Chapter 1

Introduction: The Evolution of DNA-Binding Small Molecules

The text of this chapter was taken in part from a book chapter coauthored with Eric J. Fechter, Benjamin S. Edelson, and Professor Peter B. Dervan (Caltech)

(Dervan, P. B.; Poulin-Kerstien, A. T.; Fechter, E. J.; Edelson, B. S. "Regulation of Gene Expression by Synthetic DNA-Binding Ligands." Top. Curr. Chem. **2005**, *253*, 1–31)

Abstract

During the past 20 years, polyamides have evolved from the natural product distamycin to a new class of programmable heterocyclic oligomers that bind a broad repertoire of DNA sequences with high affinity and specificity. This chapter details recent advances in this field of research, focusing on molecular recognition of DNA, and biological applications such as modulating gene expression by small molecules.

Background and Significance.

Sequence-Specific Recognition of DNA: From Natural Roots.

The natural product distamycin contains three *N*-methylpyrrole (Py) amino acids and binds in the minor groove of DNA at A,T tracts 4–5 base pairs (bp) in size.[1, 2] Distamycin inhibits DNA dependent processes, including transcription, and has antibacterial,[3] antimalarial,[4] antifungal,[5] and antiviral activities,[6] but is of limited use because of toxicity.[7]

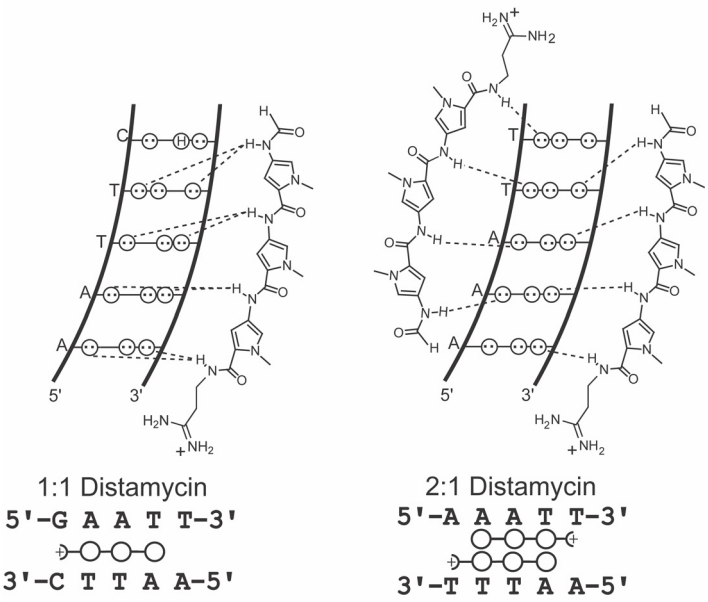


Figure 1.1. Schematic representation of the two DNA-binding modes of distamycin. Hydrogen bonds are shown as dashed lines. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Open circles below represent pyrrole amino acid units.

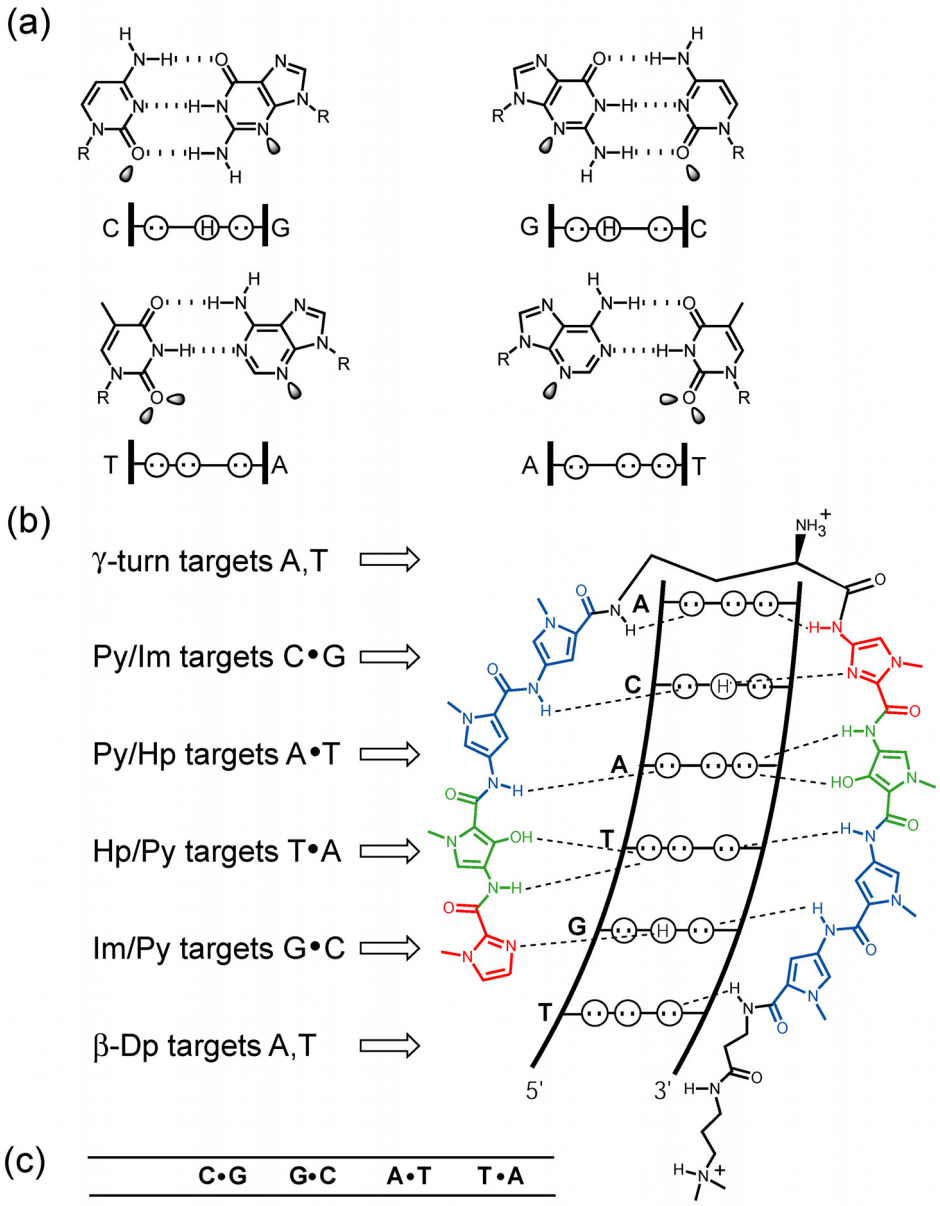
Structural studies of distamycin•DNA complexes revealed that the crescent-shaped molecule binds A,T tracts in both 2:1 and 1:1 ligand:DNA stoichiometries (Figure 1.1).[8-10] These structures revealed key ligand-DNA interactions, such as a series of hydrogen bonds between pyrrole carboxamides and the edges of the nucleobases on the adjacent DNA strand. These studies revealed that both the shape complementarity and the specific hydrogen-bonding profile of distamycin account for its affinity and

specificity toward B-form DNA. Over the nearly 20 years since the original structural work, analogues have been created and characterized that bind a large number of different DNA sequences in a predictable fashion.[11] We now have a set of 5- and 6-membered heterocyclic amino acids that can be combined as modular, antiparallel ring pairs in the minor groove of DNA to recognize predetermined sequences of DNA with affinities and specificities comparable to DNA-binding proteins.[12] Presented here are recent advances in the field of DNA recognition by polyamides.

