The hybridization-stabilization assay: a solution-based isothermal method for rapid screening and determination of sequence preference of ligands that bind to duplexed nucleic acids

Carolyn Gonzalez, Megan Moore, Sofia Ribeiro, Uli Schmitz*, Gary P. Schroth, Lisa Turin and Thomas W. Bruice

Genelabs Technologies Inc., 505 Penobscot Drive, Redwood City, CA 94063, USA

Received May 10, 2001; Revised and Accepted July 1, 2001

ABSTRACT

The gene-to-drug quest will be most directly served by the discovery and development of small molecules that bind to nucleic acids and modulate gene expression at the level of transcription and/or inhibit replication of infectious agents. Full realization of this potential will require implementation of a complete suite of modern drug discovery technologies. Towards this end, here we describe our initial results with a new assay for identification and characterization of novel nucleic acid binding ligands. It is based on the well recognized property of stabilization of hybridization of complementary oligonucleotides by groove and/or intercalation binding ligands. Unlike traditional thermal melt methodologies, this assay is isothermal and, unlike gel-based footprinting techniques, the assay also is performed in solution and detection can be by any number of highly sensitive, non-radioisotopic modalities, such as fluorescence resonance energy transfer, described herein. Thus, the assay is simple to perform, versatile in design and amenable to miniaturization and high throughput automation. Assay validation was performed using various permutations of direct and competitive binding formats and previously well studied ligands, including pyrrole polyamide and intercalator natural products, designed hairpin pyrrole–imidazole polyamides and furan-based non-polyamide dications. DNA specific ligands were identified and their DNA binding site size and sequence preference profiles were determined. A systematic approach to studying the relationship of binding sequence specificity with variation in ligand structure was demonstrated, and preferred binding sites in longer DNA sequences were found by pseudo-footprinting, with results that are in accord with established findings. This assay methodology should promote a more rapid discovery of novel nucleic acid ligands and potential drug candidates.

INTRODUCTION

Specific molecular recognition through information encoded in the sequence of nucleic acids is a requirement of fundamental processes in molecular biology including replication, transcription and translation. Enabling the regulation of gene expression at the level of transcription and the inhibition of growth and proliferation of infectious agents by small molecules that bind nucleic acids represents an important approach in the gene-to-drug discovery paradigm (1,2). However, understanding and control of nucleic acid recognition by small ligands is less developed than that of Watson–Crick base pairing of sequence complementary strands. Thus, full realization of a shift from predominately protein targeted to nucleic acid targeted small molecule drug discovery will require implementation of a comprehensive and integrated suite of new technologies.

Ligand binding to DNA (or RNA) has been studied extensively by numerous biochemical and biophysical methods (3–5). Among these are chemical and enzymatic footprinting, UV and CD spectroscopy, calorimetry, NMR spectroscopy and X-ray crystallography. These methods may require high concentrations and/or amounts of ligand and nucleic acids and/or may be tedious to perform. A commonly employed solution-based method is to measure ligand binding induced changes in the hybridization stability of nucleic acids as probed by thermal denaturation–renaturation detected by UV spectrophotometry (3–6). Numerous types of ligands, which may groove bind or intercalate, cause an increase in the T_m, the temperature corresponding to the peak of the first derivative of the UV absorbance curve for thermal denaturation–renaturation of

*To whom correspondence should be addressed. Tel: +1 650 562 1411; Fax: +1 560 368 0709; Email: ulis@genelabs.com
Present addresses:
Carolyn Gonzalez and Gary P. Schroth, PE Applied BioSystems, 850 Lincoln Centre Drive, Foster City, CA 94404, USA
Sofia Ribeiro, Berlex Biosciences, 15049 San Pablo Avenue, Richmond, CA 94804, USA
Megan Moore, Apple, 1 Infinite Loop, Cupertino, CA 95014, USA
Thomas W. Bruice, Biota Research Laboratories, Inc., 1945 Camino Vida Roble, Carlsbad, CA 92008, USA

complementary strands. These ligands include polyamines, small molecule drugs, peptides and proteins, and carbohydrates (6). This method typically requires stoichiometric binding of ligands and nucleic acids at high concentrations and commonly available instrumentation generally requires relatively large amounts of material. Most importantly, UV thermal melting experiments only provide information at the T_m of the system, which varies among systems [dependent on choice of oligonucleotides (ODNs) and ligands] and, thus, presents an impediment to automation, generally is not at a physiologically relevant temperature, and may not be a true indicator of thermodynamic binding affinities (7). We describe here a versatile, solution-based isothermal assay method that allows for the rapid study, in detail, of the molecular determinants of ligand binding to nucleic acids using relatively small amounts of material. This technology incorporates preferred attributes of established methods of nucleic acid analysis: (i) solution phase homogeneous detection of nucleic acid hybridization using covalently attached labels and (ii) the stabilizing effect of ligand binding on the duplex form of nucleic acids.

In the method described here, we use short, length matched, complementary 'indicator' ODN pairs, which are single-stranded under assay conditions, to probe for ligand binding. One strand is modified on the 5′- (or 3′-) terminus with a fluorescent dye (fluorescein), and the second strand is modified on the 3′- (or 5′-) terminus with a quenching group (dabcyl). This juxtaposition of opposing fluorophore/quencher groups that can be brought together in close proximity via hybridization has been used recently in the development of probes designed to measure target levels in PCR reactions as research tools and molecular diagnostics (e.g. TaqMan probes and molecular beacons; 8–13). Ligand binding can drive the formation of a stable duplex structure of ODN pairs. The hybridization of the complementary strands brings the two labels in close proximity, quenching the fluorescence via resonance energy transfer (FRET) only when the ligand stabilizes the duplex at the isothermal assay temperature. By using a limited variety of paired complementary ODN indicator sequences with similar overall duplex stabilities (direct hybridization-stabilization assay or dHSA), the sequence preference of ligand binding can readily be determined to be for G/C- or A/T-rich DNA. Alternatively, the preference for double-stranded DNA or RNA can be determined. Once a good 'indicator' ODN pair has been found for a particular ligand, it can be used additionally as a ligand 'depot' in competition studies (cHSA). For the latter, various fully double-stranded ODN sequences are used to compete with the indicator depot ODN pairs for ligand binding in order to show, in detail, the sequence (or local structure) preference determinants of the best ligand binding sites. Here we report the results of our initial development of the HSA in the FRET detection mode using known ligands, previously characterized by other methods. Other detection modalities have also been used successfully in the HSA, including fluorescence polarization, scintillation proximity and electrochemiluminescence and several of the assay design permutations reported herein have been automated (unpublished results).

MATERIALS AND METHODS

Materials

All commercially available ligands (netropsin, distamycin A, actinomycin D) were purchased from Sigma. Dicationic furamidines were a gift from the laboratories of David Boykin and W. David Wilson. Hairpin linked polyamides were synthesized on solid phase by modifications of published methods (14). Synthesis of ligand GL020924 will be described elsewhere. Ligands were resuspended in Tris–EDTA buffer (10 mM Tris–HCl, 0.1 mM EDTA, pH 7.2) and solution concentrations were determined spectrophotometrically. All fluorescent reporter and fluorescence quencher-labeled and unlabeled ODNs were purchased from Sigma-Genosys (The Woodlands, TX).

Hybridization-stabilization assay

Hybridization-stabilization assays were performed in 96-well plates on a CytoFluor® Series 4000, Fluorescence Multi-Well Plate Reader from PerSeptives Biosystems (MA). Emission was at 530 nm and excitation was at 485 nm and the gain was 70. Samples for dHSA consisted of 25 nM of 5′-fluorescein-labeled ODN mixed with 30 nM of complementary 3′-dabcyl quenching ODN (Sigma-Genosys) in 10 mM HEPES–HCl pH 7.2, 10 mM NaCl and 0.1 mM EDTA in a total volume typically of 200 µl. The appropriate sequences and lengths of these so-called 'indicator' ODN pairs was determined initially using T_m calculation methods (15,16), so as to give predominately single-stranded pairs at ambient temperature, and then were experimentally adjusted empirically, if necessary, for improved assay performance. Ligand was added at various concentrations, as indicated, to stabilize the hybridization of the fluorophore and quencher-labeled strands and the binding mixtures were equilibrated at room temperature. The signal was then calculated as %F of the no ligand added control divided by plus added ligand test sample. All error bars on all figure data points represent standard deviations.

