Chapter 1

Programmable DNA Binding Oligomers for Control of Transcription

The text of this chapter was taken in part from a manuscript coauthored with Raymond M. Doss and Professor Peter B. Dervan (Caltech)

(Dervan, P. B.; Doss, R. M. and Marques, M. A. "*Programmable DNA Binding Oligomers for Control of Transcription*" *Current Medicinal Chemistry* **2005** – in press.)

Abstract.

With a rapid movement toward personalized genetic medicine, tailoring treatment to the individual needs of a patient based on their genes is becoming more common. The ability to develop small molecules capable of reprogramming the cellular machinery is one approach to the difficult challenge of treating diseases that result from aberrant gene expression. Inspired by the architecture of the natural products netropsin and distamycin, polyamides are capable of binding the DNA minor groove with high affinity and fidelity. Originally composed of 5-membered heterocyclic carboxamides, polyamides have evolved in both form and function. Second-generation compounds containing novel heterocyclic recognition elements, alkylating agents, intercalators, gene activation domains, and fluorescent labels for cell trafficking studies, have recently been developed. These molecules are successful at gene repression by inhibition of the transcriptional machinery of the cell, gene activation by inhibition of repressor complexes, and activation by serving as artificial transcription factors. These rationally designed oligomers offer one approach to the challenging problem of regulating gene expression.

Introduction.

With the completion of the Human Genome Project, biology can be viewed as an informational science. The digital information in the genome (DNA) encodes the logic of life. These genes encode protein networks. Environmental cues from outside the genome (such as transcription factors) control these networks. A major goal for chemistry, biology, and human medicine would be to ask the question whether human disease could be controlled by targeting gene expression, i.e., manipulate with small molecules the information and software programs encoded in nucleic acids that control protein networks. There are two general approaches; molecules that target mRNA (antisense, RNAi, PNA) or molecules that target the gene DNA, and inhibit mRNA synthesis by interfering with components of the RNA polymerase transcription machinery. For these approaches to be successful *in vivo*, the gene regulatory agent must be cell-permeable. In the case of DNA binding molecules, the regulatory molecule must traffic to the nucleus and bind its target sequence with high affinity and specificity in the context of cellular chromatin. DNA binding must interfere with transcription of the target gene by inhibition of key transcription factors in the promoter or alternatively a steric blockade in the coding region. Nucleic acid based approaches that target either DNA or RNA (antisense, triple helix forming oligonucleotides, ribozymes and siRNA) and engineered zinc fingers have the potential for sequence specificity and can effectively inhibit transcription or translation *in vitro*: however, oligonucleotides and proteins suffer from poor cell permeability and delivery strategies, such as viral vectors, must be used for effective therapeutic outcomes in animal studies.

Small molecule approaches for gene regulation could bypass the need for delivery strategies. A number of natural and synthetic DNA binding molecules have been explored for their ability to regulate gene expression *in vitro* and in cell culture. Our laboratory has explored the development of programmable oligomers for targeting double stranded DNA with affinity and specificity comparable to that of transcription factors.[1-3] These small molecules achieve affinities and specificities of DNA binding proteins, inhibit a broad range of transcription factors, are cell-permeable, bind to chromatin and have been shown to downregulate endogenous gene expression in cell culture

Pairing Rules.

The original inspiration for programmable DNA binding oligomers is drawn from netropsin and distamycin A, which comprise two and three aromatic N-methylpyrrole (**Py**) rings.[4] These crescent-shaped natural products bind A,T tracks with both 1:1 and 2:1 ligand:DNA stoichiometries.[5] Cofacial pairs of unsymmetrical heterocycles distinguish the edges of each Watson-Crick base pair in the minor groove.[4] DNA binding polyamides containing Py, Im, and Hp amino acids form the basis of a modular code to control sequence specificity in a predictable way.[6] (Fig. 1.1) Pairs of pyrrole (**Py**), imidazole (**Im**), and hydroxypyrrole (**Hp**) rings distinguish the four Watson-Crick base pairs. Im/Py and Py/Im distinguishes G•C from C•G, Hp/Py distinguishes T•A from A•T. Each polyamide strand is usually oriented N→C with respect to the 5'→3' direction of the DNA helix.[7] The β-dimethylaminopropylamine (β-Dp) tail (from the method of synthesis) has a DNA sequence preference for A,T base pairs for steric reasons.[6]

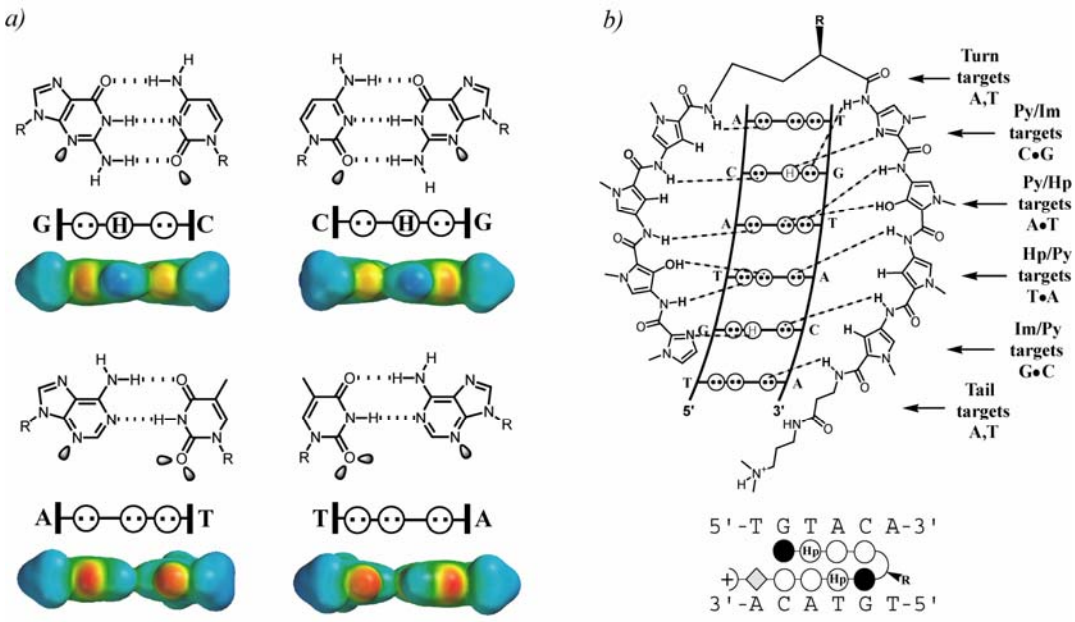

Figure 1.1 (a) Chemical structure of Watson-Crick base pairs. Circle with two dots indicates a hydrogen bond acceptor while a circle with an **H** indicates a hydrogen bond donor. Electronic surfaces of the base pair edges shown with red and blue indicating negative and positive electronic charges, respectively. (b) Chemical structure of a polyamide with the putative hydrogen bonds to the DNA minor groove is shown. Pairing rules for the 5-membered heterocycles shown to the right.

Affinity and Specificity.

Increasing polyamide affinity and specificity was accomplished by linking antiparallel polyamide dimers with a short alkyl chain **γ** to afford a single linear oligomer that folds into a hairpin structure in the minor groove of DNA.[8, 9] The turn unit enforces unambiguous ring pairing, eliminating slipped binding modes.[10, 11] Incorporation of a chiral amine at the α position of the turn unit increases affinity and disfavors reverse binding.[12, 13] Acylation of the chiral amine is effective at further increasing sequence specificity and prevents hairpin oligomers from unfolding and binding the minor groove in a linear fashion.[14] The aliphatic turn unit has a sequence preference for A,T base pairs due to an unfavorable steric clash with the exocyclic NH_2 of G,C base pairs.[14] Polyamide composition with respect to ring-number and type of heterocycles influences

the affinity. Six-ring hairpin polyamides have equilibrium association constants on the order of Ka = 10^7-10^8 M^{-1} while for eight- and ten-ring hairpins, Ka = 10^9-10^{10} M^{-1}.[15] Despite the gain in specificity, polyamides with multiple **Hp/Py** pairs exhibit reduced affinity.[16] The reduction in affinity likely results from modest distortion of the DNA upon binding as well as differences in solvation in water versus the minor groove of DNA.[17, 18]

Sequence dependent variations in the DNA microstructure play a role in the energetics of binding. A number of DNA sequences, such as purine tracts have emerged as "lower affinity" DNA targets. The relative rigidity of purine-purine steps, associated narrow minor groove, and negative propeller twist are thought to be governed by optimal

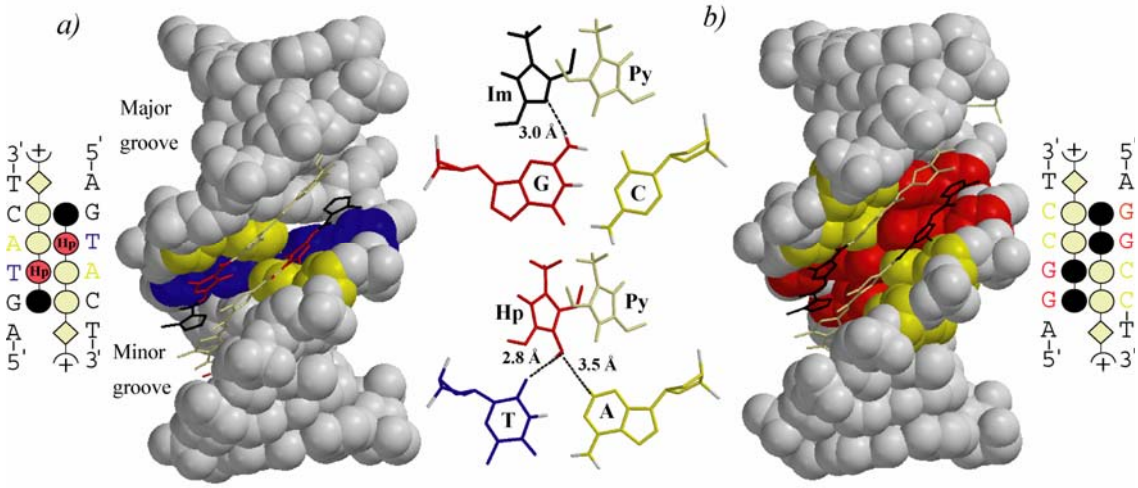

Figure 1.2. Crystal structures demonstrating the origin of G,C and T,A specificity. (a left) Ball and stick model of polyamide homodimer bound to 5'-GTAC-3' core sequence. (a center) Crystal structure of the Im-Hp-Py-Py-β-Dp homodimer bound in a 2:1 complex with the DNA. Adenine and thymine are colored yellow and blue, respectively. Pyrrole (**Py**) and imidazole (**Im**) heterocycles are colored tan and black, respectively. Hydroxypyrrole is colored red. (a right bottom) Structural basis for A-T vs. T-A discrimination by the hydroxypyrrole/pyrrole (**Hp**/**Py**) pair. (b right) Ball and stick model of polyamide homodimer bound to 5'-GGCC-3' core sequence. (b center) Crystal structure of the Im-Im-Py-Py-β-Dp homodimer bound in a 2:1 complex with DNA. Guanine and cytosine are colored red and yellow, respectively. Pyrrole (Py) and imidazole (Im) heterocycles are colored tan and black, respectively. (b left top) Structural basis for G-C vs. C-G discrimination by the imidazole/pyrrole (**Im**/**Py**) pair.