The Pairing Rules.

In a formal sense, the four Watson-Crick base pairs can be differentiated on the minor groove floor by the specific positions of hydrogen bond donors and acceptors, as well as by differences in molecular shape and electronic potential surfaces (Figure 1.2a).[12] The exocylic amine of guanine presents an unsymmetrical hydrogen bond donor "bump" on the minor groove edge of a G•C base pair. A key study in the early 1990s demonstrated that the *N*-methylimidazole (Im)-containing polyamide ImPyPy bound to the five bp sequence 5'-WGWCW-3' (where W = A or T).[13] This result was rationalized in terms of the formation of a 2:1 polyamide-DNA complex in which an antiparallel ring pairing of Im stacked against Py could specifically distinguish a G•C from the other three base pairs (Figures 1.2 and 1.3).

The Im/Py pair has been explored by extensive studies, including analyses of binding in hundreds of different sequence contexts. Crystal structures confirmed the existence of a hydrogen bond between the Im nitrogen and the exocyclic NH_2 of guanine when the Im/Py pair binds opposite the G•C base pair.[14] The preference for a linear



(c)

	C•G	G•C	A•T	T•A
Py/Im	+	−	−	−
Im/Py	−	+	−	−
Py/Hp	−	−	+	−
Hp/Py	−	−	−	+

Figure 1.2. (a) Structures of the four Watson-Crick base pairs. The R group represents the sugar-phosphate backbone of DNA, and shaded orbitals represent electron lone pairs projecting into the minor groove. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles with an H represent the exocyclic 2-amino group of guanine. (b) Schematic model for the hairpin polyamide ImHpPyPy-*γ*-ImHpPyPy-*β*-Dp bound to its match site 5'-TGTACA-3' as determined by the pairing rules for recognition of all four Watson-Crick base pairs of DNA in the minor groove by polyamides. Putative hydrogen bonds are indicated as dashed lines. (c) Table indicating the code for minor groove recognition by polyamides. Plus and minus signs indicate favored and disfavored interactions, respectively.

hydrogen bond, coupled with the unfavorable angle to an Im over the cytosine side of the base pair, provides a basis for the ability of an Im/Py pair to discriminate G•C from C•G (Figure 1.3). Thermodynamic investigations dissected binding free energies into enthalpic and entropic contributions, revealing that the sequence selectivity of the Im/Py pair is driven by a favorable enthalpic contribution.[15]

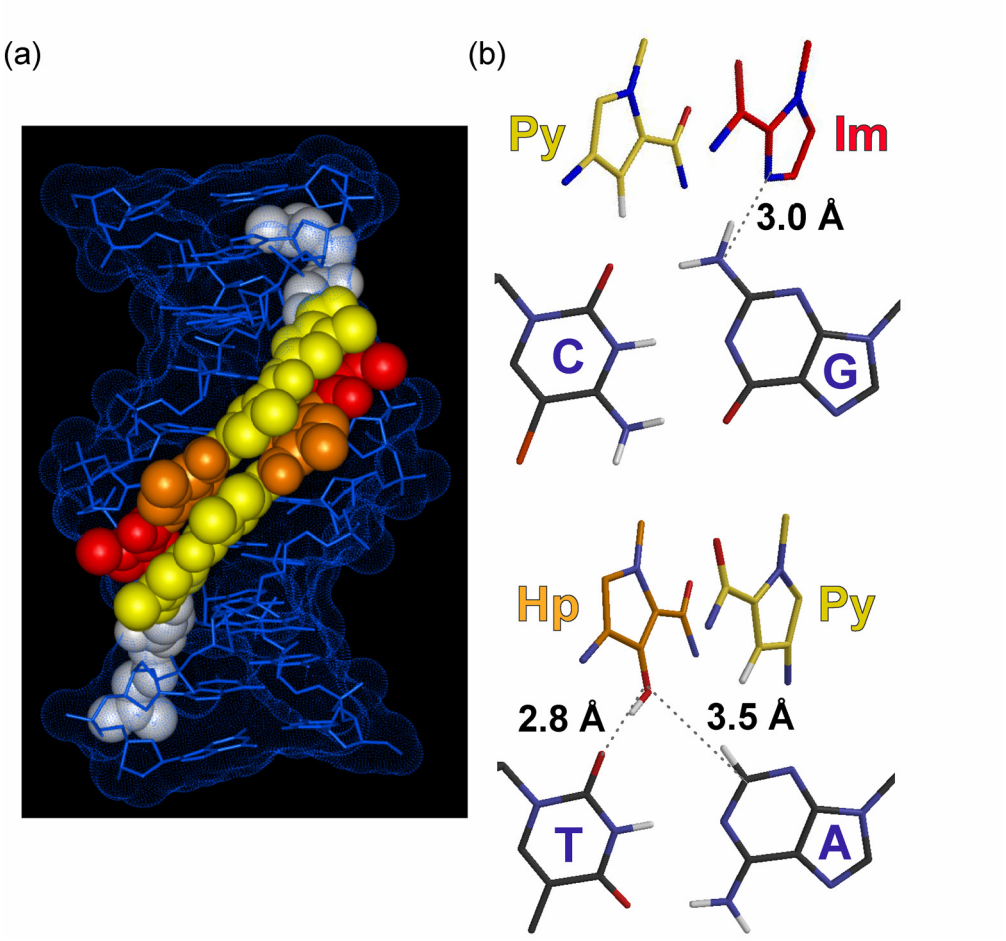


Figure 1.3. (a) X-ray crystal structure of ImHpPyPy-*β*-Dp (Dp = dimethylaminopropylamine) bound in a 2:1 complex with its target DNA site, 5'-AGTACT-3' (PDB code 407D). Im residues are red, Hp residues are orange, Py residues are yellow. (b) Detail of the Py/Im pair interacting with the C•G base pair (top) and of the Hp/Py pair interacting with the T•A base pair. Dashed lines indicate interatomic distances between carbon, nitrogen, and oxygen atoms emphasizing the close interactions responsible for specificity.

Discrimination of the A, T base pairs was not achieved until relatively recently. The A•T base pair appears fairly symmetrical, with both adenine and thymine presenting

a hydrogen bond acceptor to the floor of the minor groove (Figure 1.2a). However, closer inspection reveals that a small asymmetric cleft is formed between the thymine O2 and adenine C2. Furthermore, the N3 of adenine presents only one lone pair while the O2 of thymine presents two lone pairs capable of hydrogen bonding. Informed by high resolution crystallographic data from a polyamide-DNA complex, *N*-methyl-3-hydroxypyrrole (Hp) was designed, and subsequently proved to be a thymine-selective recognition element when paired across from Py (Figures 1.2 and 1.3).[16]

Crystal structures of two different Hp-containing polyamides, as their 2:1 complexes with DNA, have been determined at high resolution.[17, 18] The specificity of an Hp/Py pair was shown to arise from a combination of specific hydrogen bonds between the hydroxyl and the thymine O2, along with shape-selective recognition of the asymmetric cleft (Figure 1.3). Hp polyamides bind with lower affinities than their Py counterparts,[16] and a recent computational study argues that desolvation of the hydroxyl group upon insertion into the minor groove accounts for the energetic penalty.[19] Together, three rings–Py, Im, and Hp–can be combined as unsymmetrical pairs to recognize specifically each of the four Watson-Crick base pairs; Im/Py is specific for G•C and Hp/Py for T•A (Figure 1.2b). These interactions can be conveniently described as pairing rules (Figure 1.2c). The pairing rules should be considered as guidelines only. Antiparallel polyamide dimers bind B-form DNA, and there are limitations regarding sequences targeted due to the sequence-dependent microstructure of DNA.