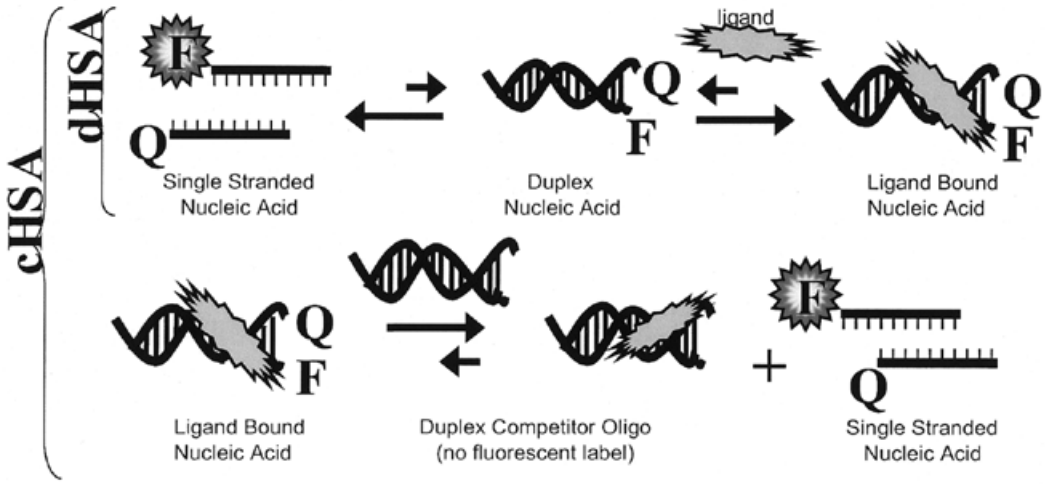
For competition analysis (cHSA) the ligand bound and fluorescence quenched indicator ODNs duplex, prepared as above for dHSA, was titrated with various unlabeled competitor duplexes of varying sequence in 10 mM HEPES–HCl pH 7.2, 50 mM NaCl and 0.1 mM EDTA. The ligand bound duplex and competitors were equilibrated overnight at room temperature (as above) and the various sequences were quantitatively ranked and assessed as %F. %F was calculated based on the following formula:

$$\%F = (F_C - F_Q)/(F_M - F_Q) \times 100$$

where F_C is the relative fluorescence observed for each mix in the presence of competitor and the ligand, F_Q is the relative fluorescence observed in the presence of the ligand without competitor (i.e., quenched), and F_M is the value observed in the absence of ligand and competitor (i.e., the maximal value). %F was calculated for each individual competitor duplex and that value was plotted. All error bars on all figure data points represent standard deviations.

Pseudo-footprinting

'Promoter walk' studies by pseudo-footprinting were carried out in 96-well plates as described above. Samples consisted of 25 nM of the indicator pair ODNs, 50 mM NaCl, 0.1 mM EDTA and 20 mM HEPES-HCl, pH 7.2 and, when indicated, 300 nM of ligand in a total volume of 200 µl. All samples were



Scheme 1.

incubated at ambient temperature and the relative fluorescence was followed until equilibrium was achieved. The competitor duplex ODNs, at 450 nM, were added to the mixture and the fluorescence was monitored (overnight) until the signal stabilized to assure that equilibrium was achieved. %F was calculated for each individual competitor duplex as described above and that value was assigned to the 18 bp sequence corresponding to that competitor ODN. The %F values for each competitor ODN denote the average of three experiments. An average curve of the promoter walk was then calculated as the mean %F value of the sum of each competitor ODN %F value observed for each individual base pair sequence position along the sequence of the promoter. The sequence of the VRE Van H promoter (–115 to 10) used as an example is as follows: TCCTAA GAGATG TATATA ATTTTTT AGGAAAA TCTCAA GGTTAT CTTTAC TTTTTC TTAGGAA ATTAAC AATTTAA TATTAA GAAACGG CTCGTT CTTACA CGGTAG ACTTAA TACCGTA GAACGA.

RESULTS AND DISCUSSION

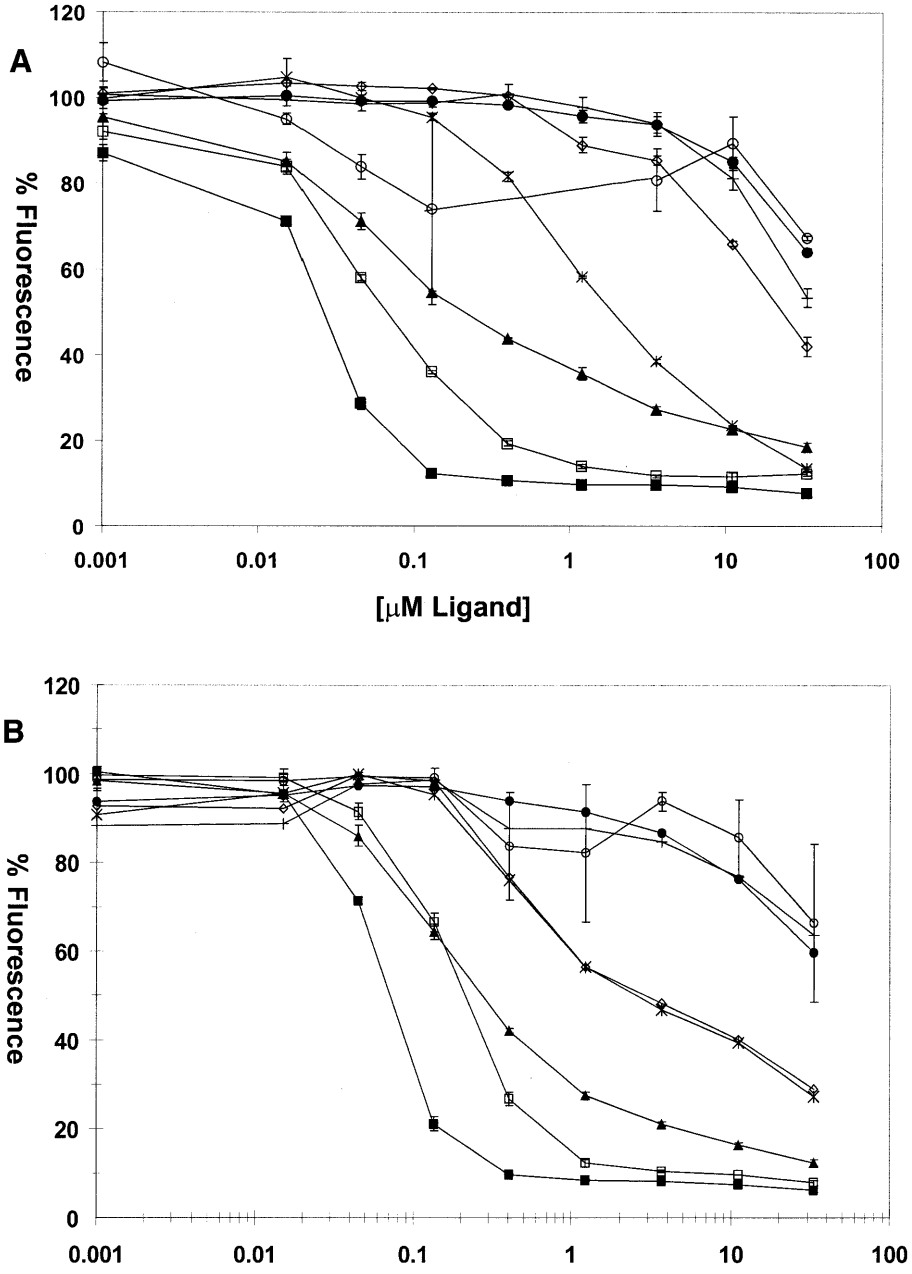
Direct binding fluorescence HSA (dHSA)

The design of the hybridization-stabilization assay, using FRET as a detection modality, is shown in Scheme 1.

When two length matched and sequence complementary ODNs, one with a 5′-fluorescein label and one with a 3′-dabcyl label (or vice versa), are hybridized, the proximity of the dabcyl quencher to the fluorescein fluorescent dye causes a decrease in fluorescence. This 'indicator' ODN pair is designed to have a T_m close to, but a little below, a convenient or relevant temperature (i.e., ambient temperature or 37°C) such that they initially remain predominately single stranded and the fluorescence yield is high. To achieve a common desired T_m of multiple indicator pairs the length of the ODNs may be adjusted, to account for sequence differences, or salt concentration may be adjusted. Ligands that bind to the duplex form of the paired indicator ODNs will promote a shift of the hybridization equilibrium towards the double-stranded, or fluorescence quenched state (dHSA). In studies of interactions of ligands with double-stranded RNA, oligoribonucleotides may be used as the indicator pair.

The ability to discriminate among different indicator ODN sequences and to identify preferred specificities of DNA binding ligands in the direct binding mode of the HSA is shown by the data in Figure 1A–C. This shows the interaction of ligands netropsin (Fig. 1A), distamycin (Fig. 1B) and actinomycin D (Fig. 1C) with indicator duplexes of varying sequence. The sequence of one strand of the duplexes is shown in the legend. Single-stranded controls (binding reactions that only contain a fluorescein-labeled strand and no complementary quenching strand) are also included and are denoted by 'ss' at the end of the sequence. Free fluorescein or 5′-fluorescein conjugated dinucleotides have also been used successfully as controls for non-specific fluorescence quenching (data not shown). The fluorescence signal plotted is normalized to the control signal without ligand. The maximum quenching expected for these duplexes is 85–90% with complete hybridization.