base stacking interactions.[19] Variation in DNA groove width, curvature, bendability, hydration, or relative position of hydrogen bond donors/acceptors all influence the DNA's ability to accommodate these shape selective ligands, and make it difficult to quantify incrementally the source of the affinity variations.[20, 21]

Sequence specific recognition of the DNA minor groove arises from the pairing of two different 5-membered heterocyclic amino acids and the interplay of a variety of direct and indirect recognition elements. The overall shape of the folded hairpin fits the shape (width, curvature, depth) of the DNA minor groove.[20, 21] The information face on the inside of the crescent-shaped oligomer may be programmed by the incremental change of atoms on the corners of the ring pairs presented to the DNA minor groove floor. [20, 21] The corners of the ring pairs read in a digital way (not unlike Braille) each of the different edges (bumps and holes) of the four Watson-Crick base pairs.[22, 23] Stabilizing and destabilizing interactions with the distinct edges of the four Watson-Crick bases are modulated by shape complementarity and specific hydrogen bonds (Figure 1.2). More specifically, the imidazole ring **Im**, which presents a lone pair of electrons via its N(3) to the DNA minor groove, can accept a hydrogen bond from the exocyclic amine of guanine.[22, 23] The 3-hydroxypyrrole ring in the **Hp/Py** pair projects an exocyclic OH group toward one side of the minor groove floor that is sterically accommodated in the asymmetric cleft of the T·A base pair, preferring to lie opposite T not A.[22, 23 24, 25] From x-ray structural analysis, it appears that **Hp** can hydrogen bond with the O(2) of thymine (Figure 1.2). Recognition of DNA by polyamides is also affected by a series of ligand-ligand or ligand-DNA interactions that take place away from the polyamide information face. Polyamide geometry with respect to overall

curvature, ligand preorganization, and van der Waalls contacts between the polyamide and the DNA minor groove walls are factors important in determining polyamide specificity.[26, 27]

Binding Site Size.

Due to the large size of the human genome, short DNA sequences 6-8 bp in size would be expected to occur millions of times in gigabase-sized DNA. Thus, it is useful to maximize the size of the targeted binding site by creating oligomers that are capable of specifically recognizing long stretches of DNA. For example, an 8-ring hairpin binds six bp of DNA. It is intuitively obvious that longer oligomers should bind larger DNA sites. It has been shown that the polyamide ligand does not match the curvature of the DNA helix after five contiguous rings.[28] The polyamide is overcurved with respect to the DNA. One solution is to incorporate the flexible β-alanine linker as a discrete unit in the hairpin.[22, 23] The **β/Py** or **β/Im** pairings function similarly to the **Py**/**Py** and **Py**/**Im** pairings respectively, with the aliphatic C-H of the **β** residue targeting A, T, or C due to steric occlusion of guanine's exocyclic amine. Unlike the rigid contiguous ring system, β-alanine provides more conformational freedom and allows relaxation of polyamide curvature (Figure 1.3) Subsequently, the polyamide is able to reregister itself and track the DNA helix.[29, 30] β-alanine's dual functionality as an effective recognition element and as a flexible extension proves useful in the design of polyamides capable of binding extended 11 to 16 base pair sites.[31]

The incorporation of β-alanine into hairpin polyamides allows for greater conformational freedom. Sequences such as 5'-G**N**G-3' that were originally targeted

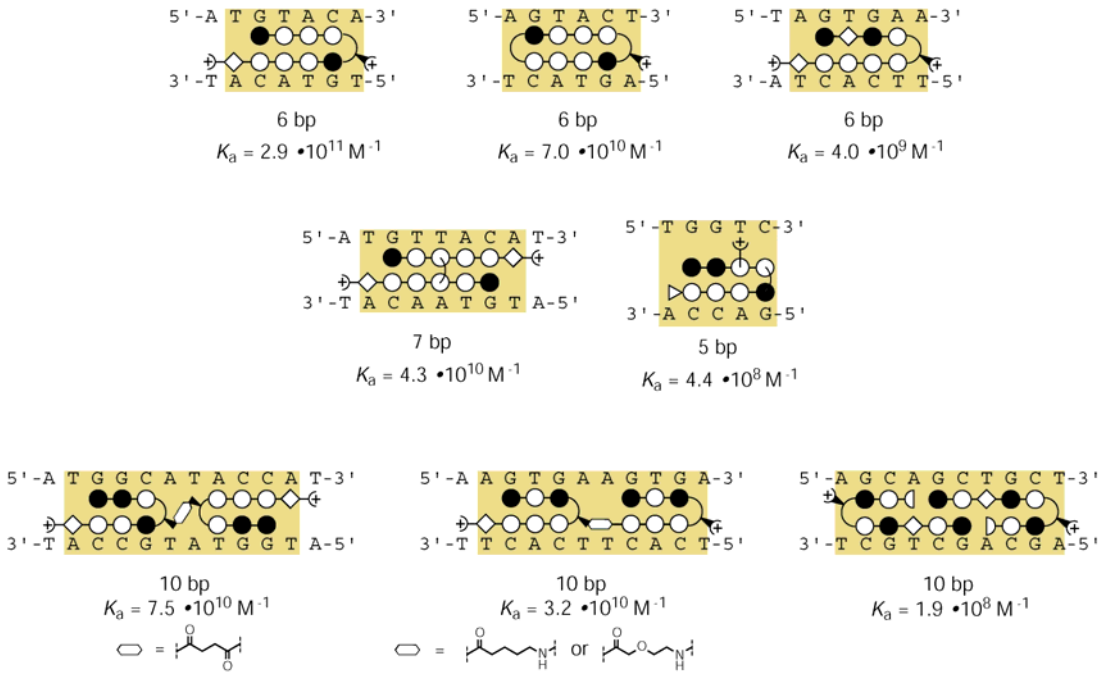

Figure 1.3. Selected polyamide binding motifs and their corresponding binding site size are shown. From left to right and top to bottom: hairpin, cycle, β-substituted, H-pin, U-Pin, turn-to-turn tandem, turn-to-tail tandem, and candy cane overlay.

using rigid contiguous ring systems Im-**X**-Im (**X** = Im, Py) at low affinity (10^8 M^{-1}) can now be coded for at high affinity (10^{10} M^{-1}) by incorporation of a flexible β linker (Im-β-Im). The aliphatic linker relaxes the overall oligomer curvature and allows the alternating imidazole rings to conform to the DNA minor groove microenvironment, maximizing hydrogen bonding with the exocyclic amine of guanine.[29, 32]

Polyamide Binding Motifs.

In addition to the hairpin motif, a series of new oligomer shapes were constructed. They include cycles, extended hairpins, H-pins, U-pins, and hairpin dimers Fig. (**3**).[29, 32] Cycles are an attractive motif due to their symmetry and pre-organization but are difficult to synthesize. H-pins and U-pins are unique in their connectivity.[33, 34] Instead of linking polyamide strands at the turn position, H-pins and U-pins link the

strands via an aliphatic bridge extending from the N-methyl substituent of paired rings. Unlike the γ turn in hairpins, the "bridging" linker is not sequence specific because it does not lie in the minor groove (Figure 1.3).

Transcription factors often bind as homo-and heterodimers to contiguous DNA sites each six bp in size often in a highly cooperative manner. Inspired by biology, we could consider coupling hairpin modules for longer binding site size recognition.[35, 36] Hairpin dimers are linked turn-to-turn or turn-to-tail, occupying a larger 10 base pair binding site and bind DNA with high affinity (Ka = 10^{10}-10^{11} M^{-1}). The benefit of the dimer motifs is the targeting of larger binding sites in comparison to the standard hairpins. There are sequence restrictions. The alkyl linkers that join the two hairpins in the tandem formation specify for A,T base pairs for steric reasons. A concern regarding the dimer motifs is the overall oligomer size, which results in poor cell uptake.

Alternative Heterocycles for DNA Minor Groove Recognition.