Expanding and Refining DNA Recognition.

Improving Affinity and Specificity.

Covalent linkage of the two antiparallel polyamide strands results in molecules with increased affinity and specificity (Figure 1.4a). Currently, the "standard" motif is the 8-ring hairpin, in which a γ-aminobutyric acid linker (γ-turn) connects the carboxylic terminus of one polyamide to the amino terminus of its antiparallel partner.

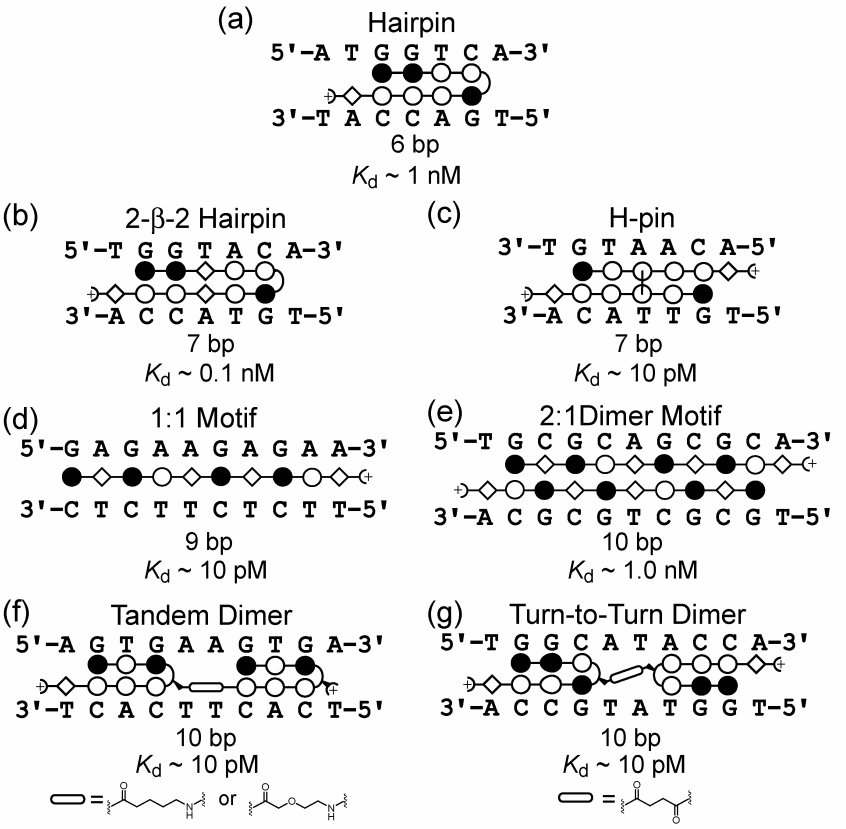


Figure 1.4. Polyamide-DNA binding motifs with approximate dissociation constants (K_d). (a) Hairpin: standard motif targeting 6 bp with high affinity and sequence specificity. (b) β-Ala-containing polyamides are able to bind longer sequences because the flexible aliphatic residue relaxes the curvature of the polyamide. (c) 2:1 Complex: β-containing polyamides are able to target long stretches of DNA as either 2:1 complexes (shown here) or 1:1 complexes depending on stoichiometry. (d) H-pin: covalent linkage of two polyamide strands is achieved at a position that is sequence-neutral. (f, g) Hairpin Dimers: linked either turn-to-tail (f) or turn-to-turn (g), these molecules are able to target long sequences of DNA with moderate specificity and high affinity. Optimized linkers for each motif are shown below. The black and open circles represent Im and Py rings, respectively; diamonds represent β-alanine residues; and plus signs next to diamonds represent Dp residues. A curved line connecting the sides of two circles represents the γ-aminobutyric acid turn, and a curved line with a wedge and a plus sign represents the chiral $(R)^{H2N}\gamma$-turn. For the H-pin, curved lines connecting the centers of two circles represent alkyl linkers attached to the N-methyl positions of the aromatic rings.

Compared to the unlinked homodimers, hairpin polyamides display ~100-fold higher affinity, with the *γ*-turn demonstrating selectivity for A,T over G,C base pairs (Figure 1.2b), presumably due to a steric clash between the aliphatic turn unit and the exocyclic amine of guanine.[20] 8-ring hairpins, which bind six bp, were shown to have affinities and specificities similar to DNA-binding proteins (i.e., K_d < 1 nM).[21] NMR studies confirmed that the *γ*-turn locks the register of the ring pairings, preventing the ambiguity of slipped dimers.[22] Hairpin polyamides retain the orientation preferences of unlinked antiparallel polyamides, aligning N→C with respect to the 5'→3' direction of the adjacent DNA strand.[23] The β-alanine-dimethylaminopropyl amine (*β*-Ala, Dp, respectively) tail substituent at the C-terminus of many hairpin polyamides is an A,T specific element (Figure 1.2b), again presumably due to a negative interaction between the aliphatic chain and the exocyclic amine of guanine.[24] The tail is thought to play a role in the orientation preference of polyamides.[25]

For some hairpins, however, "reversed binding" (a C→N alignment of the polyamide with respect to the 5'→3' direction of the adjacent DNA strand) has been observed as the preferred orientation.[23] By introducing an amino substituent at the *α* position of the γ-turn, reversed binding is disfavored because of a steric clash between the amino substituent and the floor of the minor groove.[26] Not only does the chiral turn maintain the specificity of hairpins, it increases the overall affinity, either by the addition of a positive charge, which interacts favorably with the negatively charged backbone or minor groove floor of the DNA polymer, or by sterically reducing the conformational freedom of the polyamide in solution.[26]

In other cases, polyamides containing aliphatic residues such as the *γ*-turn or *β*-

Ala have been shown to favor an extended 1:1 binding mode (Figure 1.4d). Depending on the stoichiometry, the *same* polyamide may bind *different* sequences.[27] For example, a hairpin containing internal *β*-alanines might bind in two different confirmations, a "folded" versus an "extended" conformation. An amino substituent on the γ-turn also disfavors the extended binding mode, and serves to lock the polyamide into the hairpin conformation.[28] A larger, N-acetyl group increases this effect, with an 8-ring, N-acetyl-substituted polyamide favoring the hairpin conformation over extended binding by >25,000-fold. Substitution of *β*-Ala for Py has also been shown to influence the tendency of a polyamide to bind in a hairpin conformation.[29] In each case, it is presumably a negative interaction between the substituent and the wall of the minor groove that drives the equilibrium toward the hairpin conformation.

Binding-Site Size.

