t, 2H, J = 6.4 Hz); ^{13}C NMR (CD_3OD, 125 MHz) 171.0, 164.4, 161.4, 161.2, 160.4, 127.7, 124.5, 123.9, 123.4, 123.3, 122.0, 120.9, 120.8, 119.9, 106.7, 106.5, 105.8, 37.5, 36.9, 36.8, 36.7, 34.4; IR (film) ν_{max} 3299, 2975, 1634 cm^{-1}; MALDIHRMS (DHB) *m*/*z* 504.2078 (M +Na^+, $C_{22}H_{27}N_9O_4$ requires 504.2084).

Ethidium Bromide Assay. DNA hairpin oligonucleotides[42] were purchased from Genbase Inc. (San Diego) as 880 μM (base pairs) solutions in water and stored as stock solutions at −80 °C. Prior to use, each oligonucleotide was diluted to 88 μM in water and stored at 0 °C for no longer than 2 days. Each well of a Costar black 96-well plate was loaded with Tris buffer containing ethidium bromide (0.1 M Tris, 0.1 M NaCl, pH 8, 0.44 × 10^{-5} M ethidium bromide final concentration, 88 μL). To each well was added one hairpin oligonucleotide (10 μL, 0.88 × 10^{-5} M in DNA base pairs final concentration). To each well was added distamycin A (2 μL of a 0.1 mM solution in

(42) The full set of 512 hairpin oligonucleotides may be purchased from GenBase, 6450 Lusk Blvd, Suite E. 107, San Diego, CA 92121, telephone: 858-453-8879, e-mail: genbase@aol.com.

water, 2.0×10^{-6} M final concentration) or **128** (6 μL of a 0.1 mM solution in water, 6.0×10^{-6} M final concentration). After incubation at 25 °C for 30 min, each well was read (average of 30 readings) on a Molecular Devices Spectra Max Gemini fluorescent plate reader (ex. 545 nm, em. 595 nm, cutoff filter at 590 nm) in duplicate experiments with two control wells (no distamycin = 100% fluorescence, no DNA = 0% fluorescence). Fluorescence readings are reported as % fluorescence relative to the controls. In our experience, fluorescence plate readers show a variability of ±10%, but surface effects (i.e., bubbles, dust) may contribute to larger variations requiring a second set of measurements.

Determination of Distamycin A Binding Constants with Hairpin Oligonucleotides. A 3 mL quartz cuvette was loaded with Tris buffer (0.1 M Tris, 0.1 M NaCl, pH 8) and ethidium bromide (0.44×10^{-5} M final concentration). The fluorescence was measured (ex. 545 nm, em. 595 nm) and normalized to 0% relative fluorescence (free ethidium bromide is only weakly fluorescent). The DNA hairpin oligonucleotide of interest was added (0.88×10^{-5} M in DNA base pairs final concentration), the fluorescence was measured again and normalized to 100% relative fluorescence. A solution of distamycin A (1 μL, 0.1 mM in DMSO) was added, and the fluorescence was measured following 30 min of incubation at 23 °C (each measurement was conducted four times and averaged). The addition of 1 μL aliquots was continued until the relative fluorescence had decreased to ≤50%. Binding constants (K) were calculated from $K = 1/[\text{agent}] - 0.5[\text{DNA}]/r$ where K = binding constant, [agent] = concentration at 50% reduction in fluorescence, [DNA] = DNA concentration, and $r = 1$/binding site size, using $r = 1/8$ or 0.125.

Acknowledgment. We gratefully acknowledge the financial support of the National Institutes of Health (CA78045 and CA41986), Novartis, and the award of a NIH postdoctoral fellowship (B.E.F., F32CA76718, 1998−2000). Takahiro Ishii is gratefully acknowledged for the preparation of large amounts of tripyrrole **3** and the preparation of **133** and Michael A. Dechansreiter for continued refinements of the results presented in Figure 10.

Supporting Information Available: ^{1}H NMR of **2**−**4** and distamycin A (Figure S1), experimental details for the preparation and characterization of **7b**, **8a**, **8b**, **11a**, **11b**, **12a**, **12b**, **14a**, **14b**, **15a**, **15b**, **17**, **19**, **20**, **22**, **23**, **24**, **25**, **26**, **28**, **29**, **31**, and **32**, general procedures for the preparation of individual and mixture BOC-trimers, ^{1}H NMR characterization of **34**−**48**, **59**−**78**, **99**−**108**, matrix MS characterization of the BOC-trimer libraries, general procedures for the preparation of individual and mixture DMABA-trimers, ^{1}H NMR characterization of **49**−**58**, **79**−**98**, **109**−**128**, MS characterization of the DMABA-trimer libraries, experimental details and characterization of **129**−**137**, a figure summary of the testing data from a L1210 assay of the 132 BOC-dimers (Figure S2) and experimental details for the ethidium bromide displacement assay including results of screening DMABA-trimer libraries against two hairpin oligonucleotides, and full sequence selectivity data on distamycin A against the 512 hairpin oligonucleotides (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA994192D