The pairing rules have proven useful for the recognition of hundreds of DNA sequences by designed polyamides. In addition, as this chemistry moves from cell culture to small animal models, one could imagine that a *library of heterocycles* for DNA recognition would be important regarding absorption, distribution, metabolism, toxicity and pharmacokinetics. A search was initiated for other novel heterocycle recognition elements that offer improved affinity/specificity and are chemically robust relative to the original Py analogs. Initial efforts were directed towards developing new pairs within the context of 5-membered heterocyclic amino acid pairs.[37, 38] Figure 1.4 shows a family of 5-membered, aromatic, heterocyclic residues grouped in columns by the type of

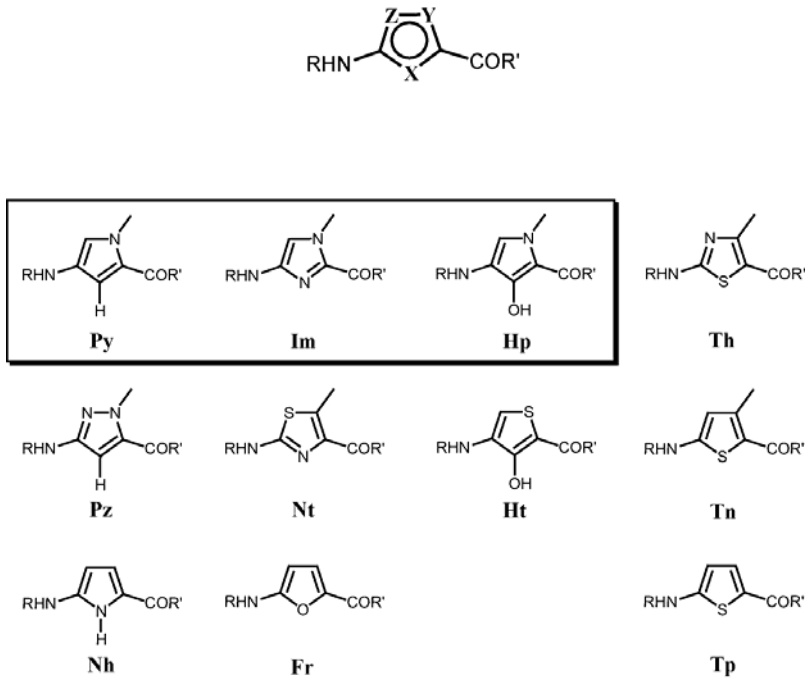

Figure 1.4. Family of heterocyclic amino acids studied here. Centered above is a schematic showing the five-membered heterocyclic framework with the variable positions labeled X, Y, and Z. The parent Im, Py, and Hp residues are boxed. All residues are shown with the functionality that faces the DNA minor groove pointed down (X).

functionality directed toward the floor of the DNA minor groove. **Py**, pyrazole (**Pz**) and 1H-pyrrole (**Nh**) project a hydrogen with positive potential toward the DNA; **Im**, 5-methylthiazole (**Nt**), and furan (**Fr**) project an sp^2 lone pair from nitrogen or oxygen; **Hp** and 3-hydroxythiophene (**Ht**) project a hydroxyl group; and 4-methylthiazole (**Th**) and 4-methylthiophene (**Tn**) project a large, polarizable sulfur atom.

Comparative analysis of new residues within this 5-membered heterocyclic framework enabled us to retain overall ligand shape and to observe the effects of small structural changes, such as single atom substitution, on DNA base pair specificity. Each heterocyclic amino acid has an inherently unique bonding geometry, which results in varying degrees of curvature complementarity between the polyamide and the DNA minor groove (Figure 5). Curvature effects are amplified in contiguous ring polyamides, where continuous π-conjugation limits conformational flexibility. Given that **Py** is over-

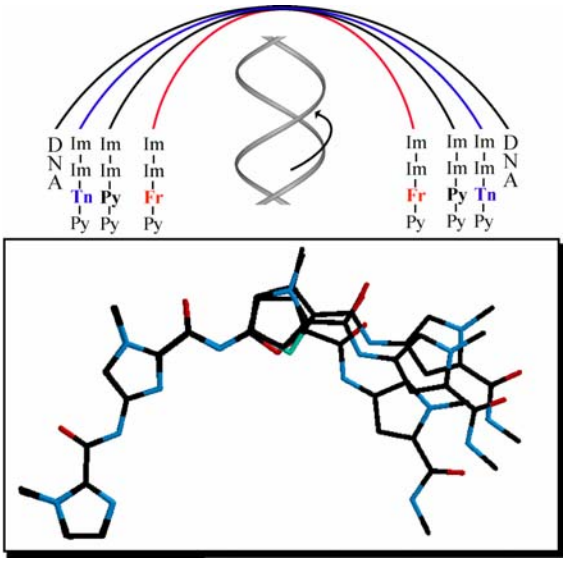

Figure 1.5. (Top) Schematic illustrating the curvatures of 4-ring polyamide subunits containing Tn, Py and Fr heterocycles with respect to one another and the DNA helix. (Box) *Ab inito* models of polyamide subunits (Im-Im-**X**-Py, **X** = Tn, Py and Fr) superimposed to demonstrate the significant difference in curvature resulting from atomic substitution. Hydrogen not shown.

curved with respect to the DNA helix, reducing heterocycle curvature should increase the polyamide-DNA fit. Consequently, the polyamide would have greater sensitivity to changes in DNA structure and therefore greater DNA sequence selectivity, demonstrated by the **Tn**/**Py** pairing.[39] In contrast, increasing heterocycle curvature can decrease sensitivity to changes in DNA sequence or abolish binding completely as in the case of the **Fr**/**Py** and **Ht**/**Py** pairs.[38] These results suggest that considering the functional group facing the minor groove floor is insufficient for an accurate prediction of DNA recognition behavior and that recognition can be impacted by substitution of heterocyclic atoms projecting *away from* the DNA minor groove. Simply, the most conservative atomic substitutions that influence geometry can have a substantial impact on oligomer specificity, indicating how small the acceptable chemical space is for the creation of novel recognition elements.

N-Terminus Pairings.

The limited success at developing novel 5-membered heterocycle recognition elements at the internal position of hairpin polyamides prompted the search for new rings which could impart specificity at the N-terminus (cap) position. Earlier studies had shown that both **Im/Py** and **3-Pz/Py** pairs were capable of selecting for G > C at the cap position with good selectivity and affinity.[40] While **Im/Py** pairings show comparable specificity for G•C at both *terminal* and *internal* positions, and a 3-**Pz/Py** (pyrazole) shows near 100-fold specificity for G > C at the N-terminal position, N-terminal pairings capable of binding T•A with affinity and specificity comparable to those of the G•C specific residues had not been established.[41] Knowing that the specificity of cofacial aromatic amino acid pairings depend on their context (position) within a given hairpin polyamide, a library of heterocyclic carboxylic acids was screened for favorable recognition properties.

Due to the absence of a second "groove-anchoring" carboxamide, rings at the cap position can adopt different conformations and are thought to be allowed to bind DNA in either of two rotamers. For example, a terminal **Hp** residue can exist in two rotamer forms where one orients the hydroxyl group into the minor groove while a second orients the ring with the hydroxyl recognition element oriented away from the floor of the minor groove. This second orientation could be stabilized by intramolecular hydrogen bonding between the C3-OH and the carbonyl oxygen of the 2-carboxamide. For terminal 2-hydroxybenzamide residues, some measure of T•A selectivity can be recovered by creating steric bulk at the 6-position to force the hydroxyl recognition element into the groove.[41]

The fidelity of minor groove recognition by N-terminal **Im/Py** pairs suggests that the rotamer that projects N(3) into the groove is the preferred orientation in hairpin polyamides. This observation can be rationalized by a combination of both stabilizing and destabilizing forces that favor the rotamer with N(3) in the groove and N-methyl out. Rotation of a terminal **Im** residue in the opposite conformer, orienting N(3) away from the minor groove, would create unfavorable lone pair interactions with the proximal carboxamide oxygen, disrupt a favorable hydrogen bond with the exocyclic amine of G, and project an N-methyl group to the DNA floor that is sterically unfavorable.

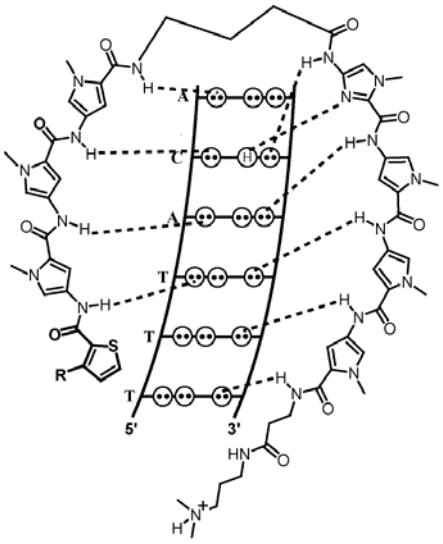

Figure 1.6. Proposed binding models for hairpin polyamides with 5'-T**X**TACA-3' site. A circle enclosing two dots represents lone pairs of N3 of purines and O2 of pyrimidines. A circle containing an H represents the exocyclic amine of guanine. Putative hydrogen bonds are indicated by dashed lines. N-terminal residue drawn in "sulfur down" syn conformation.

Using the modest 3-fold specificity of the **Tn/Py** pair for T•A as a starting point, a library of thiophene rings derivatized at the 3-position was designed to impart a shape selective mode of recognition for thymine. From a series of thiophene caps including 3-H, -CH_3, -NH_2, -NHAc, -OH, -OCH_3, -F and –Cl, it was found that N-terminal 3-methoxy (**Mt**) and 3-chloro (**Ct**) thiophene-2-carboxamide residues, when paired with **Py**, demonstrate selectivity for T•A versus A•T (Figure 6).[42] Three- and 4-ring polyamides containing a variety of heterocycles at the N- and C-terminus have demonstrated good *in vitro* potency against Gram-positive bacteria.[43]

Fused Bicycles – Benzimidazole Analogues.

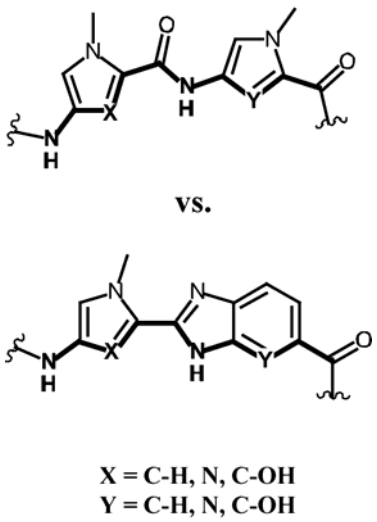

Figure 1.7. Structures of the fused 6-5 bicyclic benzimidazole building blocks in comparison with their respective five-membered ring systems. Hydrogen bonding surfaces presented to the DNA-minor-groove are bolded.