Netropsin (Scheme 2) is a well-studied, dipyrrole polyamide, natural product, minor groove ligand which binds preferentially to A/T-rich regions of DNA in a 1:1 (ligand:single DNA binding site) binding mode (2,17). As shown in Figure 1A, netropsin strongly prefers binding to the indicator sequence 5′-CTTTATTATTTT-3′, as well as to 5′-CTTTTTTT-3′, while it does not bind well to G/C-rich DNA (5′-CGCGCC-3′) or to RNA (5′-rCUAGAUCUGA-3′). Netropsin also binds well to the sequence 5′-GCGGTATTT-3′, which is a useful 'universal' indicator probe as it has short A/T- and G/C-rich regions, and A/T–G/C interface regions. The mixed DNA sequence 5′-CTAGATCTGAAC-3′ binds netropsin moderately well, with an EC_{50} ~50-fold higher than for the all A/T sequence, as it contains 4–5 bp subsites enriched with three to four A/T, although there are no more than 2 bp of contiguous A/T (2,18,19). Figure 1B shows the same type of experiment performed with distamycin A (Scheme 2), a structurally closely related natural product ligand, with similar results. Figure 1C shows the same type of experiment performed with actinomycin D (Scheme 2), an intercalating drug with a weak preference for binding to G/C-rich DNA (1). At concentrations of up to 20 μM, actinomycin D shows a clear preference for



binding to G/C containing, versus A/T containing, duplexes and shows a moderate preference for a poly(C) and mixed sequence duplexes.

The dHSA format may be used to ascertain whether ligands are structurally specific or non-specific nucleic acid binders. As shown in Figure 1, all of the ligands tested are specific for double-stranded DNA over RNA [analogous data for ligands with structural specificity for double-stranded RNA (20) over single-stranded and double-stranded DNA is not shown]. Thus, secondary assay filters may be established in order to assure selection of ligands with desired target specificities (either for RNA or DNA) from initial collections of primary hits obtained from high throughput dHSA screening of libraries of novel compounds (unpublished results).

Competition binding fluorescence HSA (cHSA)

Once the general preference for nucleic acid species (double-stranded DNA or double-stranded RNA) and sequence (if there is a preference) are obtained from the direct binding assay, specific sequence preferences for DNA binding ligands may be determined more precisely by the competition binding mode of the HSA (Scheme 1). A fluorescein/dabcyl-labeled ODN pair, identified using the dHSA as described above, is used as the 'ligand-depot indicator duplex' and various unlabeled, longer, hybridized stable duplexes or folded ODN structures (e.g., stem–loop hairpin ODNs) are used to probe for the ligand's binding site size and sequence preferences. Ligand binding discrimination of subtle structural changes and chemical modifications of nucleic acids also may be probed with appropriate

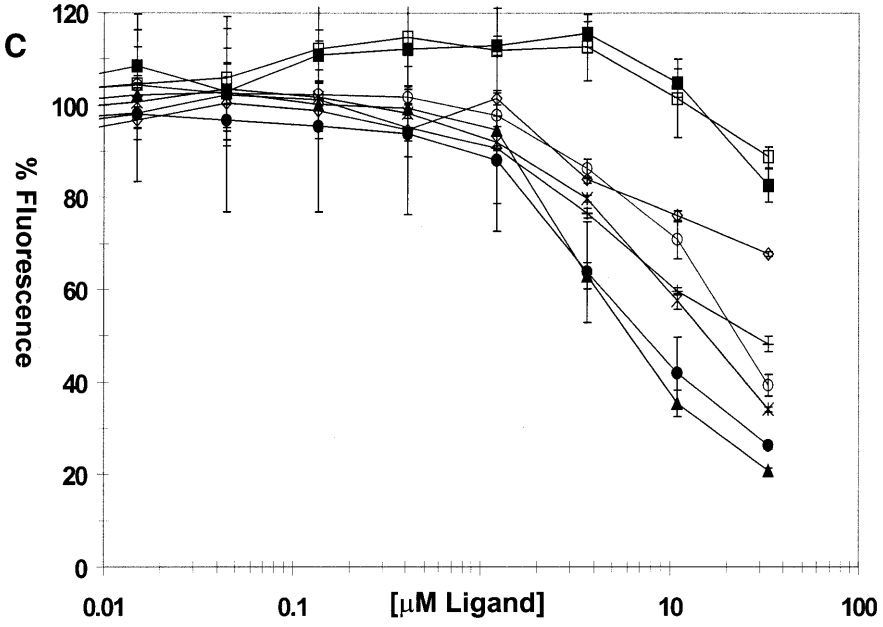


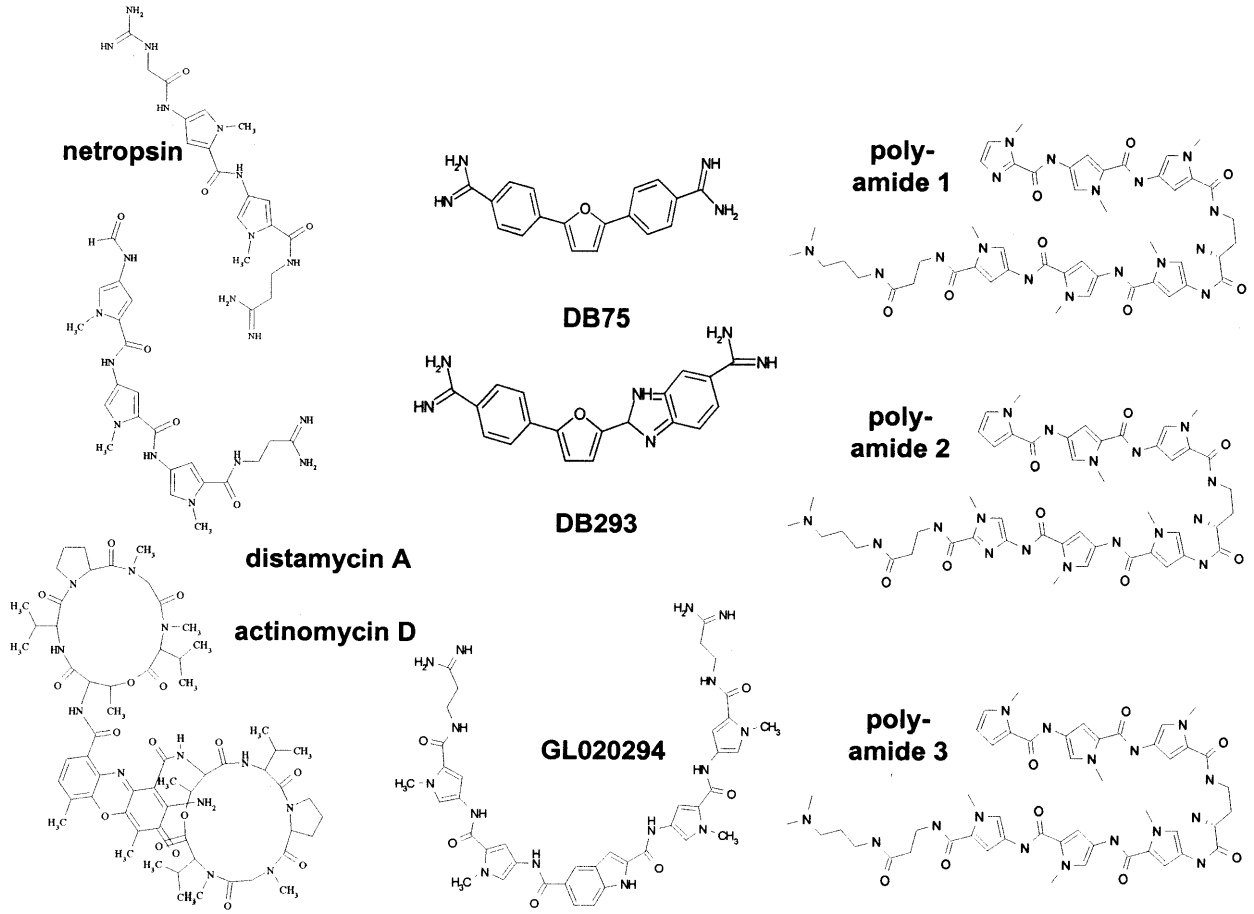
Figure 1. (Previous page and above) dHSA of DNA ligands with various indicator ODN pairs and single strand (ss) ODN and oligoribonucleotide (r) controls. (**A**) Netropsin, (**B**) distamycin A and (**C**) actinomycin D were varied from 0.015 to 33 μM in 210 μl of 10 mM HEPES buffer, pH 7.2, 10 mM NaCl, 0.1 mM EDTA with 25 nM 5′-fluorescein ODN ± 30 nM 3′-dabcyl complementary ODN (where indicated). Mixes were equilibrated at ambient temperature overnight and then 0.2 μg/ml tRNA (final) was added as described in Materials and Methods. Compounds were tested in triplicate and error bars are SD of the mean. Only sequences of the 5′-fluoresceinated strand of indicator ODN pairs and controls are shown: closed squares, CTTATTATTTT; open squares, CTTTTTTTTT; triangles, GCGGTATTT; closed circles, CGCGCC; diagonal crosses, CTAGATCTGAAC; diamonds, rCUAGAUCUGA; vertical crosses, CCCCCCC; open circles, CTTTTTTTTTss. In (C) open circles, CGCGCCss.

competitor ODNs (unpublished results). In this cHSA approach, the concentration of the short, complementary indicator ODNs is held constant. Generally, ligand concentration is fixed at a value giving a high fractional saturation (and fluorescence quenching) of binding to the indicators (i.e., 75–95%) so as to allow for both a large signal response upon addition of high affinity competitors (return of fluorescence) and sensitivity to competitors (by starting at less than ligand saturation binding). Upon addition of duplex competitors an increase in fluorescence signal indicates that the ligand is binding to the unlabeled competitor duplex, releasing the 'indicator' duplex, and causing a shift in the equilibrium back towards a single-stranded conformation.