While the standard 8-ring hairpin recognizes DNA with high affinity and specificity, it targets only six bp. For biological applications, binding site size may be critical because longer sequences would be expected to occur less frequently in the genome. Early attempts to increase the targeted site size by simply extending the number of aromatic rings resulted in polyamides with decreased affinity.[30] Crystal structures of polyamide/DNA complexes have shown that the polyamide rise-per-residue matches the pitch of the B-DNA helix–that is, the spacing of the pyrrole and imidazole rings matches the spacing of the DNA base pairs.[14, 17, 18] However, the inherent crescent-shaped curvature of polyamides is slightly tighter than the curvature of the minor groove. Beyond five contiguous rings, the shape of a polyamide is no longer complementary to

DNA, and the resulting loss of specific contacts accounts for the observed loss in affinity and specificity.[14]

The hypercurvature of polyamides can be overcome by the introduction of the flexible *β*-alanine residue. When introduced as a Py replacement,[31] its flexibility relaxes the curvature of the polyamide, restoring complementarity to DNA. *β*-Ala-containing polyamides have been used to target up to 9 bp with high affinity and specificity (Figure 4b, d, e).[32]

Another strategy to target longer sequences of DNA is to covalently link two hairpin polyamides with a flexible linker. These dimeric polyamides, linked both turn-to-tail and turn-to-turn, have been shown to bind longer sequences with high affinity (Figure 1.4f, g).[33-35] Both turn-to-turn and turn-to-tail dimers with optimized linkers showed good selectivity for a 10 bp site (over 11 and 12 bp sites) but exhibited poor specificity (expressed in terms of affinity for match over single base pair mismatch sites). Nonetheless, an impressive application of the tandem motif was demonstrated by Laemmli and coworkers, who used tandem-hairpins to stain insect or vertebrate telomeres, $(TTAGG)_n$ or $(TTAGGG)_n$ repeats, respectively, with remarkable selectivity in fixed cells and chromosome spreads.[36]

H-Pin and U-Pin Motifs.

Polyamides also can be linked, via the ring nitrogens, with an alkyl spacer that projects away from the minor groove. When placed in the center of a polyamide, the resultant branched molecule has been termed an H-pin (Figure 1.4c); when placed at the end, a U-pin. H-pin polyamides bind with high affinity and good specificity.[37] Recent

efforts to improve the synthetic methods for H-pins using alkene metathesis on a solid support have enabled a detailed study of the optimal alkyl linker length, demonstrating that four and six methylene units provide the highest affinities.[38] U-pin polyamides behave similarly.[39] The affinity of an 8-ring U-pin is most comparable to a 6-ring hairpin polyamide, likely due to a loss of two hydrogen bond donors upon removal of the γ-turn element. Thus, the dimeric Py-Im U-turn element may be thought of as a C•G specific replacement for the *γ*-turn. In combination with removal of the *β*-Ala-Dp tail, H-pin and U-pin polyamides could potentially bind purely G,C sites, a sequence type that has been difficult to target with other polyamide motifs.

1:1 Polyamide:DNA Complexes.

Homopurine tracts have been a challenging target for polyamide binding because these DNA sequences have particularly narrow minor grooves.[40] While the width of the minor groove of these DNA sequences may be too narrow for hairpin polyamides, a single Py-Im-β-containing strand may be accommodated at such sequences.

In 1:1 polyamide•DNA complexes, β-linked polyamides appear to prefer a single orientation, N→C with respect to the 3'→5' direction of the purine-rich strand.[41] Footprinting of 1:1 complexes has shown that Im residues do not distinguish G,C from A,T whereas Py and β residues prefer A,T over G,C base pairs.[41]

A high-resolution 1:1 solution NMR structure of ImPy-*β*-Im-*β*-ImPy-*β*-Dp elucidated the role of *β*-alanine in minor groove recognition (Figure 1.5).[42] *β*-Ala allows both Im rings in the *β*-Im-*β*-Im subunit to adapt to the relatively large dihedral required for hydrogen bonding. Additionally, close contacts of *β*-alanine to the floor of the minor

groove provided a structural explanation for its observed A/T specificity.

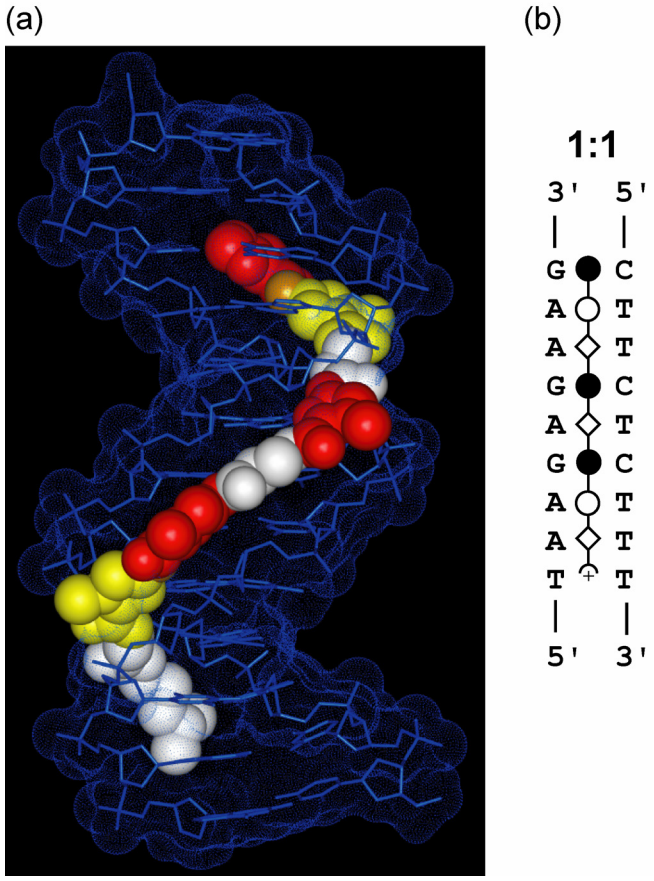


Figure 1.5. (a) NMR Structure of a 1:1 polyamide:DNA complex. The DNA backbone and bases are shown in blue. Im residues are red, Py residues are yellow, *β*-Ala and Dp residues are white. (b) Schematic representation of the structure shown in (a).

While the entire recognition code in the 1:1 motif has not been fully elucidated, this type of binding mode is uniquely suited to targeting homopurine sequences.[41, 43] Laemmli and coworkers reported a striking example of a 1:1 motif polyamide targeted to the satellite regions of *Drosophila melanogaster* being able to induce specific gain-of-function and loss-of-function phenotypes when fed to developing flies.[44]

Improving Synthetic Methodology.

The investigation of minor groove-binding polyamides was greatly accelerated by

the implementation of solid-phase synthesis.[45] Originally demonstrated on Boc-*β*-Ala-PAM resin with Boc-protected monomers, it was also shown that Fmoc chemistry could be employed with suitably protected monomers and Fmoc-*β*-Ala-Wang resin.[46] Recently, Pessi and coworkers used a sulfonamide-based safety-catch resin to prepare derivatives of hairpin polyamides.[47] Upon activation of the linker, resin-bound polyamides were readily cleaved with stoichiometric quantities of nucleophile to provide thioesters or peptide conjugates.