A movement away from the classic 5-membered heterocyclic carboxamides led to the incorporation of benzimidazole analogues as effective DNA recognition tools.[44, 45] The benzimidazole 6-5 bicyclic-ring structure, while having different curvature from the 5-membered heterocyclic carboxamides, presents an "inside edge" with a similar readout and shape to the DNA minor groove floor, effectively mimicking **Py**, **Im**, and **Hp** heterocycles (Figure 1.7). A series of experiments showed that the benzimidazole (**Bi**), imidazopyridine (**Ip**), and hydroxybenzimidazole (**Hz**) analogues, when placed into 8-ring hairpins as **Bi**/**Py**, **Ip**/**Py**, and **Hz**/**Py** pairs, are as effective at recognizing the DNA minor groove as their 5-membered counterparts (Figure 1.8).[46, 47]

Of particular interest is the **Hz**/**Py** pairing. Designed to be a hydroxypyrrole mimic, the **Hz**/**Py** pair places the same direct readout functionality to the floor of the DNA minor groove as the **Hp**/**Py** pair. Like the **Hp**/**Py** pair, the **Hz**/**Py** pair is capable of effectively discriminating between A,T Watson-Crick base pairs such that **Hz**/**Py** codes for T·A and **Py**/**Hz** codes for A·T (Table 1.2). The **Hz**/**Py** pair demonstrates an increase in binding affinity for its match sites and is comparable in specificity to the **Hp**/**Py** pairing. The **Hz**/**Py** pair also discriminates more effectively against G,C base pairs than

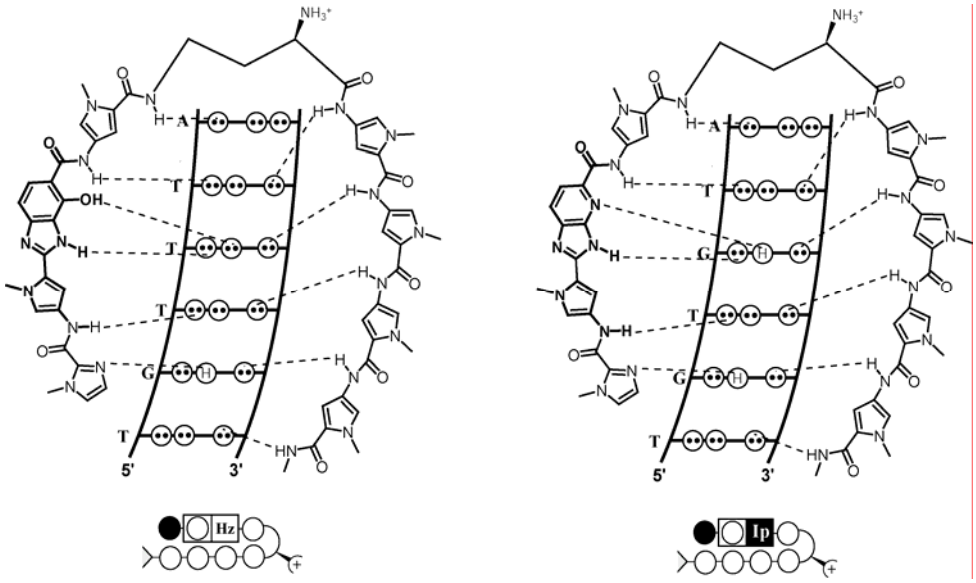

Figure 1.8. Proposed binding models for hairpin polyamides containing a hydoxybenzimidazole (**Hz**) ring system (Left) and a imidazopyridine (**Ip**) ring system (Right). A circle enclosing two dots represents lone pairs of N3 of purines and O2 of pyrimidines. A circle containing an H represents the exocyclic amine of guanine. Putative hydrogen bonds are indicated by dashed lines.

Hp/**Py**. While the "recognition edges" of **Hz** and **Hp** are the same, there are significant differences in overall ligand geometry and electronics. As is the case with all of the 6-5 ring systems **Bi**, **Ip**, and **Hz**, a higher degree of rigidity and pre-organization of the fused hydroxybenzimidazole structure, coupled with a lower degree in curvature that may be more complementary to the inherent curvature of the DNA helix, likely play roles in the increased affinity and specificity. Further, the benzimidazole moiety has a greater aromatic surface and hydrophobicity that may alter both the DNA-ligand van der Waals interactions, and the inter-strand π-stacking. Thus, by going from the 5-membered heterocyclic system to the fused 6-5 system, changes associated with the indirect readout of the DNA minor groove may be responsible for the evident changes in recognition.

β-Linked Polyamides: A Special Case for DNA Recognition of $(GAA)_n$ Tracts.

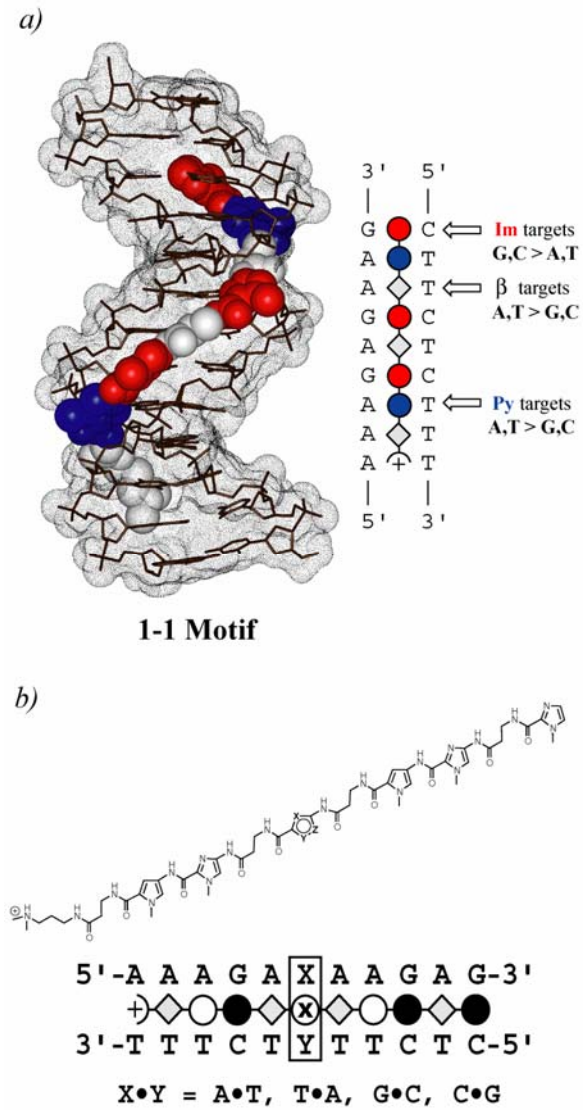

Figure 1.9. 1:1 Polyamide:DNA motif. (a) (Left) High resolution 1:1 polyamide:DNA complex determined by NMR. (a) (Right) Ball and stick model of polyamide:DNA complex. Imidazole (**Im**), Pyrrole (**Py**), and β-alanine (**β**) are colored red, blue, and white respectively. (b) Chemical structure of the polyamide system used for examining the recognition profile of novel 5-membered heterocycles in a 1:1 polyamide:DNA complex.

Although the energetics and structure of the 2:1 complex has been explored extensively, there is less understood about 1:1 polyamide recognition beyond the initial studies on netropsin and distamycin.[48] Laemmli and coworkers reported that certain β-linked **Py**/**Im** polyamides bind GAGAA tracks in a 1:1 stoichiometry with a single orientation.[49] Previously, purine tracts have been difficult sequences to target using 2:1 binding hairpin polyamides. The 1:1 complex is important for expanding the sequence repertoire for DNA targeting, but the fact that β-linked **Py**/**Im** polyamides can bind *both* 1:1 and 2:1 in the minor groove raises important design issues for the field.[50] In an effort to further characterize the 1:1 mode of binding, specificity studies were conducted to

determine if β-linked polyamides in 1:1 stoichiometry complexes can discriminate any of the four Watson-Crick base pairs.[51] For 1:1 recognition we find that **Py** and **β** target A,T > G,C and **Im** targets G,C > A,T.

The structure of β-linked polyamides binding GAA purine tracts was studied by high resolution NMR (Figure 1.9).[52] The complex reveals B-form DNA with a narrow minor groove and a large degree of negative propeller twist. Stabilization of the negative propeller twist by bifurcated hydrogen bonds donated from each polyamide NH group to proximal purine N(3) and pyrimidine O(2) atoms, in addition to the inherently rigid and narrow minor groove, is thought to be the reason polyamides bind 1:1 in polypurine sequences. Second, there is a G/C dependent orientation such that the polyamide is oriented N-C with respect to the 3'-5' direction of the guanine-containing strand. Finally, the ensemble reveals specific hydrogen bonds between Im-N(3) and G-NH2 that could only be made due to the flexibility imparted by the β residue.

Synthetic Methods.

The synthesis of polyamides is traditionally accomplished by the stepwise addition of Boc-protected amino acid monomers and dimers to either the Boc-β-Ala PAM resin or the Kaiser oxime resin (Figure 1.10).[53] The benzyl ester linkage, which binds the growing oligomer to the PAM resin, has been found to be stable to the established polyamide coupling and Boc-deprotection conditions. Monomers and dimers are coupled onto a growing polyamide chain by deprotection of the solidphase bound amine with TFA, followed by addition of an activated ring with either DCC/HOBt or HBTU. Deprotection and addition steps are repeated in a stepwise fashion until the

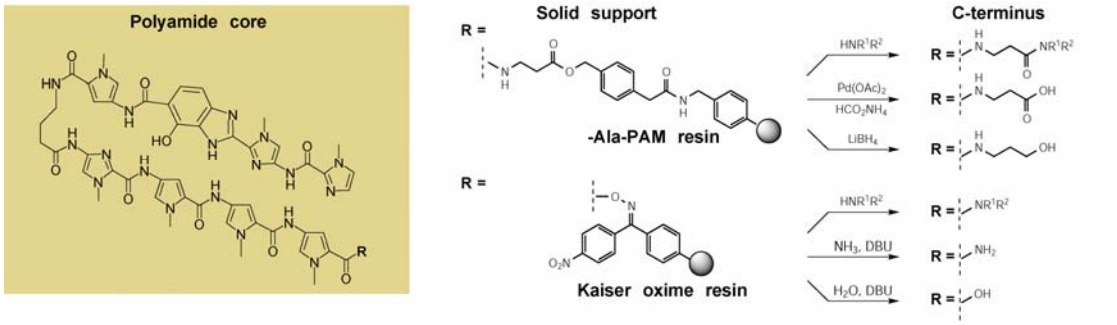

Figure 1.10. (Left) Oligomer synthesized using solid phase methodology. (Center) Resin systems used for polyamide synthesis. (Right) Cleavage conditions for modulating the chemical functionality at the C-terminus of a polyamide.

oligomer synthesis is complete, at which point the final product is cleaved off the resin by a nucleophile of choice.