Dependence on concentration of competing duplex. Figure 2 shows the results of an experiment where a set of duplex competitor nucleic acids, with 4 bp test sequences centrally embedded in conserved, flanking G/C-rich sequences, were titrated over a range of concentration centered on a fixed concentration of the ligand netropsin (Scheme 2). When the concentration of the competitors is close to that of the binding ligand, effective differential partitioning of the ligand, dependent on the sequence (and/or structure) of the competitors, is observed. Typically, the effective range of competitor duplexes is ~0.1–1.9 times the concentration of ligand. Importantly, within this range, the partitioning is insensitive to the actual concentration of the competitors. That is, any single chosen concentration of competitors within this range will give essentially the same relative preferences of ligand binding as any other concentration within the range. Thus, automated high throughput analyses (unpublished results) of relative binding preferences are facilitated, since full titrations of competitors may be avoided and rapid assaying of multiple ligands against multiple competitor sequences (and/or structures) may be performed at fixed concentrations of competitors, ligands and indicators.

Confirmation and extension of dHSA. The experiments of the following four sections show the increase in fluorescence response when fixed concentrations of various competitor DNA ODN duplexes are added to complexes of either netropsin or distamycin bound to an A/T-rich DNA 'indicator' duplex. It is clear that netropsin and distamycin show a strong preference for A/T-rich competitor DNA, supporting results of the dHSA experiments of Figure 1A and B. In addition, more subtle distinctions among A/T sequences also are revealed, consistent with published reports using other assay methods.

Four base pair site competitors. For DNA minor groove binding ligands with a strong preference for A/T over G/C sequences it is feasible to examine the relative binding preferences to all 4 bp A/T binding site sequences. Since neither netropsin nor distamycin A tolerates multiple G/C base pair containing sequences at ligand concentrations that show good binding to A/T sequences, a number of distal, consecutive G/C base pairs may be used to define placement of, or to 'clamp', single 4 bp sites embedded within longer duplexes. Using competitor ODN duplexes of this design (Fig. 3) it was found for both distamycin and netropsin that the most favored 4 bp sequences are AAAA and AATT, consistent with known preferences for uniformly narrow minor grooves (2). However, distamycin is longer than netropsin and A/T sequence variation at only a 4 bp



Scheme 2.

site is unlikely to completely define binding site preferences equivalently for both ligands.

A-tract size determination. In order to more finely characterize the preferred length of the binding site of netropsin and distamycin A, a competition study was done using G/C sequence clamped poly(A)-tracts of increasing length, from 1–8 bp. This data is shown in Figure 4. Both ligands bind relatively weakly to A-tracts of 3 bp or less, but binding sharply increases for 4 bp A-tracts. The minimum length poly(A)-tract supporting uncompromised binding is 4 bp for netropsin and 5 bp for distamycin A. This is consistent with their molecular length and known interactions of the distal functionalities (amidinium, guanidinium, formamide) with the groove and is in accord with previous DNase I footprinting results (2,19,21,22). One would expect that for some ligands the determinations of binding site size by cHSA and chemical or enzymatic footprinting may differ. Footprinting methods measure the number of base pairs physically covered by a ligand that occlude access to the groove by nuclease activity, whereas cHSA is a functional method that determines the base pairs that actually contribute to the energetics of binding. Thus, the two methods are complementary and not redundant.

Nearest neighbors analysis. Since the results of Figures 3 and 4 indicate that distamycin interacts with more than 4 bp, sequences immediately flanking (i.e., nearest neighbors to) the high affinity 4 bp AATT site must be important. To study the sequence determinants at the edges of the 4 bp core elements, a set of ODN duplexes was designed containing all possible sequences in the arrangement 5′-XAATTY-3′, where X and Y are any base. Since AATT is a palindrome there are 10 unique combinations of sequences in this set. For instance, if X and Y are the bases C and A, then the actual binding site is 5′-CAATTA-3′ on one strand, and 5′-TAATTG-3′ on the other strand. The results of testing these 10 new sites in competition binding assays with netropsin and distamycin are shown in Figure 5. These results show that distamycin is much more sensitive to the bases at the edges of the AATT site than is netropsin, which binds reasonably well to all 10. The common elements in all of these sites are summarized as follows. Longer A/T sites are best, that is 6 versus 5 bp, but not greatly so. For X and Y of 6 bp sites, T ≈ A. Within a defined 5 bp site, X and Y show the trend: A > T >> G/C at X and T ≥ A >> G/C at Y. From this we conclude that AAATT is a good 5 bp binding site for distamycin.

Five base pair site competitors. In order to unequivocally define which 5 bp (A+T) binding site is best for these two ligands, the duplexes shown in Figure 6 were tested in the cHSA. This demonstrates that distamycin has relatively strong preferences among the various 5 bp A/T binding sites, and

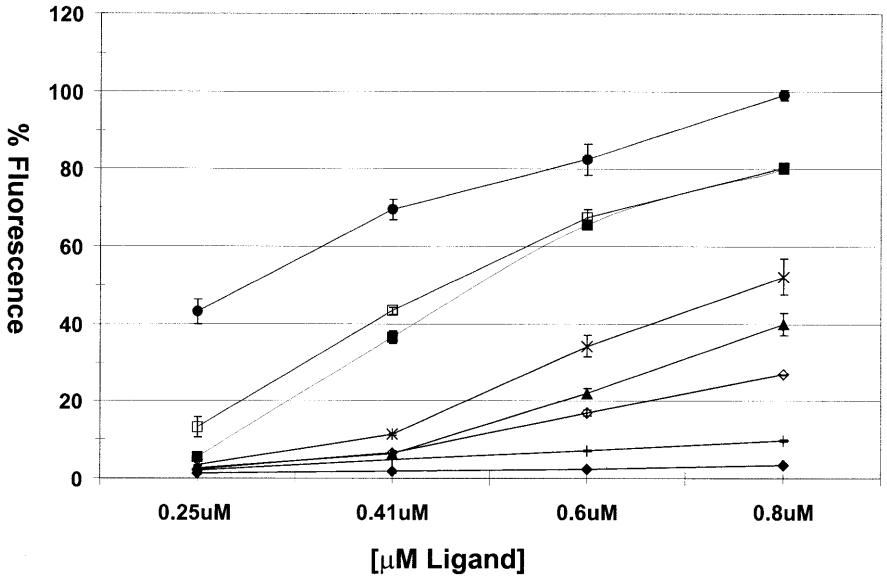


Figure 2. Dependence of netropsin binding on competitor duplex ODN concentration in cHSA. Binding mixes were made and equilibrated as described in Materials and Methods, containing 0.5 μM netropsin and 25 nM 5′-fluorescein CTTATTATTTT/30 nM 3′-dabcyl bottom strand complement indicator ODN pair in 10 mM HEPES pH 7.2, 50 mM NaCl, 0.1 mM EDTA and 0.25–0.8 μM competitor duplex ODN with centrally placed 4 bp A/T binding site clamped by distal G/C bp, CCCG(A/T)$_4$CGCC (top strand): closed square, -AAAA-; octagon, -AATT-; closed diamond, -TATA-; open square, -ATAT-; triangle, -TATT-; open diamond, -TTAA-; diagonal cross, -TAAT-; vertical cross, -TGTT-. Compounds were tested in duplicate.

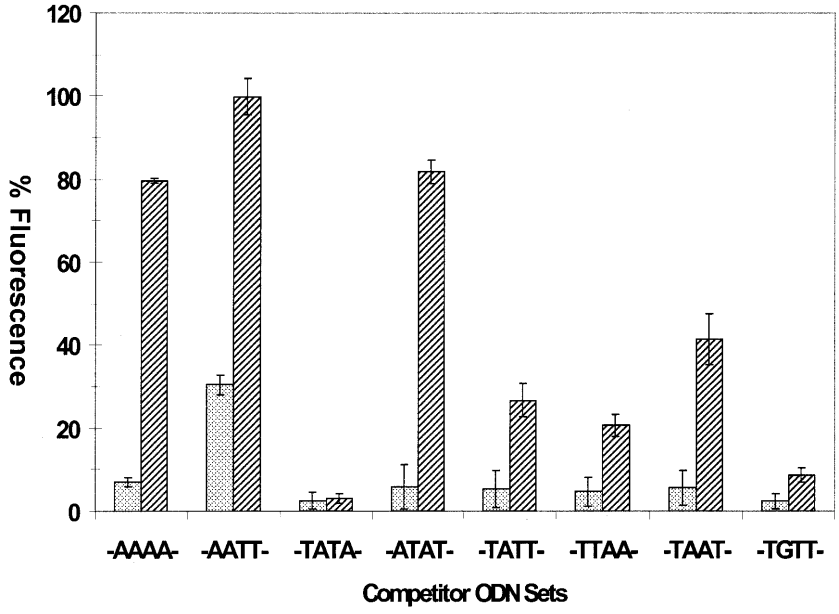


Figure 3. Dependence of distamycin A (dotted bars) and netropsin (hatched bars) binding to 4 bp A/T sequences in competitor ODNs by cHSA. Binding mixes were made and equilibrated as described in Materials and Methods containing 0.5 μM netropsin or distamycin A and 25 nM 5′-fluorescein/30 nM 3′-dabcyl indicator ODN pair (as in Fig. 2) in 10 mM HEPES pH 7.2, 50 mM NaCl, 0.1 mM EDTA and 0.6 μM competitor duplex ODN with centrally placed 4 bp A/T binding site clamped by distal G/C bp, CCCG(A/T)$_4$CGCC (top strand). Compounds were tested in duplicate.