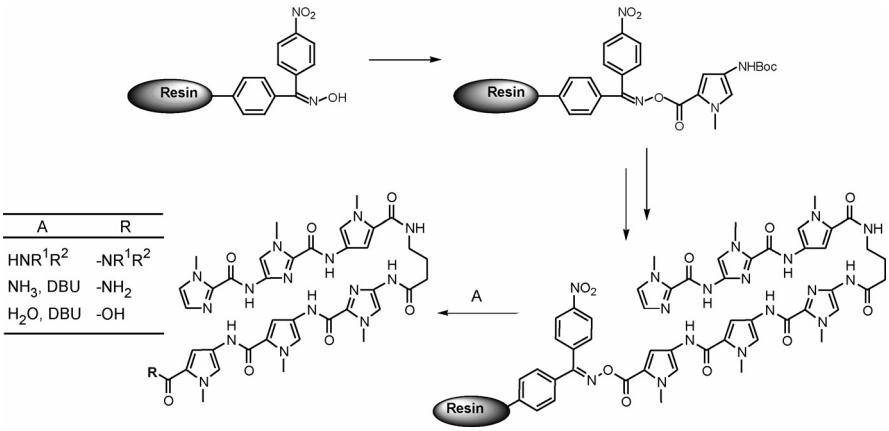


Figure 1.6. Scheme for synthesis of polyamides on the Kaiser oxime solid-support resin. Cleavage from resin with various reagents (**A**) can result in polyamides with shorter C-terminal groups than molecules prepared on *β*-Ala-PAM resin. The amine HNR^1R^2 may be a primary or secondary alkyl amine.

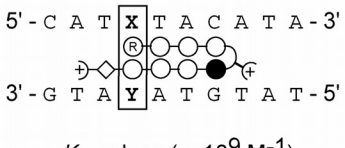
While allowing rapid preparation of a range of polyamides, these resins install a T,A selective *β*-Ala residue at the C-terminus, which places limits on the DNA sites that can be targeted.[24] The shortest tail available from these resins is a propanolamide, obtained by reductive cleavage. Polyamides prepared on Boc-Gly-PAM resin can be reductively cleaved to obtain ethanolamide tails, but it was expected that further truncation of the C-terminus would be necessary for tolerance of G,C at the tail position.[24] The Kaiser oxime resin was therefore adapted to polyamide synthesis, allowing the preparation of polyamides with shorter C-termini (Figure 1.6). These

molecules display the desired tolerance for G,C bases while maintaining high affinities.[48]

New Ring Systems.

The specificity of cofacial aromatic amino acid pairings is highly dependent upon their position within a given polyamide. For example, an Im/Py pairing is specific for G•C at both internal and terminal positions.[13] In contrast, Hp/Py and Py/Py pairings, while specific for T•A and A,T, respectively at internal positions, lose all specificity when incorporated at the *N*-terminal cap position. The loss of specificity at the cap position is presumably a result of conformational freedom caused by the absence of a second "groove-anchoring" carboxamide, allowing terminal rings to bind DNA in either of two rotamers.

(a)



K_a values (x 10^9 M^{-1})

X·Y				
T·A	3.8	5.1	**2.0**	**13**
A·T	2.8	3.1	0.32	3.7
C·G	3.2	0.26	≤ 0.01	0.21
G·C	**70**	1.1	≤ 0.01	0.31

(b)

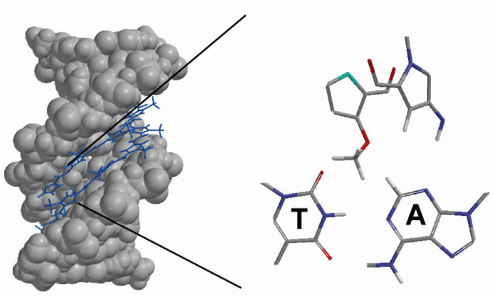


Figure 1.7. (a) Table indicating the equilibrium association constants (K_a) for polyamides containing the indicated ring at the N-terminal position of the 8-ring hairpin polyamide pictured above at sequences containing each of the four Watson-Crick base pairs. When paired with Py, Im shows a preference for guanine while 3-substituted thiophenes show a preference for thymine. (b) Model created using the PC Spartan (Wavefunction, Inc.) illustrating the shape-selective recognition of thymine by the 3-methoxy substituent of thiophene.

Recently, a library of substituted five-membered aromatic carboxylic acids was incorporated at the *N*-terminal position of an 8-ring polyamide, and *N*-terminal specificity

was probed.[49] It was found that 3-chlorothiophene (Ct) and 3-methoxythiophene (Mt) rings, when paired against Py, imparted a moderate degree of specificity for T•A over A•T (6- and 3-fold, respectively), and a large degree of specificity over G,C base pairs (>200-fold) (Figure 1.7a). While the Mt/Py imparts a higher degree of specificity for T•A over A•T, the Ct/Py pair imparts modest specificity as well as higher affinity. Molecular modeling of Mt- and Ct-containing polyamides indicated that the rotamer that places the 3-substituent into the floor of the minor groove is energetically favored. This substituent fills the asymmetric cleft created by the N2 of thymine and the C2 of adenine, accounting for the observed preference for T•A (Figure 1.7b).

While most polyamide research has focused on 5-membered aromatic ring systems, other scaffolds have been shown to bind DNA. The benzimidazole ring system represents a different structural framework that is amenable to functionalization on the 6-membered ring and appears to impart a curvature that is complementary to DNA.[50] Indeed, the classic minor groove-binding Hoechst dyes are composed of benzimidazole units, and a number of derivatives of these molecules have been prepared. We have incorporated benzimidazole derivatives into the backbones of hairpin polyamides in a manner that preserves critical hydrogen bonding contacts and overall molecular shape (Figure 1.8).[51, 52] The imidazopyridine (Ip) and hydroxybenzimidazole (Hz) rings are introduced into polyamides as the dimeric subunits PyIp and PyHz, respectively, in which the Py ring is directly connected to the benzimidazole derivative without an intervening amide bond.[51, 52] DNase I footprinting indicates that the Ip/Py and Hz/Py pairs are functionally identical, at least in some sequence contexts, to the analogous five-