Though use of the PAM resin allows for rapid preparation of a range of polyamides, it installs a T,A-selective β-alanine residue at the C-terminus, which places limits on the DNA sites that can be targeted. To address this problem, the Kaiser oxime resin was adapted to polyamide synthesis, allowing the preparation of polyamides with incrementally shorter C-termini.[54] Tails as short as methyl amide can be obtained. Molecules lacking the C-terminal β-alanine residue display the desired tolerance for G,C bases while maintaining high affinities. Moreover, removing the β residue has proven to be instrumental in nuclear uptake of polyamides in live cells. One caveat of the oxime resin is the significantly weaker aromatic oxime linker when compared to the beta-alanine benzyl ester used in PAM resin. High temperature monomer couplings needed to couple less reactive monomers – such as the less nucleophilic **Im** amine – and high concentrations of TFA for Boc deprotections tend to cleave the oligomer/resin bond, resulting in poor overall yields. In order to circumvent these shortcomings, lower concentrations of TFA are used for the Boc deprotection of solid-phase bound amines.

Also, solution-prepared dimers and trimers are used to ensure that activated monomers and dimers are coupled only to reactive resin-bound amines.

Pessi and co-workers have recently used a sulfonamide-based safety-catch resin to prepare derivatives of hairpin polyamides. Unlike other resins that have been employed, activation of the safety-catch resin linker is necessary before nucleophilic cleavage of the compound can be accomplished. Pessi demonstrated that resin-bound polyamides were readily cleaved with stoichiometric quantities of nucleophile providing a practical route to thioesters or polyamide-peptide conjugates.[55]

Although Boc-peptide chemistry is generally employed in solid phase polyamide synthesis, it should also be noted that that Fmoc chemistry has been effectively used with suitably protected monomers and Fmoc-β-Ala-Wang resin.[56]

Nucleosomes.

Biological DNA in the nucleus of a cell is condensed one-million fold and compacted in chromatin. The nucleosome core particle (NCP), the elemental repeating unit of all eukaryotic chromatin, consists of 146 base pairs of DNA wrapped around two copies each of four histone proteins. Structural differences between DNA on chromatin and B-form DNA coupled with the diminished accessibility of the minor groove could adversely affect the ability of polyamides to bind their designed sites, prompting much research in that area.[57] By using six different hairpin polyamides, however, it was shown that sites on nucleosomal DNA facing away from the histone octamer, or even partially facing the octamer, are fully accessible.[58] In fact, one section of 14 consecutive base pairs (which constitutes more than one full turn of the DNA helix) were accessible to

high-affinity polyamide binding. It was noted that polyamides whose binding sites were near the N-terminal tails of histone H3 and histone H4 were unable to effectively bind. In order to probe whether or not the histone tails were responsible for the diminished binding ability of polyamides, each tail was removed, allowing polyamides access to their designed sites.[58]

The structures of three of the polyamide-NCP complexes were determined by x-ray crystallography (Figure 1.11) and it was observed that while the histone octamer is unaffected by polyamide binding, the nucleosomal DNA undergoes significant structural changes at the polyamide binding sites and at adjacent regions.[59] The effect of polyamides on the dynamic behavior of the nucleosome in solution was also probed by testing their effect on the repositioning of the histone octamer, induced by heating. It has been established that after salt-gradient dialysis, the histone octamer shifts to an off-centered position on the 146 base-pair satellite DNA. After heating reconstituted NCPs for two hours at 37 ^{o}C, the octamer transfers to a more thermodynamically centered position. Experiments have shown that the binding of one of these polyamides (ImPyPyPy-γ-PyPyPyPy-β-Dp) to the NCP completely inhibits heat-induced repositioning thus implicating the ligands in modulating the structure of the nucleosome.[58] The exact

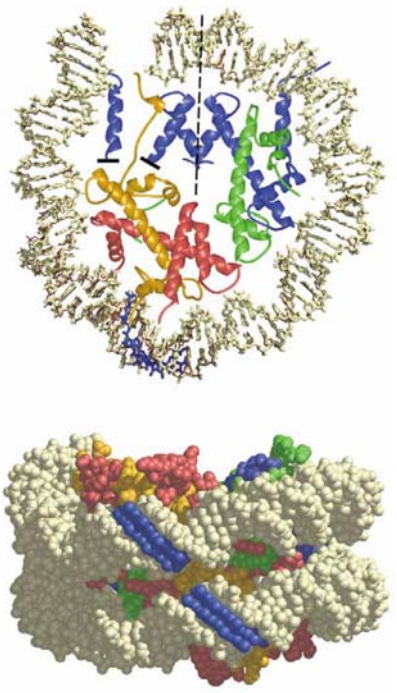

Figure 1.11. High resolution crystal structure of polyamides bound to the DNA minor groove of the nucleosome core particle (NCP).

mechanism of nucleosome repositioning is not yet fully understood, but it appears that, in some cases, DNA can "roll" over the histones. It is reasonable to assume that certain polyamides, when bound to their target sites, can act as chocks that prevent the DNA from moving.[59, 60]

To investigate potential effects on transcription through a nucleosome, hairpin polyamides were targeted to sites on the nucleosome positioning sequence of the sea urchin 5S gene.[60] The two molecules that blocked heat-induced nucleosomal translocation also blocked transcription by T7 RNA polymerase. Each of these polyamides binds with high affinity (Kd ~ 1 nM) to a single site in the nucleosome construct, potentially implying that placement is critical. Nonetheless, the positions of these sites are distinct from those occupied by the compound in the crystal structure. These polyamides prevent repositioning of the histone octamer by RNA polymerase, in turn inhibiting passage of the elongating polymerase through nucleosomal DNA.[60] These results establish with good certainty that the NCP is not only accessible to polyamides, but that regulation of downstream events is possible.

Cell and Nuclear Uptake.

While DNA binding polyamides have been shown to inhibit and influence a wide variety of protein-DNA interactions in solution, in order to see similar effects in cell culture, access to the nucleus is critical. Nuclear uptake of hairpin polyamides has proven to be dependent on cell type. In order to visualize the localization of polyamides in live cells, a series of fluorescently labeled polyamides has been prepared to analyze the intracellular distribution of these molecules in a panel of cell lines. It has been

shown that polyamide–Bodipy conjugates stain the nuclei of T lymphocytes, a cell type that has shown robust responses to polyamides *in vivo*, but to no other cell type tested.[61] In fact, polyamide-Bodipy conjugates most commonly produce a punctuate cytoplasmic staining pattern with no appreciable levels of nuclear staining. Other studies have shown that polyamide–fluorescein conjugates can uniformly exhibit favorable nuclear uptake properties in several human cancer cell lines.[62] Presence of a C-terminal β-alanine residue, a feature of polyamides synthesized on the PAM resin, seems to be a negative determinant for nuclear uptake. Polyamide–FITC conjugates, which only differ in whether or not they contain a β-alanine at the C-terminus, tend to exhibit different uptake properties with des-β compounds showing improved uptake properties.[62] It has also been observed that the **Im** content and location within a polyamide can affect the level of nuclear localization. A clear trend, however, has yet to be determined. What has been concretely established is that manipulation of the linking residues, ring content, and choice of dye, can generate compounds that display nuclear localization of polyamide-dye conjugates across a broad range of mammalian cells.[62]

Modulation of Gene Expression.

Diversity of cellular function within an organism is not dependent upon the genomic information contained within the cells but instead on the choice of which genes, and at what frequency, are transcribed. With the sequencing of the human genome complete, the stage is set to explore how manipulation of individual genes determines cell fate. In general, small molecule antigene therapeutics such as polyamides can be grouped into two different categories based on their function: transcriptional repression

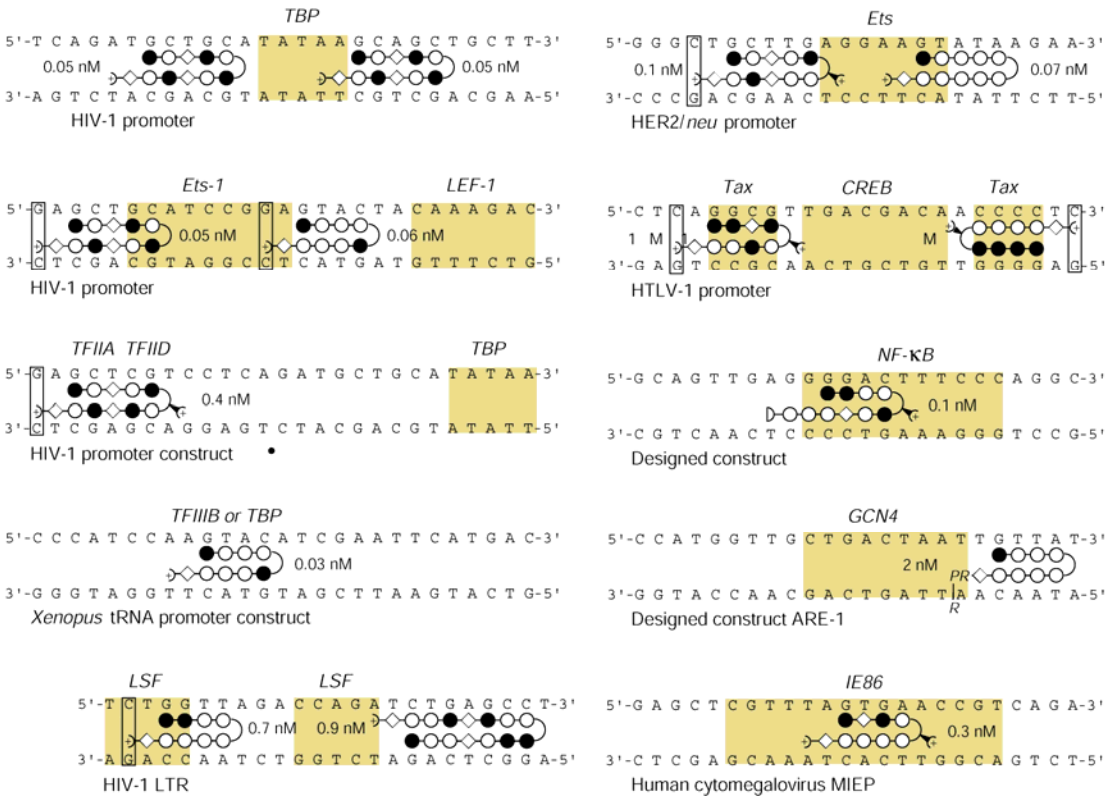

Figure 1.12. Binding sites for an array of important transcriptional regulatory proteins and the polyamides that modulate their binding are depicted.

or transcriptional activation. Repression involves downregulation of a gene by inhibiting the assembly of the necessary transcriptional machinery on the DNA. Small molecules could inhibit repressor proteins resulting in gene upregulation by derepression. Gene activation involves designing molecules with two separable modular domains. DNA binding oligomers with an activation domain (AD) recruit transcriptional machinery to the promoter of specific genes (Figures 1.12 & 1.13).