appears to bind best to the 5′→3′ sites AAAAA > AAATT ≥ AAAAT ≈ AATTA, generally consistent with the experiment discussed above (Fig. 5). It is known that 5′→3′ TA steps create disruptions in the structure of the otherwise uniformly narrow minor groove of A/T sequences most preferred by the natural polyamides netropsin and distamycin (2,23). The results of Figure 6 also show that a TA step at the 3′-end of the ligand binding site is least damaging. A TA step at the 5′-end or TA steps in the central portion of the binding site are the most damaging to binding. In this experiment netropsin demonstrates less discrimination than distamycin between the various 5 bp sites, with at least moderate binding to all sites and a minor preference for AAAAA, AAATT, TAAAT, AATTA and AAAAT sites (Fig. 6). The degeneracy of binding

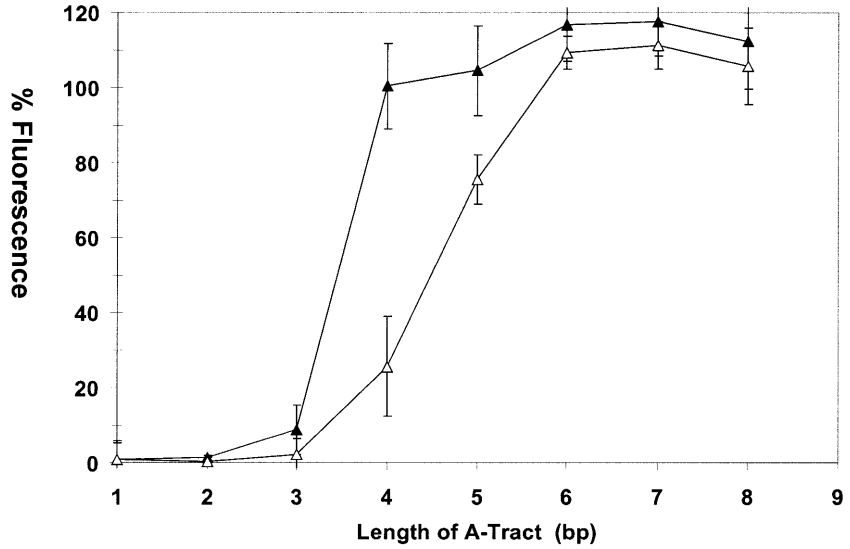


Figure 4. Determination of minimum A-tract binding site length for netropsin (closed triangle) and distamycin A (open triangle) by cHSA. Binding mixes were made and equilibrated as described in Materials and Methods, containing 0.5 μM netropsin or distamycin A and 25 nM 5′-fluorescein/30 nM 3′-dabcyl indicator ODN pair (as in Fig. 2) in 10 mM HEPES pH 7.2, 50 mM NaCl, 0.1 mM EDTA and 0.45 μM competitor duplex ODN containing 0–6 A (top strand): CCCGGA $_{0–6}$ CGCGCC. Experiments were performed in triplicate.

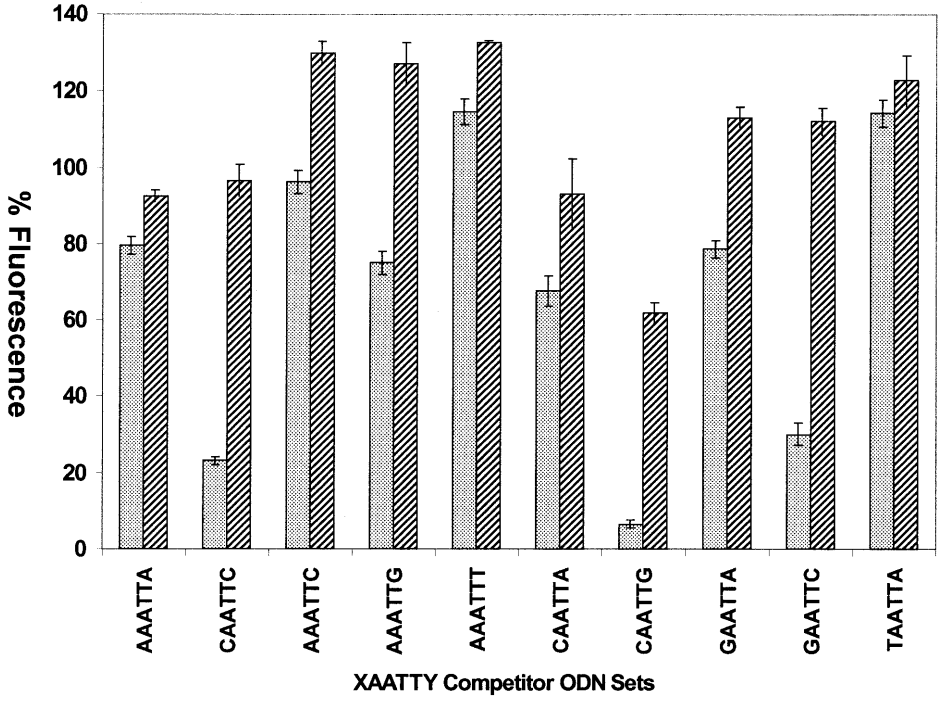


Figure 5. Nearest neighbors analysis of binding of distamycin A (dotted bars) and netropsin (hatched bars) to the AATT sequence by cHSA. Binding mixes were made and equilibrated as described in Materials and Methods, containing 0.5 μM netropsin or distamycin A and 25 nM 5′-fluorescein/30 nM 3′-dabcyl indicator ODN pair (as in Fig. 2) in 10 mM HEPES pH 7.2, 50 mM NaCl, 0.1 mM EDTA and 0.45 μM competitor duplex ODN (top strand): CCCXAATTYGCC. Experiments were performed in triplicate.

register within a 5 bp A/T binding site may obscure some of the sequence specificity of the 4 bp A/T binding netropsin (Figs 2 and 3) compared to the specificity of binding of distamycin, with a single register in a site of this size.

Unlinked, stacked ligands. It should prove possible to use the competition format of the HSA to ascertain the binding preferences of ligands of divergent designs and modes of recognition. The only requirement should be that the nucleic acid binding site [or flanking region(s)] has at least partial duplex structure that is stabilized upon ligand binding. Recently, Wilson and co-workers (24,25) have designed and synthesized an interesting new class of dicationic DNA minor groove binding heterocycles that are not conjugated through

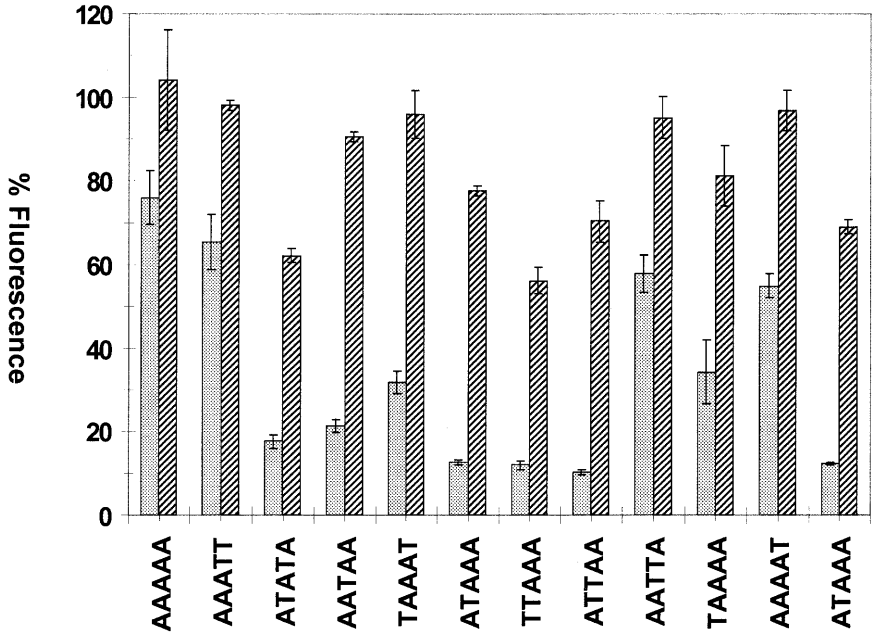


Figure 6. Binding of distamycin A (dotted bars) and netropsin (hatched bars) to 5 bp A/T sequence sites by cHSA. Binding mixes were made and equilibrated as described in Materials and Methods, containing 0.5 μM netropsin or distamycin A and 25 nM 5′-fluorescein/30 nM 3′-dabcyl indicator ODN pair (as in Fig. 2) in 10 mM HEPES pH 7.2, 50 mM NaCl, 0.1 mM EDTA and 0.45 μM competitor duplex ODN (top strand): $CCCG(A/T)_5CCG$. Experiments were performed in triplicate.

polyamides. In particular, one asymmetric furamidine dication, DB293 (Scheme 2), was shown by surface plasmon resonance, DNase I footprinting, CD and thermal melt analysis to bind in a non-covalently self-assembled, ring-stacked, 2:1 stoichiometry/binding site with a preference for G/C containing sites. The results of cHSA analysis of this furamidine and a related A/T sequence preferring, symmetric furamidine dication, DB75, as well as of netropsin and distamycin for comparison, are shown in Table 1. The cHSA preferences obtained closely parallel the specificities of the original published report. Compound DB293 displays a modest preference for G/C containing sequences and all the other compounds prefer A/T sequences.