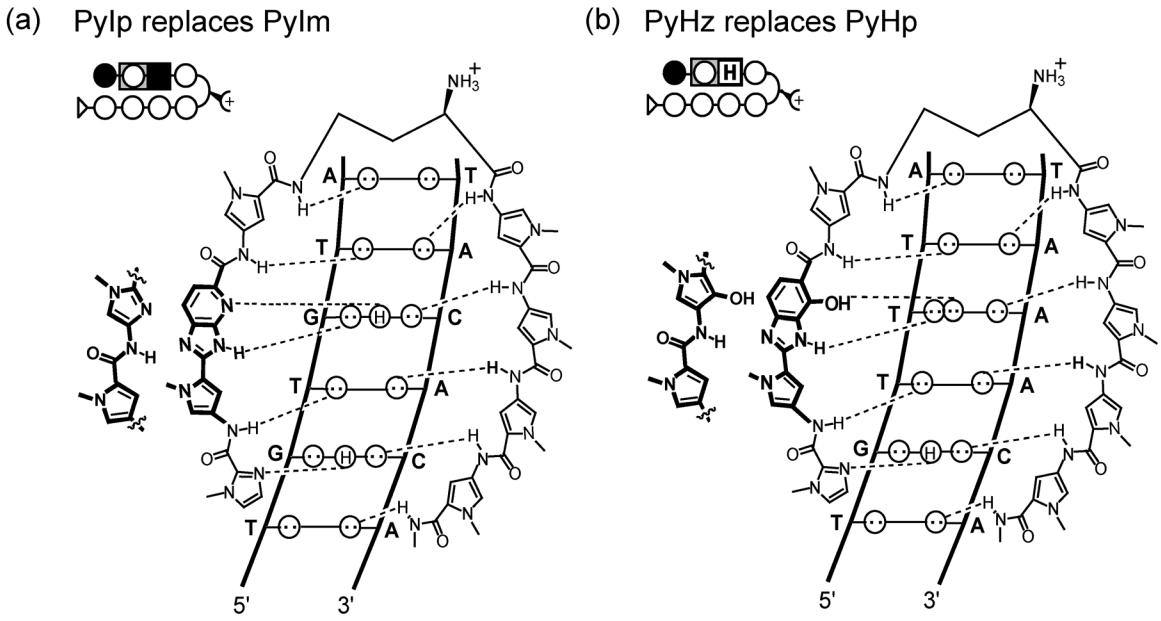


Figure 1.8. Recognition of the DNA minor groove with benzimidazole-derivatives. (a) Structure of polyamide containing Py-hydroxybenzimidazole (PyHz). (b) Structure of polyamide containing Py-imidazopyridine (PyIp). The 5-membered ring analogues of the dimeric benzimidazole derivatives are shown to the right of each model. Putative hydrogen bonds are indicated with dashed lines. Rectangles containing a white circle and a black square indicate the PyIp unit. Rectangles containing a white circle and a square with an H indicated the PyHz unit. Other symbols are defined in figures 1.2 and 1.4.

membered ring pairs Im/Py and Hp/Py.[52] Significantly, we have found that the Hp-containing polyamides can degrade over time in the presence of acid or free radicals, whereas the analogous Hz-containing compounds are chemically robust. Thus, the Hz/Py pair is a strong candidate for replacing Hp/Py in biological studies.

Secondary Effects of Polyamides.

Displacement of DNA-Binding Proteins.

Polyamides bind with high affinity to a wide range of DNA sites and can competitively displace many proteins from DNA. This can have an effect on gene expression, as DNA-binding proteins are often involved in the regulation of transcription.

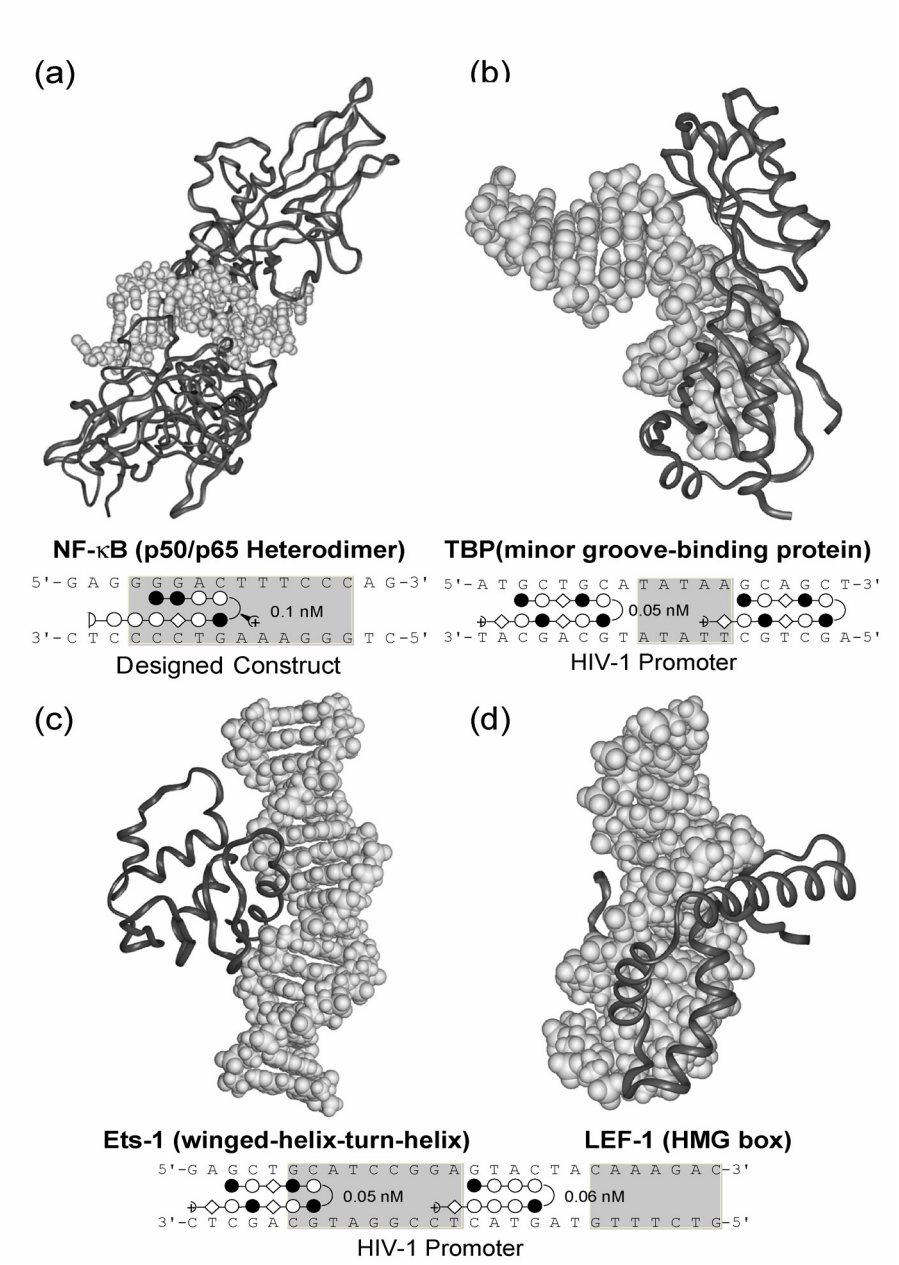


Figure 1.9. X Ray crystal structures of four different protein-DNA complexes that have been inhibited by polyamides. Below each structure is illustrated the context for inhibition, with the protein binding sites shaded and the polyamides responsible for inhibition shown bound to their match sites. Listed beside each polyamide is its dissociation constant (K_d). Symbols are as defined in figure 1.4.