Repression.

One of the experiments that set the stage for later work in gene inhibition involved the selective inhibition of RNA pol III transcription of a 5S RNA gene in

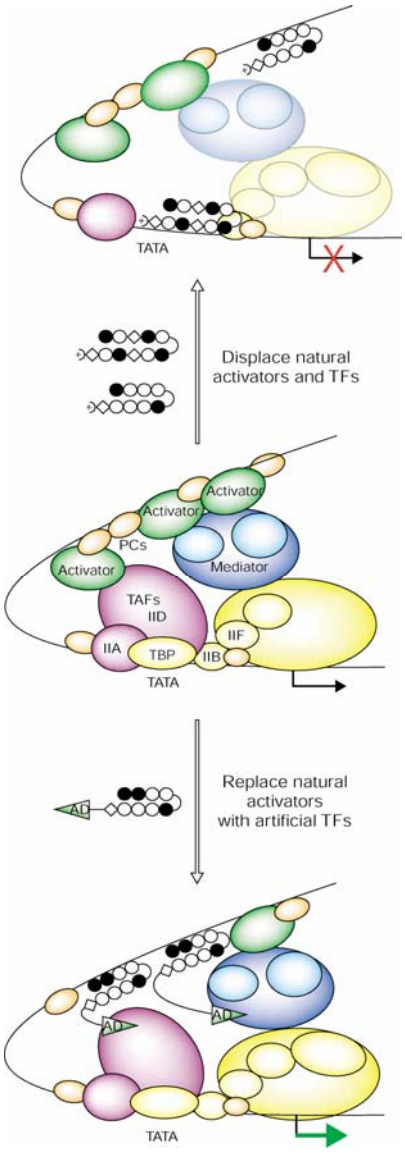

Figure 1.13. (Center) Assembly of transcription factor machinery prior to transcription is shown. (Top) Addition of polyamides that bind specific sequences in at the promoter site can disrupt the assembly of transcriptional machinery and block transcription. (Bottom) polyamides attached with small molecules that are specific for the recruitment of activator or mediator complexes can function as an artificial transcription factor (ATF) by promoting transcription.

Xenopus kidney cells.[63] *In vitro* experiments determined that sequence specific polyamides are capable of both preventing TFIIIA from binding its target sequence as well as disrupting the pre-formed TFIIIA:DNA complex. Nuclear transcription experiments were also able to demonstrate that polyamides are effective *in vivo*, meaning that the oligomers were able to cross the cell membrane, traffic to the nucleus, and disrupt the transcriptional complex on the chromosomal 5S RNA genes. A zinc finger protein, TFIIIA is known to have both major and minor groove DNA contacts. Further work was done to elucidate the nature of the polyamide:TFIIIA interaction and it was determined that the oligomers were able to inhibit TFIIIA binding by blocking one of the transcription factors' minor groove contacts.[64] Work done with TFIIIB, which contains a TATA box binding protein (TBP), was also conducted and demonstrated that disruption of key TBP minor groove contacts resulted in

transcription inhibition.[65]

Earlier work with the possibility for therapeutic applications involved the use of polyamides to inhibit RNA pol II transcription. Multiple DNA-binding transcription factors are required by HIV-1 for RNA synthesis. Design of sequence specific oligomers that target binding sites adjacent to the transcription factors Ets-1, Lef-1 and TBP proved successful in inhibiting viral replication by > 99% in isolated human peripheral blood lymphocytes with no observable cell toxicity.[66] Additional studies were conducted on the HIV-1 promoter to examine the efficacy of polyamide binding fixed distances proximal and distal of the TBP site. The experiments showed that transcription inhibition can be achieved by targeting polyamides to promoter sequences distant from the TATA element.[67] Site-specific alkylation, as a means of irreversible gene regulation, along the HIV-1 promoter was also examined using polyamide-chlorambucil (CHL) conjugates (Figure 1.14a).[68] Site-specific alkylation with polyamide-alkylator conjugates has also been reported by Sugiyama and coworkers.[69]

The success of polyamides at inhibiting binding of Ets-1 led to further studies examining, the cooperative relationship between Ets-1 and NF-κB, and Ets-1 as a regulatory element in transcription of the HER2/*neu* oncogene. In the previous case, Ets-1 and NF-κB binding to the DNA was shown to be cooperative with the ternary complex subject to inhibition by polyamides.[70] NF-κB regulation is clinically important due to its role in several disease-causing viruses to regulate viral gene expression, including HIV-1, HTLV-1, HSV, and CMV. Further, NF-κB plays a role in mediating inflammatory response in humans. Inhibition of the NF-κB:DNA complex by polyamide oligomers *in vitro* has recently been accomplished.[71] In the latter case, polyamides were successful

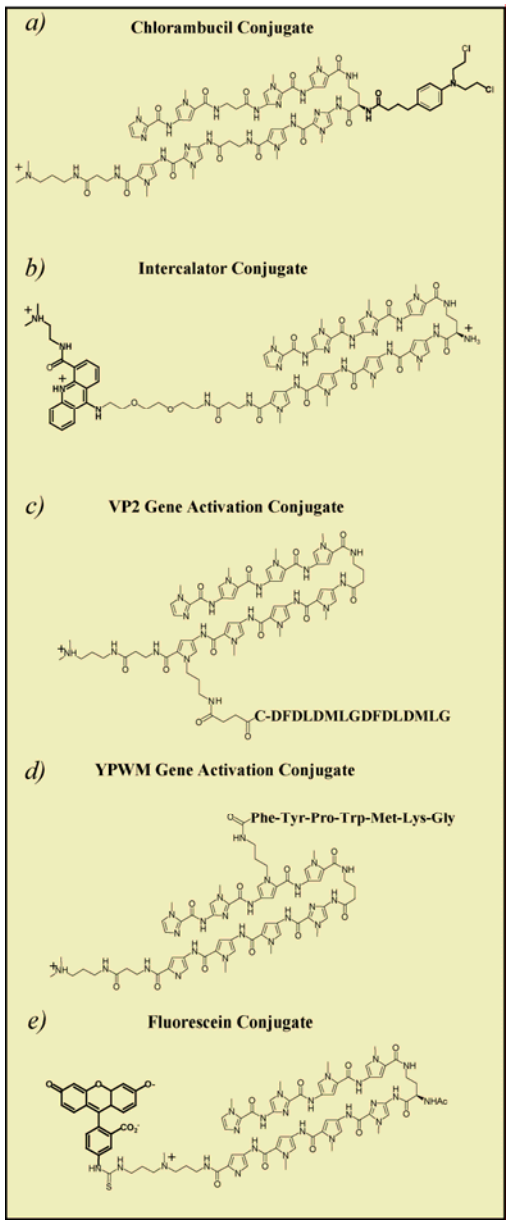

Figure 1.14. A variety of polyamide conjugates. (a) Polyamides functionalized with a chlorambucil (CHL) alkylator group. (b) Polyamides functionalized with a dye intercalator. (c) Polyamides functionalized with a peptidic gene activation domain. (d) Polyamides functionalized with a YPWM activation sequence. (e) Polyamide functionalized with a fluorescein label for visualization and uptake studies.

inhibitors of HER2/*neu* transcription *in vitro*.[72] Other important targets for regulation of viral replication include viral cyclic AMP response elements (CREs).[73]

Thus far, the inhibition of proteins that bind DNA in the minor groove have been discussed. A significant improvement in the functional utility of polyamides as candidates for the regulation of gene expression was reached when the inhibition of major groove binding bZIP protein, GCN4, was accomplished. GCN4 inhibition was accomplished by multiple polyamide designs. The first method employed polyamide Arg-Pro-Arg conjugates that inhibited GCN4 by a mixture of steric occlusion, disruption of phosphate contacts, and charge neutralization of the major groove.[74] Later, this disruptive interaction was refined using polyamides conjugated to a positively charged alkylamino side chain. These “positive

patch" polyamides were 10-fold more potent than their Arg-Pro-Arg counterparts.[75] A second method, which was more universal, involved the development of sequence specific polyamide intercalator conjugates (Figure 1.14b). The polyamide-intercalator conjugates were able to unwind and distort the DNA, effectively inhibiting GCN4 binding at low concentration (IC_{50} ~ 10nM).[76] The sequence specific conjugates offer a novel and powerful approach for disrupting protein:DNA complexes. Further examples of allosteric inhibition of major groove-binding zinc finger proteins by polyamides has also been reported.[77]

Regulation of Endogenous Genes in Cell Culture.