Covalently linked, hairpin stacked ligands. Progress towards the design of minor groove binding compounds capable of recognizing mixed DNA sequences with single base pair resolution has recently been reviewed for covalently linked, ring-stacked, hairpin polyamides of five-membered aromatic nitrogen heterocycles pyrrole (Py) and imidazole (Im) (and *N*-methyl and hydroxyl functionalizations thereof) (2,26). This class of compounds can achieve a remarkable level of base pair recognition through a side-by-side, stacked binding mode in the minor groove, in which the Im/Py stack binds preferentially to G:C base pairs, via specific hydrogen bonding of Im and G, whereas the Py/Py stack prefers A/T base pairs (27). It would be of interest and potential value to ascertain if the recognition code can be expanded beyond these two ring systems (2,26,28–30).

The specificity of a stacked polyamide is defined as the ratio of its thermodynamic dissociation constants for two different—a matched and a mismatched—DNA sequences. The binding constants typically are determined using DNase I footprinting (30,31). It would be desirable to efficiently characterize the binding specificity of a large number of novel compounds of combinatorial syntheses against a comprehensive set of sequences, a task for which the gel electrophoresis-based DNase I footprinting is ill suited.

Table 1. Recognition of G/C base pairs by non-covalent self-stacking of a furamidine dication minor groove binding ligands determined by cHSA

Sequence	Ligand			
	DB293	DB293	Netropsin	Distamycin A
CccTAT**G**Accc[a]	82	37	24	30
cccAT**G**ATTccc	79	44	21	28
cccTAT**G**ACCAcc	83	44	17	23
cgAA**TT**cg-(HP)	57	67	50	51

[a]Binding mixtures were made and equilibrated as described in Materials and Methods, containing 2 μM ligand and 25 nM 5′-fluorescein/30 nM 3′-dabcyl indicator ODN pair (as in Fig. 2) in 10 mM HEPES pH 7.2, 50 mM NaCl, 0.1 mM EDTA and 1.75 μM competitor duplexes ODN (top strand as shown: HP indicates a hairpin competitor ODN was used). Experiments were performed in triplicate.

Figure 7 illustrates how cHSA could be implemented in order to systematically determine the relative sequence preferences for binding of novel variants of stacked hairpin polyamide ligands. The six-ring compounds **1–3** of Scheme 2 stack with three rings side-by-side with three other rings (3 + 3) in the minor groove of a hairpin ODN, which is designed such that there are strong constraints on binding register and polarity, with the N-terminus oriented in the 5′-direction (Fig. 7A). Using all 16 base pair combinations for positions NN of the hairpin ODN, a complete profile can be obtained for the

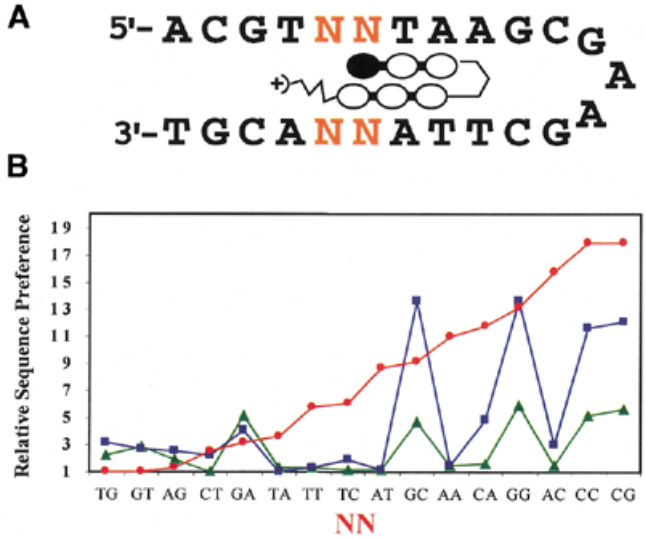


Figure 7. SPPs of hairpin linked polyamides by cHSA. Structures of hairpin polyamides **1–3** are shown in Scheme 2. (**A**) Sequence of 25 nt DNA hairpin used as competitor ODN in cHSA experiments. All 16 bp combinations were realized for NN. Hairpin competitor ODNs were designed to exhibit $T_m > 75°C$ to minimize interference with indicator ODN. Taking into consideration that the diaminobutyric acid linker and the β-alanine dimethylpropylamine tail of stacked polyamides **1–3** (Scheme 2) have a known preference for A/T base pairs, polyamides **1–3** are expected to bind with the unique register shown. (**B**) SPPs from cHSA for polyamides **1** (cirlces), **2** (triangles) and **3** (squares). Each data point for every NN sequence for each compound represents the ratio of the cHSA competitor hairpin ODN concentration observed to give 50% fluorescence recovery (IC_{50}) and the IC_{50} for the one NN sequence with the highest affinity interaction for a given compound. In (B) data are sorted according to an ascending profile (descending NN sequence preference) of the chosen reference compound, polyamide **1**. Indicator ODN pair used was CTTATTATTTT at a fixed concentration of 30 nM. Competitor hairpin 25 nt ODNs were used at 10 μM final concentration.

sequence preference of new test recognition elements placed at one, or both, of the paired, stacked N- and C-terminal ring positions of a 3 + 3 hairpin polyamide (Fig. 7A). In similar fashion, sequence preferences of new test recognition elements could be queried at other positions of the 3 + 3 hairpin polyamide or at various positions of alternative hairpin polyamides (i.e., 4 + 4, 3 + 4, etc.). Such systematic analyses readily lend themselves to automation (unpublished results).

From published DNase I footprinting studies, polyamide 1 (32) would be expected to have the highest affinity for hairpin ODNs of Figure 7A having NN = AG and TG, involving hydrogen bonding of Im of the Im/Py stack with G of the G:C base pair. Specificity of the tail for A/T should position sequences NN = GN and CN towards the other end of the affinity scale and sequences NN = CC and GC would be expected to be the lowest affinity, as they accommodate neither the tail nor the Im/Py stack. The sequence preference profile (SPP) of polyamide **1** (Fig. 7B) indeed matches these expectations. All four sequences with solely A/T pairs have ~4–12 times lower relative preference for this 'guanine specific' polyamide **1**. The fact that NN = GT also is among the most preferred sequences is easily explained by a shift of the binding register by 1 bp towards the 5′-end.

Although polyamides **2** and **3** are obvious structural variants in a systematic study of Im and Py based 3 + 3 hairpin polyamides, no published data is available for direct, external comparison. According to the stacked polyamide base pair recognition code (32) sequences with solely A/T pairs would be expected to be most preferred for polyamide **3** and sequences with NN = AC and TC should be predicted to bind most tightly to polyamide **2**. SPPs in Figure 7B clearly show that for both of these polyamides the specificity is much less pronounced than is expected from the published 'rules'. Remarkably, polyamides **2** and **3** have very similar SPPs and no guanine specificity is observed for polyamide **2** (NN = TA, TT, AT and AA bind just as well as NN = TC and AC). This result is very similar to what was described recently for a 4 + 4 eight-ring polyamide (33). In this case, traditional footprinting methods revealed the loss of G specificity when Im assumed the position next to the tail. For our all-Py compound **3**, values for NN = TC and AC are only marginally higher than for the solely A/T containing sequences. On the other hand, polyamides with a Py in the N-terminal position (**2** and **3**) bind less tightly to sequences with NN = TG and AG, showing that Py cannot easily be placed against a G, similar to the finding with polyamide **1** that Im does not like to be placed against A/T. Most surprising, however, is that the Im at the C-terminus of the polyamide **2** has lost all specificity for G. Thus, even this simple study reveals that, although the hairpin polyamide sequence recognition rules were not broken, per se (i.e., there were no observed contrary specificities), they appear to be more like guidelines. Whether there is realization of specificity discriminations of individual binding elements predicted by the rules is dependent in complex ways on polyamide structure and target sequence context.

cHSA preferences versus thermodynamic specificities. The cHSA requires non-ideal, high concentrations of competitor ODNs, typically ~$10 \times K_d$ of ligands for preferred sequences: if indicator ODN duplexes are bound with high affinity, initially are 90% fractionally bound with ligand, and competitor ODN duplexes are added at concentrations equimolar with ligand concentrations. Nevertheless, the values for the relative sequence preferences (the ratios of IC_{50} values for matched versus mismatched sequences) from cHSA of the ring stacked dicationic furamidine compounds (<2-fold; Table 1) and the linked hairpin polyamides (up to ~10-fold; Fig. 7B) are not hugely compressed compared to the values for the relative sequence specificities (ratios of apparent K_d values for matched versus mismatched sequences) obtained from thermodynamic studies [<2.5-fold (24) and 8.6-fold (34), respectively]. This conclusion is supported further by the correctly discerned subtle A/T sequence preferences obtained by cHSA for the natural polyamide ligands netropsin and distamycin. Thus, the experimental design requirement of cHSA for competitor duplex concentration to be at, or above, the K_d for binding of the ligand to preferred sequences does not appear to significantly compromise the quality of sequence discrimination results that can be obtained with the method.