One approach to modifying gene expression involves inhibition of key transcription factor (TF)-DNA complexes in a designated promoter, thus interfering with recruitment of RNA polymerases.[53] Significantly, because there are considerably fewer oncogenic TFs than potentially oncogenic signaling proteins, TF inhibition represents a uniquely

promising approach to cancer treatment.[54] The transcription factor TFIIIA was chosen as a first target because it regulates a relatively small number of genes and because the contacts between the nine zinc-finger protein and the minor groove had been established. A polyamide bound in the recognition site of TFIIIA suppressed transcription of 5 S RNA genes by RNA polymerase III *in vitro* and in cultured *Xenopus* kidney cells (Figure 1.9a).[53] Further studies used polyamides in combination with recombinant derivatives of TFIIIA subunits to elucidate essential minor groove contacts for the binding of this TF.[55]

Polyamides were then used to target viral genes transcribed by RNA polymerase II. The HIV-1 enhancer/promoter contains binding sites for multiple transcription factors, including TBP, Ets-1, and LEF-1. Two hairpin polyamides designed to bind DNA sequences immediately adjacent to the binding sites for LEF-1 and Ets-1 specifically inhibited binding of each transcription factor and prevented HIV-1 transcription in a cell-free assay (Figures 1.9b, c and d).[56] In human blood lymphocytes, treatment with the two polyamides in combination inhibited viral replication by 99%, with no significant decrease in cell viability. Inhibition of viral replication is indirect evidence for specific transcription inhibition by polyamides, because other modes of action could be involved, such as modulation of T-cell activation pathways. However, RNase protection assays indicated that the two polyamides did not alter the RNA transcript levels of several cytokine and growth-factor genes, suggesting that polyamides do affect transcription directly.

This early biological result spurred a variety of biochemical studies of the interactions of polyamides with the basal transcription machinery and TF-DNA complexes. Two studies have used promoter scanning to identify sites where polyamide

binding inhibits transcription.[57, 58] The method uses a series of DNA constructs with designed polyamide binding sites at varying distances from the transcription start site. Essential minor groove contacts were identified for a subunit of TFIIIB (possibly TBP) in a *Xenopus* tRNA promoter,[58] as well as for TFIID-TFIIA and TBP in the HIV-1 core promoter.[57] The binding of the homodimeric basic-helix-loop-helix TF Deadpan was investigated using a variant of promoter scanning.[59] A series of duplex oligonucleotides based on a *Drosophila* neural promoter were designed, incorporating polyamide binding sites on different sides of the Deadpan recognition sequence and in different orientations. The TF-DNA complex was inhibited only by a polyamide binding upstream of the homodimer, establishing an asymmetric binding mode for this TF.

In the human T-cell leukemia virus type 1 (HTLV-1) promoter, polyamides targeted to G,C-rich regions flanking the viral CRE sites inhibited binding of the Tax protein and Tax transactivation *in vitro* (Figure 1.10).[60] This example illustrates several important polyamide•DNA•protein interactions. HTLV-1 genes are regulated by the major groove-binding protein CREB (CRE Binding Protein). CREB-mediated transcription is enhanced by the binding of the viral protein Tax, which makes contacts with CREB and to the minor groove at sites flanking the CREB binding site. Tax then recruits CREB binding protein (CBP) via a KIX domain on CBP, which then induces transcription. Researchers found that addition of two polyamides designed to target the Tax recognition elements inhibited Tax from associating to the CREB•DNA complex, and Tax-induced transcription was abolished. Interestingly, these polyamides bind only a few base pairs away from the CRE site, yet CREB is able to co-occupy the DNA, with CREB-meditated basal transcription remaining intact. Thus, polyamides are able to

interfere very specifically with some protein•DNA interactions while leaving other nearby interactions unaffected.

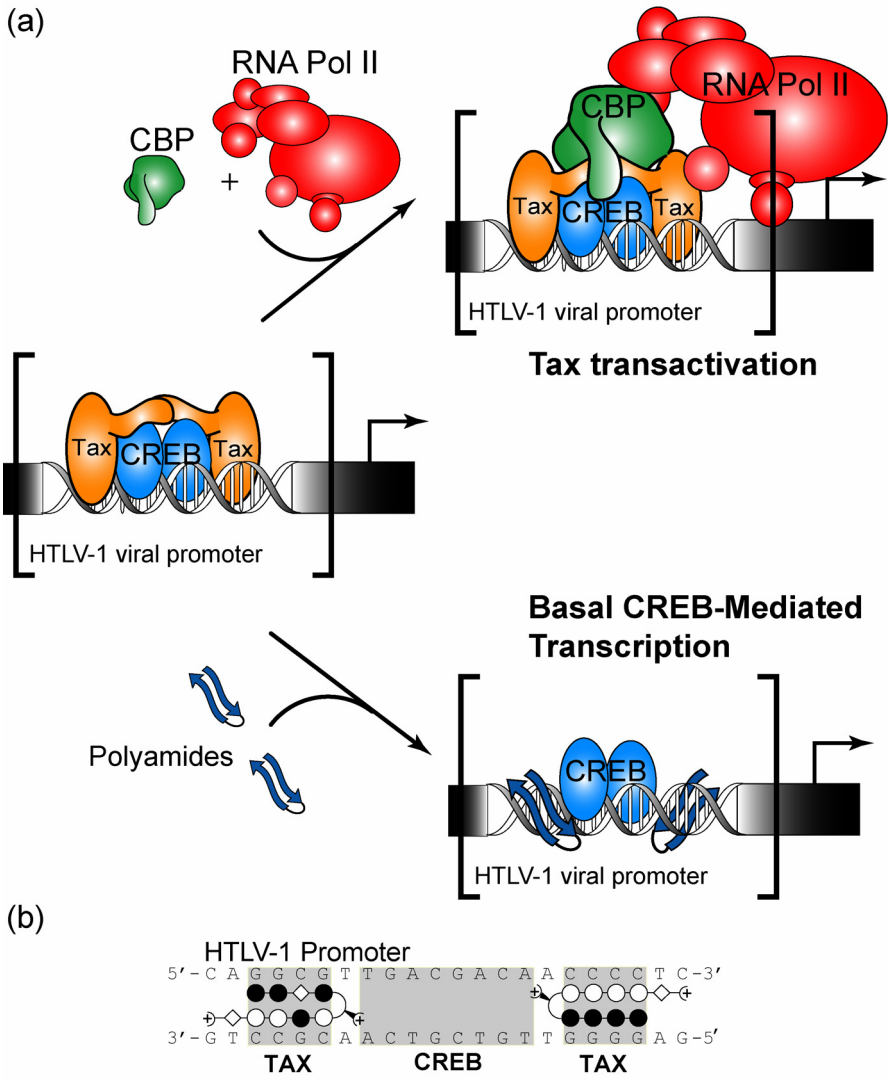


Figure 1.10. (a) Polyamide inhibition of Tax transactivation. Tax and CREB are bound to the HTLV-1 viral promoter. The trimeric complex recruits CBP and the transcriptional machinery. Polyamides specifically inhibit Tax but not CREB from binding to DNA, thus abolishing Tax transactivation while leaving basal CREB-mediated transcription unaffected. (b) Model of the sequence recognized by Tax and CREB with the structures of the polyamides used to inhibit Tax shown bound to their targeted sequences. All symbols are as defined in figure 1.4, with Tax and CREB binding sites shaded.