Vascular endothelial growth factor (VEGF) and its receptors have been implicated as key components in tumor angiogenesis that are up-regulated under hypoxic conditions. Polyamides have been shown to repress the hypoxia-inducible transcription of VEGF by selectively binding the hypoxia response element (HRE) DNA sequence, and disrupting the binding of hypoxia-inducible factor (HIF) to its target HRE site (Figure 1.15).[78] Disruption of HIF binding resulted in reduction of VEGF mRNA and secreted protein levels in cultured HeLa cells. Furthermore, the observed effects were polyamide-specific and dose-dependent. Microarray analysis demonstrated a number of hyoxia-inducible genes to be downregulated, providing polyamides as a new approach for targeting angiogenesis.

A small library of polyamide-DNA alkylator conjugates were screened for effects on morphology and growth characteristics of human colon carcinoma cells.[79] Chlorambucil conjugates (Figure 1.14a) arrested neoplastic cells in the G2/M stage of the

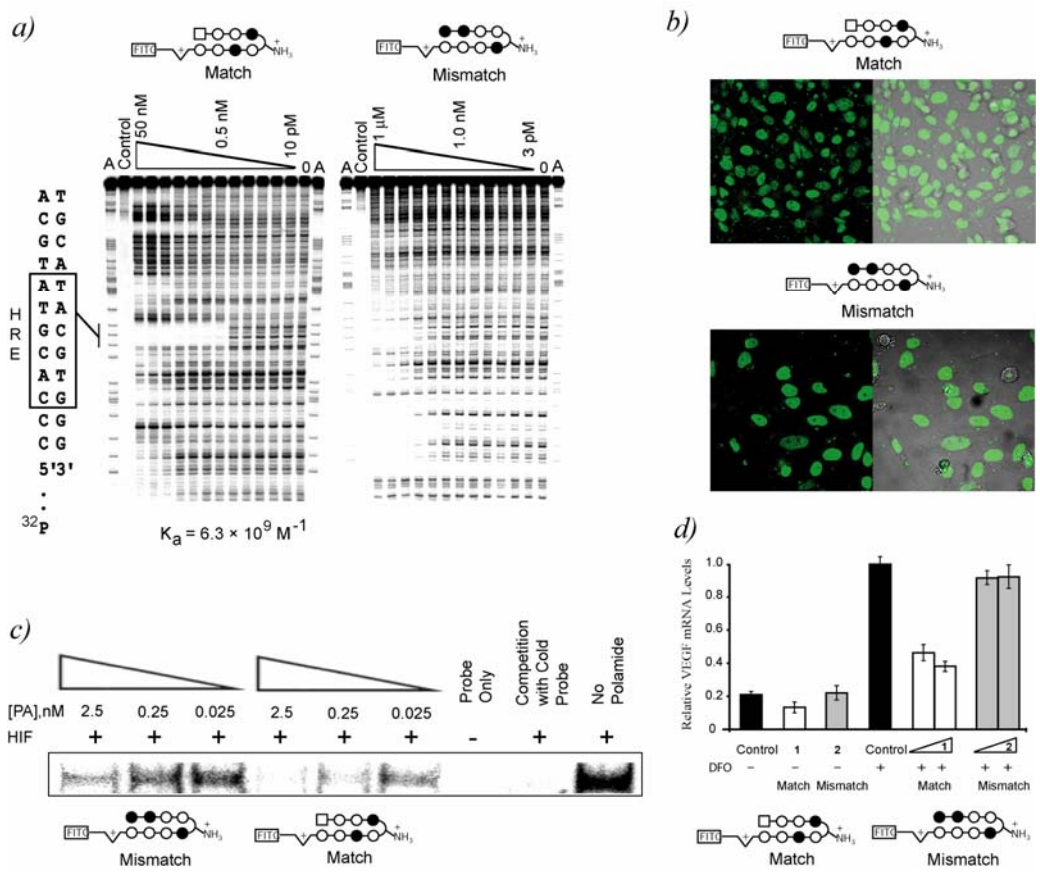

Figure 1.15. (a) DNase I footprinting titration showing fluorescein labeled match polyamide binds to the HRE site with high affinity and specificity. The boxed sequence (left) represents the HRE site. (b) Cellular localization of match and mismatch polyamides in HeLa cells with fluorescent (left) and visible light (right) fields shown. (c) Storage phosphor autoradiogram from EMSA experiment demonstrating the match polyamides ability to inhibit the binding of HIF-1α/ARNT heterodimer to the HRE site. (d) Match polyamide blocks VEGF induction by hypoxia. The relative mRNA levels of expression of the VEGF gene as measured by real-time quantitative RT-PCR are shown.

cell cycle. Microarray, RT-PCR and Western blotting analysis demonstrate down-regulation of the histone H4c gene. LM-PCR confirmed sequence specific alkylation at the H4c gene-coding region. Of particular interest, the polyamides demonstrated activity in a wide range of cancer cell lines, and prevented tumor formation in nude mice with no obvious toxicity.

Activation.

In addition to using polyamides for downregulation of gene expression, recent experiments have focused on inhibiting repressor proteins and, subsequently, gene activation. The first case involves inhibition of repressor protein IE86, which is responsible for repression of human cytomegalovirus (CMV) major immediate early promoter (MIEP).[80] The second case uses polyamides to prevent host factor LSF, a repressor protein, from binding HIV-1 long terminal repeat (LTR) sequences, and subsequently decreases recruitment of histone deacetylase (HDAC1), a protein that is deemed responsible for viral latency.[81] Identifying latent virus has considerable therapeutic applications. Furthermore, LSF is a host protein, making it less susceptible to the frequent mutations that impart HIV-1 with multi-drug resistance.

A more complex problem is the direct recruitment and assembly of the transcription machinery by small molecules or artificial transcription factors (ATFs).[82] In general, transcriptional activators minimally comprise at least two domains, a DNA binding domain and a separate activation domain. In many cases, proteins also possess a dimerization module. In an early attempt to develop a general approach for transcriptional activation, polyamides linked with activating peptide AH (amphipathic helix) PEFPGIELQEQELQALLQQ were shown to activate transcription in yeast nuclear extracts by 13-fold over basal levels.[83] Of note, gene activation was found to be polyamide dependent. More specifically, transcription was not upregulated unless the polyamide targeted its DNA match site upstream of the promoter. Polyamides without an activation domain also had no effect on activation. Further work towards a “minimal motif” reduced the linker length between the polyamide and activation domain from 36

to 8 atoms. The AH signal peptide was also abridged by incorporation of a shorter 16 residue domain taken from the viral potent activator VP16. The polyamide-VP16 conjugate activated transcription over two-fold better than the previous AH conjugate (Figure 1.14c).[84] Replacement of the flexible linker connecting the activation domain to the polyamide with a rigid polyproline linker of fixed lengths (6, 9, 12, and 15 L-proline residues) was used as a "molecular ruler" to determine the optimal distance between the DNA and assembling transcriptional machinery.[85] As each Pro residue measures approximately 3 Å, it was determined that maximal activation occurs as the linker length approaches 12 Pro residues (36 Å).

The most recent development using polyamides as ATFs has been the successful recruitment of homeobox (Hox) protein binding partner extradenticle (Exd) to a cognate DNA site. Hox proteins are known to play a critical role in developmental embryology.[86] By themselves, Hox proteins display poor affinity and sequence specificity for DNA *in vitro*. Hox proteins often interact with the TALE (Three Amino acid Loop Extension) class of proteins and together bind as heterodimers to target DNA with high affinity and fidelity. Recently, x-Ray data have shown that the Hox protein UBX, in conjunction with a TALE domain of sequence YPWM, recruits protein extradenticle (Exd). Polyamides conjugated with the YPWM peptide are able to effectively mimic UBX and improve the affinity of Exd for its cognate site by at least 200-fold, as well as enhancing specific binding of Exd to a target site (Figure 1.14d).[87] The ability of the polyamide-YPWM conjugate to recruit Exd more effectively than its natural Hox protein binding partner (UBX) demonstrates the efficacy of structure-based design for the development of new artificial transcription factors.

Future Directions.

Inspired by the architecture of the natural products netropsin and distamycin A, a new class of programmable sequence-specific DNA-binding oligomers has been invented. Whether these synthetic molecules will allow external control of endogenous gene expression in living systems in a predictable manner, requires the successful integration of man-made chemistry with the complexity of living biological systems. Can we understand how these small molecules access DNA in the chromatin context and interact with the transcriptional machinery in the nucleus to reprogram gene expression? Currently, our research efforts are aimed at a context broader than medicinal chemistry. Perhaps our findings will contribute to the basic biology of gene regulation in eukaryotic cells, probe chromatin and promoter accessibility in the nucleus, and guide the way toward new thinking regarding the potential therapeutic value of small molecule regulation of gene expression.[87]

References:

[1] Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2001,** 9, (9), 2215-2235.

[2] Dervan, P. B.; Burli, R. W., *Current Opinion in Chemical Biology* **1999,** 3, (6), 688-693.

[3] Dervan, P. B.; Edelson, B. S., *Current Opinion in Structural Biology* **2003,** 13, (3), 284-299.

[4] Arcamone, F.; Nicolell.V; Penco, S.; Orezzi, P.; Pirelli, A., *Nature* **1964,** 203, (494), 1064-&.

[5] Pelton, J. G.; Wemmer, D. E., *Proceedings of the National Academy of Sciences of the United States of America* **1989,** 86, (15), 5723-5727.

[6] Swalley, S. E.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1999,** 121, (6), 1113-1120.

[7] White, S.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1997,** 119, (38), 8756-8765.

[8] Wade, W. S.; Mrksich, M.; Dervan, P. B., *Journal of the American Chemical Society* **1992,** 114, (23), 8783-8794.

[9] Mrksich, M.; Parks, M. E.; Dervan, P. B., *Journal of the American Chemical Society* **1994,** 116, (18), 7983-7988.

[10] Baliga, R.; Baird, E. E.; Herman, D. M.; Melander, C.; Dervan, P. B.; Crothers, D. M., *Biochemistry* **2001,** 40, (1), 3-8.

[11] Pilch, D. S.; Poklar, N.; Baird, E. E.; Dervan, P. B.; Breslauer, K. J., *Biochemistry* **1999,** 38, (7), 2143-2151.

[12] Herman, D. M.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1998,** 120, (7), 1382-1391.