Pseudo-footprinting generated promoter walks. The relative binding preferences of ligands along longer DNA sequences, such as those known to encompass functionally important protein binding sites of transcription factors in promoter regions of genes of interest, may prove useful for studies of gene structure, functional genetics, and small molecule control of gene expression. The solution-based cHSA format alternative

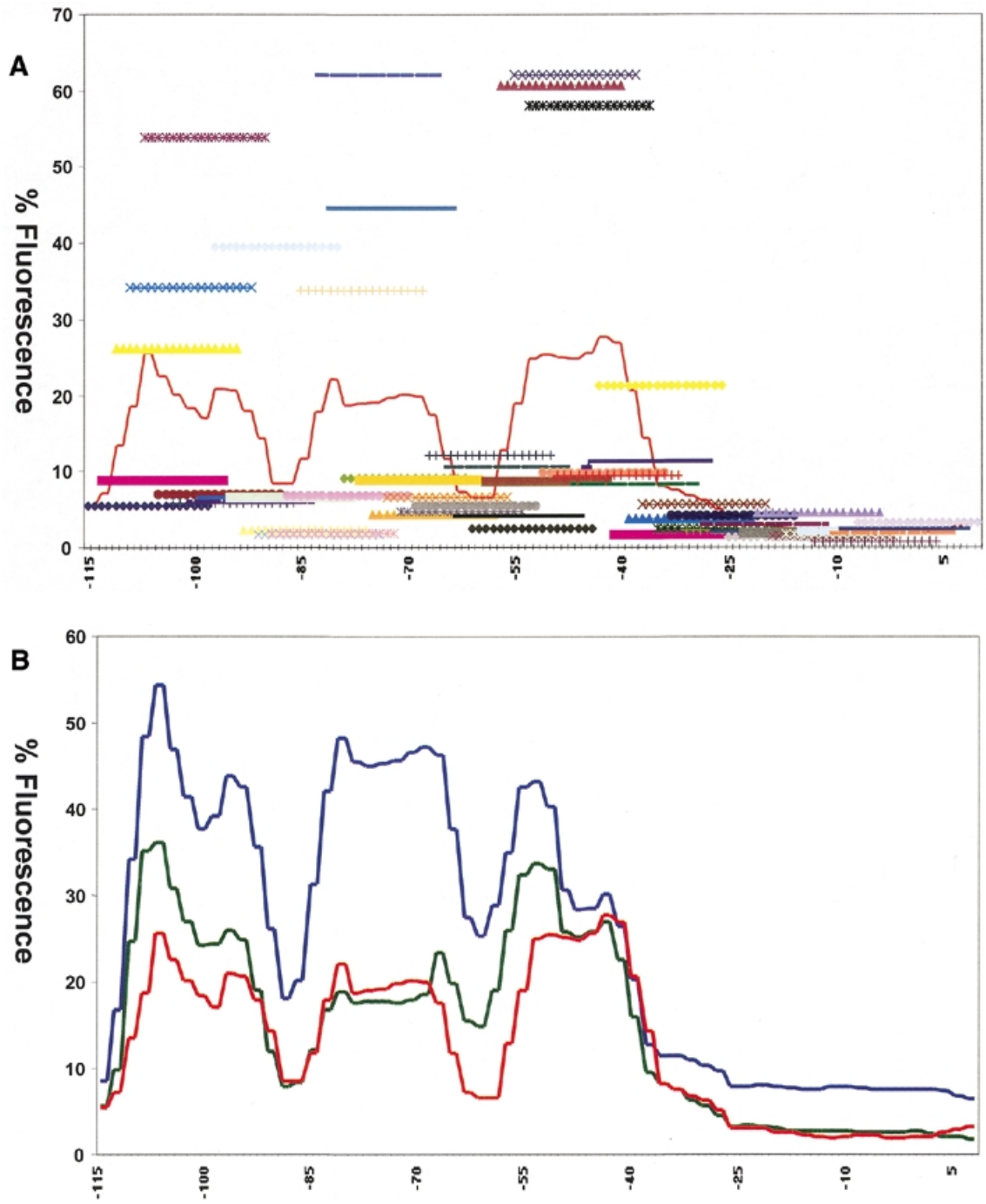


Figure 8. VRE Van H promoter walk by cHSA pseudo-footprinting. Experiments and analyses were performed as described in Materials and Methods. (**A**) Randomly colored horizontal bars represent the cHSA measured %F values of 18mer ODN competitor duplexes utilizing compound GL020294 (see Scheme 2). Competitor duplexes correspond to the 18 bp portion of the Van H promoter they lie directly above. From –115 to +10 of the Van H promoter consecutive 18mer ODN duplexes were walked, overlapping by 16 bp at each step. The curve of the average %F value at 2 bp resolution for compound GL020294 is shown in red. (**B**) Average curves, at 2 bp resolution, of the average %F value of all (i.e., 9, except at the ends of the walk) 18mer ODN competitors at each base pair position are shown for netropsin (blue), distamycin A (green) and GL020294 (red) for reference.

to gel-based footprinting is easy to perform and may be multiplexed and automated.

An example of the type of results that may be obtained is shown in Figure 8. In this experiment an A/T-rich indicator DNA duplex (5′-CTTTATTATTT-3′) was chosen that was shown to be stabilized by the addition of ligands of interest. A series of sequence degenerate competitor ODN pairs, corresponding to the highly sequence conserved region –115 to +10 from the Van H promoter of a clinical isolate of vancomycin resistant *Enterococcus faecium* (VRE) (unpublished results), were designed. Each competitor duplex was 18 bp long and overlapped the preceding one by 16 bp. All competitor

duplexes, considered together, generate a 'walk', in 2 bp steps, in the region of the Van H promoter examined. The smoothed curves, at 2 bp resolution, represent the average of the individual values of nine different competitor sequences that encompass each 2 bp step (Fig. 8A). They allow for the ready visualization of preferred sites in the sequence of interest, very much like digitized presentations of footprinting data, hence the term pseudo-footprinting. Promoter walks of numerous different ligands may conveniently be performed in parallel without the use of radioactive labels and gels and the results can be depicted simultaneously to facilitate direct comparisons (Fig. 8B). In our example in Figure 8, results are shown for netropsin (a dipyrrole polyamide), distamycin (a tripyrrole polyamide) and GL020924 (an indole-linked dimer of netropsin-like dipyrrole polyamides) (Scheme 2), of which all are A/T binding ligands. Contiguous, minimum A/T sequence binding site size requirements for these ligands are: netropsin, 4 bp and distamycin, 5–6 bp, and GL020924, 8 bp (35). In the nearly 80% A/T content Van H promoter of VRE there are multiple binding sites for these A/T specific ligands (Fig. 8B). Interpretation of the pseudo-footprinting patterns on the single base resolution scale is not always straightforward, probably due to a complex interplay of multiple factors, including modulation of ligand affinities by (i) sequence dependent structural polymorphism in the minor groove (36), (ii) distortion of minor groove structure by ligand binding simultaneously at multiple A/T rich sites (37) and (iii) alteration of ligand binding site structure by adjacent sequences (7). However, some generalities, with potentially important implications for control of gene expression by small molecules, are evident. Although the binding patterns are similar for all three polyamides, the overall effect of going from minimal A/T binding site size requirements of 4 to 5–6 to 8–10 bp is a progressive smoothing and dampening of differential sequence discrimination.

Extrapolating from these observations, it is probable that appropriately designed minor groove ligands with somewhat longer binding sites (8–12 bp) and higher affinities for most preferred sites will provide for greater biological selectivity of modulation of gene expression at the level of transcription than 4–6 bp binding ligands due to the combined effects of attenuated binding at all of the vastly greater number of suboptimal sequences and potentiated binding to the longer, hence, rarer optimal binding sequence sites.

General applicability of HSA methodology. Although the direct implementation of our assay (dHSA) is more readily set up and interpreted, it is clearly limited to nucleic acid systems for which ligand binding produces a significant shift in the hybridization equilibrium. This, however, does not necessarily imply that only canonical DNA and RNA systems can be studied, which indeed embodied our main interest. For example, ligand binding to duplexes with mismatches, bulged nucleotides or loop motifs should be readily detectable if the hybridization equilibrium is affected. In cases where mismatches and bulged motifs lead to decreased thermodynamic stability, a ligand's contribution to duplex stability is even more easily determined, such that the dHSA methodology should be especially suitable for ligand screening against RNA motifs involving small bulges or mismatches.

The competition format of the hybridization-stabilization assay provides a versatile method to characterize ligand preferences for a panel of nucleic acids. The diversity of the panel could range from single nucleotide mutations (sequence specificity profiling) to vastly different nucleic acid motifs (e.g., RNA versus DNA or modified ODNs). Since the cHSA methodology relies on the duplex indicator system it is essential to use competitor nucleic acids that do not interact with the indicator ODNs. In general, choosing competitor nucleic acids that are almost completely hybridized at the assay temperature should take care of this problem. Therefore, longer duplexes and hairpins, if applicable, are most suitable as competitor nucleic acids.

A possible source of error arises from the fluorescence-based method of detection; if the chemical nature of a ligand induces changes in the fluorescence of the labeled indicator system, results are difficult to interpret. Whenever ligand titration profiles in dHSA experiments deviate from ideal sigmoidal curves, fluorescence interference might be involved. In those cases, ligand interaction with the individual labeled ODN strands should be analyzed. Note that due to the comparative nature of the cHSA format fluorescence interference is much less of a problem in the latter case.

CONCLUSIONS

In this report we have introduced a simple and versatile assay for identifying and characterizing ligands that bind to (substantially) duplexed nucleic acids and have provided validation using select ligands representative of diverse structural types previously well characterized by other methods. The assay is based on the known stabilization of hybridization of sequence complementary nucleic acid strands upon ligand binding. Unlike the traditional thermal melt methods for measuring this effect, the assay described here is isothermal. It also is solution-based and amenable to detection by multiple modalities, such as FRET, described herein. Thus, it is readily adaptable to multiplexing and modern formats of automation for rapid screening for prospective novel ligands and for subsequent elucidation of ligand binding preferences for nucleic acid structure and sequence, in fine detail. Screening is accomplished using the dHSA and a finer characterization with a layered dHSA-to-cHSA approach. The assay has been used here to identify DNA specific ligands of different chemical classes, to determine DNA binding site size and sequence preference profiles of select ligands, to perform systematic SAR of hairpin polyamide ligands for effects of chemical variation on binding sequence preferences (with some surprising results), and to identify preferred ligand binding sites in longer DNA sequences of a bacterial resistance gene promoter by pseudo-footprinting. Although the cHSA format is performed using thermodynamically non-ideal concentrations of competitor nucleic acids, it has been shown to provide a relative quantitative assessment of binding sequence preferences that accurately recapitulates known A/T from G/C, and even subtle A versus T, discriminations without adverse compression of sequence preference values in comparison to values for sequence specificities obtained from thermodynamic methods. This new assay methodology should facilitate the rapid discovery of novel nucleic acid ligands as research reagents

and as potential small molecule drugs useful for the control of gene expression and inhibition of gene function.

ACKNOWLEDGEMENTS

The authors wish to thank Profs David W. Boykin and W. David Wilson for the gracious gift of dicationic aromatic minor groove binding compounds, Mark Velligan for the sequence of the Van H promoter of VRE, Janos Botyanszki for synthesis of hairpin polyamides, Chris Roberts for synthesis of GL020924, Kirsten Nygren, Jason Lacombe and Nanci Alexi for using the assay in Genelab's screening program. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) Unconventional Pathogens Countermeasures (UPC) program, grant N65236-98-1-5400.

REFERENCES

1. Turner,P.R. and Denny,W.A. (2000) The genome as a drug target: sequence specific minor groove binding ligands. *Current Drug Targets*, **1**, 1–14.
2. Wemmer,D.E. (2000) Designed sequence-specific minor groove ligands. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 439–461.
3. Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules*. W. H. Freeman, New York, NY.
4. Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry Part II: Techniques for the Study of Biological Structure*. W. H. Freeman, New York, NY.
5. Cantor,C.R. and Schimmel,P.R. (1980) *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*. W. H. Freeman, New York, NY.
6. Wilson,W.D., Tanious,F.A., Fernandez-Saiz,M. and Rigl,C.T. (1997) In Fox,K.R. (ed.), *Methods in Molecular Biology*. Humana Press, Totawa NJ, Vol. 90, pp. 219–240.
7. Bostock-Smith,C.E., Harris,S.A., Laughton,C.A. and Searle,M.S. (2001) Induced fit DNA recognition by a minor groove binding analogue of Hoechst 33258: fluctuations in DNA A tract structure investigated by NMR and molecular dynamics simulations. *Nucleic Acids Res.*, **29**, 693–702.
8. Cantor,C.R. (1996) Lighting up hybridization. *Nat. Biotechnol.*, **14**, 264.
9. Harrison,J.G. and Balasubramanian,S. (1998) Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates. *Nucleic Acids Res.*, **26**, 3136–3145.
10. Heid,C.A., Stevens,J., Livak,K.J. and Williams,P.M. (1996) Real time quantitative PCR. *Genome Res.*, **6**, 986–994.
11. Livak,K.J., Flood,S.J., Marmaro,J., Giusti,W. and Deetz,K. (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl.*, **4**, 357–362.
12. Morrison,L.E. and Stols,L.M. (1993) Sensitive fluorescence-based thermodynamic and kinetic measurements of DNA hybridization in solution. *Biochemistry*, **32**, 3095–3104.
13. Tyagi,S., Bratu,D.P. and Kramer,F.R. (1998) Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.*, **16**, 49–53.
14. Baird,E.E. and Dervan,P.B. (1996) Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.*, **118**, 6141–6146.
15. Owczarzy,R., Vallone,P.M., Gallo,F.J., Paner,T.M., Lane,M.J. and Benight,A.S. (1997) Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers*, **44**, 217–239.
16. SantaLucia,J.,Jr, Allawi,H.T. and Seneviratne,P.A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
17. Bailly,C. and Chaires,J.B. (1998) Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues. *Bioconjug. Chem.*, **9**, 513–538.
18. Portugal,J. and Waring,M.J. (1987) Comparison of binding sites in DNA for berenil, netropsin and distamycin. A footprinting study. *Eur. J. Biochem.*, **167**, 281–289.
19. Fox,K.R. and Waring,M.J. (1987) The use of micrococcal nuclease as a probe for drug-binding sites on DNA. *Biochim. Biophy. Acta*, **909**, 145–155.
20. Chen,Q., Kuntz,I.D. and Shafer,R.H. (1996) Spectroscopic recognition of guanine dimeric hairpin quadruplexes by a carbocyanine dye. *Proc. Natl Acad. Sci. USA*, **93**, 2635–2639.
21. Fox,K.R. and Waring,M.J. (1987) Footprinting at low temperatures: evidence that ethidium and other simple intercalators can discriminate between different nucleotide sequences. *Nucleic Acids Res.*, **15**, 491–507.
22. Portugal,J. and Waring,M.J. (1987) Hydroxyl radical footprinting of the sequence-selective binding of netropsin and distamycin to DNA. *FEBS Lett.*, **225**, 195–200.
23. Bewley,A.B., Gronenborn,A.M. and Clore,G.M. (1998) Minor groove-binding architectural proteins: structure, function and DNA recognition. *Annu. Rev. Biophy. Biomol. Struct.*, **27**, 105–131.
24. Wang,L., Bailly,C., Kumar,A., Ding,D., Bajic,M., Boykin,D.W. and Wilson,W.D. (2000) Specific molecular recognition of mixed nucleic acid sequences: an aromatic dication that binds in the DNA minor groove as a dimer. *Proc. Natl Acad. Sci. USA*, **97**, 12–16.
25. Wang,L., Carrasco,C., Kumar,A., Stephens,C., Bailly,C., Boykin,D.W. and Wilson,W.D. (2001) Evaluation of the influence of compound structure on stacked-dimer formation in the DNA minor groove. *Biochemistry*, **40**, 2511–2521.
26. Wemmer,D.E. and Dervan,P.B. (1997) Targeting the minor groove of DNA. *Curr. Opin. Struct. Biol.*, **7**, 355–361.
27. White,S., Baird,E.E. and Dervan,P.B. (1997) On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem. Biol.*, **4**, 569–578.
28. Ellervik,U., Wang,C. and Dervan,P. (2000) Hydroxybenzamide/pyrrole pair distinguishes T·A from A·T base pairs in the minor groove of DNA. *J. Am. Chem. Soc.*, **122**, 9354–9360.
29. Minehan,T., Gottwald,K. and Dervan,P. (2000) Molecular recognition of DNA by Hoechst benzimidazoles: Exploring beyond the pyrrole-imidazole-hydroxypyrrole polyamide-pairing code. *Helv. Chim. Acta*, **83**, 2197–2213.
30. Zhan,Z.Y. and Dervan,P.B. (2000) Alternative heterocycles for DNA recognition: a 3-pyrazole/pyrrole pair specifies for G.C base pairs. *Bioorg. Med. Chem.*, **8**, 2467–2474.
31. Mrksich,M., Wade,W.S., Dwyer,T.J., Geierstanger,B.H., Wemmer,D.E. and Dervan,P.B. (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.
32. White,S., Szewczyk,J.W., Turner,J.M., Baird,E.E. and Dervan,P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468–471.
33. Wang,C., Ellervik,U. and Dervan,P. (2001) Expanding the recognition of the minor groove of DNA by incorporation of beta-alanine in hairpin polyamides. *Bioorg. Med. Chem.*, **9**, 653–657.
34. Pilch,D.S., Poklar,N., Gelfand,C.A., Law,S.M., Breslauer,K.J., Baird,E.E. and Dervan,P.B. (1996) Binding of a hairpin polyamide in the minor groove of DNA: sequence-specific enthalpic discrimination. *Proc. Natl Acad. Sci. USA*, **93**, 8306–8311.
35. Laurance,M., Starr,D., Michelotti,E., Cheung,E., Gonzalez,C., Tam,A., Deikman,J., Edwards,C. and Bardwell,A. (2001) Specific down-regulation of an engineered human cyclin D1 promoter by a novel BRA-binding ligand in intact cells. *Nucleic Acids Res.*, **29**, 652–661.
36. Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer Verlag, New York, NY.
37. Benight,A.S., Gallo,F.J., Paner,T.M., Bishop,K.D., Faldasz,B.D. and Lane,M.J. (1995) Sequence context and DNA reactivity: application to sequence-specific cleavage of DNA. *Adv. Biophys. Chem.*, **5**, 1–55.