[13] Rucker, V. C.; Melander, C.; Dervan, P. B., *Helvetica Chimica Acta* **2003,** 86, (6), 1839-1851.

[14] Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L., *Journal of the American Chemical Society* **2002,** 124, (10), 2148-2152.

[15] Turner, J. M.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1997,** 119, (33), 7636-7644.

[16] White, S.; Turner, J. M.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1999,** 121, (1), 260-261.

[17] Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R., *Journal of the American Chemical Society* **2001,** 123, (21), 5044-5049.

[18] Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Journal of Molecular Biology* **2000,** 295, (3), 557-567.

[19] Hunter, C. A., *Journal of Molecular Biology* **1993,** 230, (3), 1025-1054.

[20] Wemmer, D. E., *Annual Review of Biophysics and Biomolecular Structure* **2000,** 29, 439-461.

[21] Goodsell, D. S.; Kopka, M. L.; Cascio, D.; Dickerson, R. E., *Proceedings of the National Academy of Sciences of the United States of America* **1993,** 90, (7), 2930-2934.

[22] Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C., *Nature Structural Biology* **1998,** 5, (2), 104-109.

[23] Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E., *Science* **1994,** 266, (5185), 646-650.

[24] Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Science* **1998,** 282, (5386), 111-115.

[25] Urbach, A. R.; Szewczyk, J. W.; White, S.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1999,** 121, (50), 11621-11629.

[26] Kool, E. T., *Chemical Reviews* **1997,** 97, (5), 1473-1487.

[27] Bremer, R. E.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2000,** 8, (8), 1947-1955.

[28] Kelly, J. J.; Baird, E. E.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, (14), 6981-6985.

[29] Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1998,** 120, (25), 6219-6226.

[30] Wang, C. C. C.; Ellervik, U.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2001,** 9, (3), 653-657.

[31] Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1998,** 120, (14), 3534-3535.

[32] de Clairac, R. P. L.; Seel, C. J.; Geierstanger, B. H.; Mrksich, M.; Baird, E. E.; Dervan, P. B.; Wemmer, D. E., *Journal of the American Chemical Society* **1999,** 121, (13), 2956-2964.

[33] Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1999,** 121, (6), 1121-1129.

[34] Melander, C.; Herman, D. M.; Dervan, P. B., *Chemistry-a European Journal* **2000,** 6, (24), 4487-4497.

[35] Weyermann, P.; Dervan, P. B., *Journal of the American Chemical Society* **2002,** 124, (24), 6872-6878.

[36] Kers, I.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2002,** 10, (10), 3339-3349.

[37] Nguyen, D. H.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2001,** 9, (1), 7-17.

[38] Marques, M. A.; Doss, R. M.; Urbach, A. R.; Dervan, P. B., *Helvetica Chimica Acta* **2002,** 85, (12), 4485-4517.

[39] Doss, R. M., Marques, M. A., Foister, S., Dervan, P. B., *Chemistry & Biodiversity* **2004**, submitted.

[40] Zhan, Z. Y. J.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2000,** 8, (10), 2467-2474.

[41] Ellervik, U.; Wang, C. C. C.; Dervan, P. B., *Journal of the American Chemical Society* **2000,** 122, (39), 9354-9360.

[42] Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2003,** 11, (20), 4333-4340.

[43] Burli, R. W.; Jones, P.; McMinn, D.; Le, Q.; Duan, J. X.; Kaizerman, J. A.; Difuntorum, S.; Moser, H. E., *Bioorganic & Medicinal Chemistry Letters* **2004,** 14, (5), 1259-1263.

[44] Minehan, T. G.; Gottwald, K.; Dervan, P. B., *Helvetica Chimica Acta* **2000,** 83, (9), 2197-2213.

[45] Burli, R. W.; McMinn, D.; Kaizerman, J. A.; Hu, W. H.; Ge, Y. G.; Pack, Q.; Jiang, V.; Gross, M.; Garcia, M.; Tanaka, R.; Moser, H. E., *Bioorganic & Medicinal Chemistry Letters* **2004,** 14, (5), 1253-1257.

[46] Briehn, C. A.; Weyermann, P.; Dervan, P. B., *Chemistry-a European Journal* **2003,** 9, (9), 2110-2122.

[47] Renneberg, D.; Dervan, P. B., *Journal of the American Chemical Society* **2003,** 125, (19), 5707-5716.

[48] Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E., *Proceedings of the National Academy of Sciences of the United States of America* **1985,** 82, (5), 1376-1380.

[49] Janssen, S.; Durussel, T.; Laemmli, U. K., *Molecular Cell* **2000,** 6, (5), 999-1011.

[50] Dervan, P. B.; Urbach, A. R., *The Importance of β-Alanine for Recognition of the Minor Groove of DNA*. ed.; Verlag Helvetica Chimica Acta: Zurich, 2001; 'Vol.' p 327-339.

[51] Urbach, A. R.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **2001,** 98, (8), 4343-4348.

[52] Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B., *Journal of Molecular Biology* **2002,** 320, (1), 55-71.

[53] Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1996,** 118, (26), 6141-6146.

[54] Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2002,** 10, (8), 2767-2774.

[55] Fattori, D.; Kinzel, O.; Ingallinella, P.; Bianchi, E.; Pessi, A., *Bioorganic & Medicinal Chemistry Letters* **2002,** 12, (8), 1143-1147.

[56] Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Organic Letters* **2001,** 3, (8), 1201-1203.

[57] Dudouet, B.; Burnett, R.; Dickinson, L. A.; Wood, M. R.; Melander, C.; Belitsky, J. M.; Edelson, B.; Wurtz, N.; Briehn, C.; Dervan, P. B.; Gottesfeld, J. M., *Chemistry & Biology* **2003,** 10, (9), 859-867.

[58] Gottesfeld, J. M.; Melander, C.; Suto, R. K.; Raviol, H.; Luger, K.; Dervan, P. B., *Journal of Molecular Biology* **2001,** 309, (3), 615-629.

[59] Suto, R. K.; Edayathumangalam, R. S.; White, C. L.; Melander, C.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., *Journal of Molecular Biology* **2003,** 326, (2), 371-380.

[60] Gottesfeld, J. M.; Belitsky, J. M.; Melander, C.; Dervan, P. B.; Luger, K., *Journal of Molecular Biology* **2002,** 321, (2), 249-263.

[61] Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2002,** 10, (10), 3313-3318.

[62] Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **2003,** 100, (21), 12063-12068.

[63] Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Nature* **1997,** 387, (6629), 202-205.

[64] Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B.; Gottesfeld, J. M., *Journal of Molecular Biology* **1997,** 274, (4), 439-445.

[65] McBryant, S. J.; Baird, E. E.; Trauger, J. W.; Dervan, P. B.; Gottesfeld, J. M., *Journal of Molecular Biology* **1999,** 286, (4), 973-981.

[66] Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **1998,** 95, (22), 12890-12895.

[67] Ehley, J. A.; Melander, C.; Herman, D.; Baird, E. E.; Ferguson, H. A.; Goodrich, J. A.; Dervan, P. B.; Gottesfeld, J. M., *Molecular and Cellular Biology* **2002,** 22, (6), 1723-1733.

[68] Wurtz, N. R.; Dervan, P. B., *Chemistry & Biology* **2000,** 7, (3), 153-161.

[69] Takahashi, R.; Bando, T.; Sugiyama, H., *Bioorganic & Medicinal Chemistry* **2003,** 11, (12), 2503-2509.

[70] Dickinson, L. A.; Trauger, J. W.; Baird, E. E.; Dervan, P. B.; Graves, B. J.; Gottesfeld, J. M., *Journal of Biological Chemistry* **1999,** 274, (18), 12765-12773.

[71] Wurtz, N. R.; Pomerantz, J. L.; Baltimore, D.; Dervan, P. B., *Biochemistry* **2002,** 41, (24), 7604-7609.

[72] Chiang, S. Y.; Burli, R. W.; Benz, C. C.; Gawron, L.; Scott, G. K.; Dervan, P. B.; Beerman, T. A., *Journal of Biological Chemistry* **2000,** 275, (32), 24246-24254.

[73] Lenzmeier, B. A.; Baird, E. E.; Dervan, P. B.; Nyborg, J. K., *Journal of Molecular Biology* **1999,** 291, (4), 731-744.

[74] Bremer, R. E.; Baird, E. E.; Dervan, P. B., *Chemistry & Biology* **1998,** 5, (3), 119-133.

[75] Bremer, R. E.; Wurtz, N. R.; Szewczyk, J. W.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2001,** 9, (8), 2093-2103.

[76] Fechter, E. J.; Dervan, P. B., *Journal of the American Chemical Society* **2003,** 125, (28), 8476-8485.

[77] Nguyen-Hackley, D. H.; Ramm, E.; Taylor, C. M.; Joung, J. K.; Dervan, P. B.; Pabo, C. O., *Biochemistry* **2004,** 43, (13), 3880-3890.

[78] Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **2004,** 101, (48), 16768-16773.

[79] Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M., *Chemistry & Biology* **2004,** 11, (11), 1583-1594.

[80] Dickinson, L. A.; Trauger, J. W.; Baird, E. E.; Ghazal, P.; Dervan, P. B.; Gottesfeld, J. M., *Biochemistry* **1999,** 38, (33), 10801-10807.

[81] Coull, J. J.; He, G. C.; Melander, C.; Rucker, V. C.; Dervan, P. B.; Margolis, D. M., *Journal of Virology* **2002,** 76, (23), 12349-12354.

[82] Ansari, A. Z.; Mapp, A. K., *Current Opinion in Chemical Biology* **2002,** 6, (6), 765-772.

[83] Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (8), 3930-3935.

[84] Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., *Chemistry & Biology* **2001,** 8, (6), 583-592.

[85] Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., *Journal of the American Chemical Society* **2002,** 124, (44), 13067-13071.

[86] McGinnis, W.; Krumlauf, R., *Cell* **1992,** 68, (2), 283-302.

[87] Arndt, H. D.; Hauschild, K. E.; Sullivan, D. P.; Lake, K.; Dervan, P. B.; Ansari, A. Z., *Journal of the American Chemical Society* **2003,** 125, (44), 13322-13323.