Several other protein•DNA interactions have been inhibited with polyamides. Bacterial gyrase recognizes a short 5'-GGCC-3' site, and a polyamide targeted to this sequence inhibited gyrase-catalyzed strand cleavage at nanomolar concentrations.[61] NF-

κB is a TF crucial for development, viral expression, inflammation, and anti-apoptotic responses. The most common form is a p50-p65 heterodimer, which binds DNA in the major groove, making several phosphate contacts throughout the binding site. Polyamides targeted to the minor groove opposite p50, but not p65, inhibit DNA binding by NF-κB (Figure 1.9a).[62] In a different study, polyamides were shown to bind very near the 3' processing end of Moloney murine leukemia virus (M-MuLV) long terminal repeat (LTR) sequences, thereby inhibiting retroviral integration catalyzed by M-MuLV Integrase (IN).[63]

The binding of Ets-1 to the HIV-1 enhancer was examined in greater detail, and polyamides were shown to inhibit the formation of a ternary Ets-1–NF-κB–DNA complex.[64] Ets-1 is a winged-helix-turn-helix TF, and its key phosphate contacts on either side of the major groove can be disrupted by a polyamide in the adjacent minor groove. The report provided evidence for cooperative DNA binding by Ets-1 and NF-κB to the HIV-1 enhancer sequence. A different Ets binding site in the HER2/*neu* promoter was targeted with hairpin polyamides that successfully blocked Ets•DNA complex formation and transcription of the HER2/*neu* oncogene in a cell-free system.[65]

Recently, researchers were able to inhibit the binding of human papilloma virus (HPV) transcription factor E2 using a tandem hairpin polyamide.[66] The E2 homodimer binds exclusively in the major groove and bends the DNA towards the body of the protein. A polyamide targeted to the E2 binding site prevents such bending, thereby destabilizing the E2•DNA complex. The topological change to the DNA caused by polyamide binding is thought to be the mechanism for E2 inhibition, and illustrates how polyamides may interfere with DNA-binding proteins without actually contacting the

protein.

Other purely major-groove binding TFs, such as the basic-region leucine zipper (bZIP) protein GCN4, have been shown to co-occupy the DNA helix in the presence of polyamides.[67] Strategies employing polyamides functionalized with helix-distorting moieties have been successful at inhibiting such proteins. Polyamides with an attached Arg-Pro-Arg tripeptide can interfere with major-groove binding proteins by disrupting key phosphate contacts, distorting the DNA by charge neutralization, or sterically invading the major groove. An Arg-Pro-Arg polyamide conjugate successfully inhibited the binding of GCN4 to DNA,[67] and further optimization yielded a polyamide derivative with an alkyl diamine substituent that was 10-fold more potent.[68]

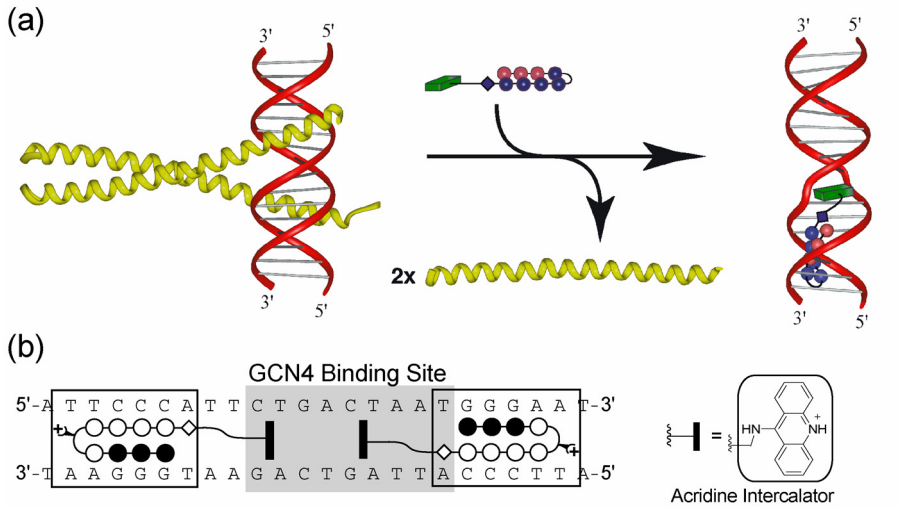


Figure 1.11. Model for allosteric inhibition of a protein-DNA complex by a polyamide-intercalator conjugate. (a) The GCN4 homodimer is displaced by the intercalating moiety of the polyamide conjugate. (b) Ball-and-stick model of the polyamide conjugate binding the target site (boxed) adjacent to the binding of the protein GCN4 (shaded). The structure of the acridine intercalator is shown at right. All other symbols are as defined in figure 1.4.

Polyamide-intercalator conjugates that distort the DNA at specific, targeted sequences by insertion of an intercalator have also proved to be highly potent inhibitors of major groove-binding proteins. Polyamides conjugated to the intercalator acridine

disrupt the DNA microstructure via unwinding, and were shown to significantly inhibit GCN4 binding when bound to sites adjacent to the GCN4 binding site, placing their acridine moieties into the GCN4 recognition element (Figure 1.11). Such molecules are promising candidates for site-selective inhibition of any DNA-binding protein.[69]

Polyamides can also upregulate transcription by inhibition of a repressor protein (derepression). For example, a hairpin polyamide was shown to block binding of the repressor IE86 to DNA, thereby upregulating transcription of the human cytomegalovirus MIEP.[70] A more complex case involves derepression of the integrated HIV-1 long terminal repeat (LTR). The human protein LSF binds in the promoter region at the LTR and recruits YY1, which then recruits histone deacetylases (HDACs). HDACs subsequently maintain LTR quiescence, which has been implicated in HIV latency, by maintaining a silent stock of pathogen. Three different live-cell models demonstrated that polyamides can inhibit LSF binding and increase expression of integrated HIV-1 promoter.[71] As with other systems, only polyamides matched to the correct protein binding site induced significant effects. Several existing drug treatments can reduce HIV-1 levels in the blood to below detectable amounts, yet the virus inevitably returns in infected patients. Derepression by inhibition of LSF-DNA binding may eventually allow HIV to be fully eradicated by drug treatments. This approach is particularly promising because LSF is a human protein, which could make the target less susceptible to resistance by HIV-1 mutations.

Recruitment of DNA-Binding Proteins

Polyamides have also been shown to affect DNA structure and function by recruiting proteins to specific, targeted sites. Most transcription factors have a DNA-binding domain and a separate domain that recruits the transcriptional machinery to that site (often called the activation domain). A polyamide can be thought of as an artificial DNA binding domain that can be linked to an activation domain. Such artificial transcription factors have been synthesized and evaluated in cell-free transcription assays.[72, 73] A hairpin polyamide tethered by a 36-atom straight-chain linker to the short (20-residue) peptide activation domain AH gives robust activation of transcription, with a size of only 4.2 kDa. Replacing the AH peptide with the shorter yet more potent activator VP2 (derived from the activator domain of the viral activator VP16) and reducing the linker from 36 atoms to eight provided a "minimal" polyamide-peptide conjugate, 3.2 kDa in size, which activated transcription slightly more effectively than the larger analogue (Figure 1.12a).[73] Since the linker length had been shown to influence activation efficiency, a set of molecules with rigid oligoproline linkers between the polyamide and the activation domain was synthesized.[74] The oligoproline linkers act as "molecule rulers," and optimal activation was observed with a Pro_{12} linker, about 36Å in length.

Many genes are influenced by multiple pathways and thus rely on the binding of several proteins. One example is the Hox (Homeobox) family of transcriptional regulators, which plays a vital role in the developmental fate of an organism. However, Hox proteins generally display poor affinity and sequence-specificity towards DNA. Instead, they are recruited to DNA by the strong, specific binding of members of the

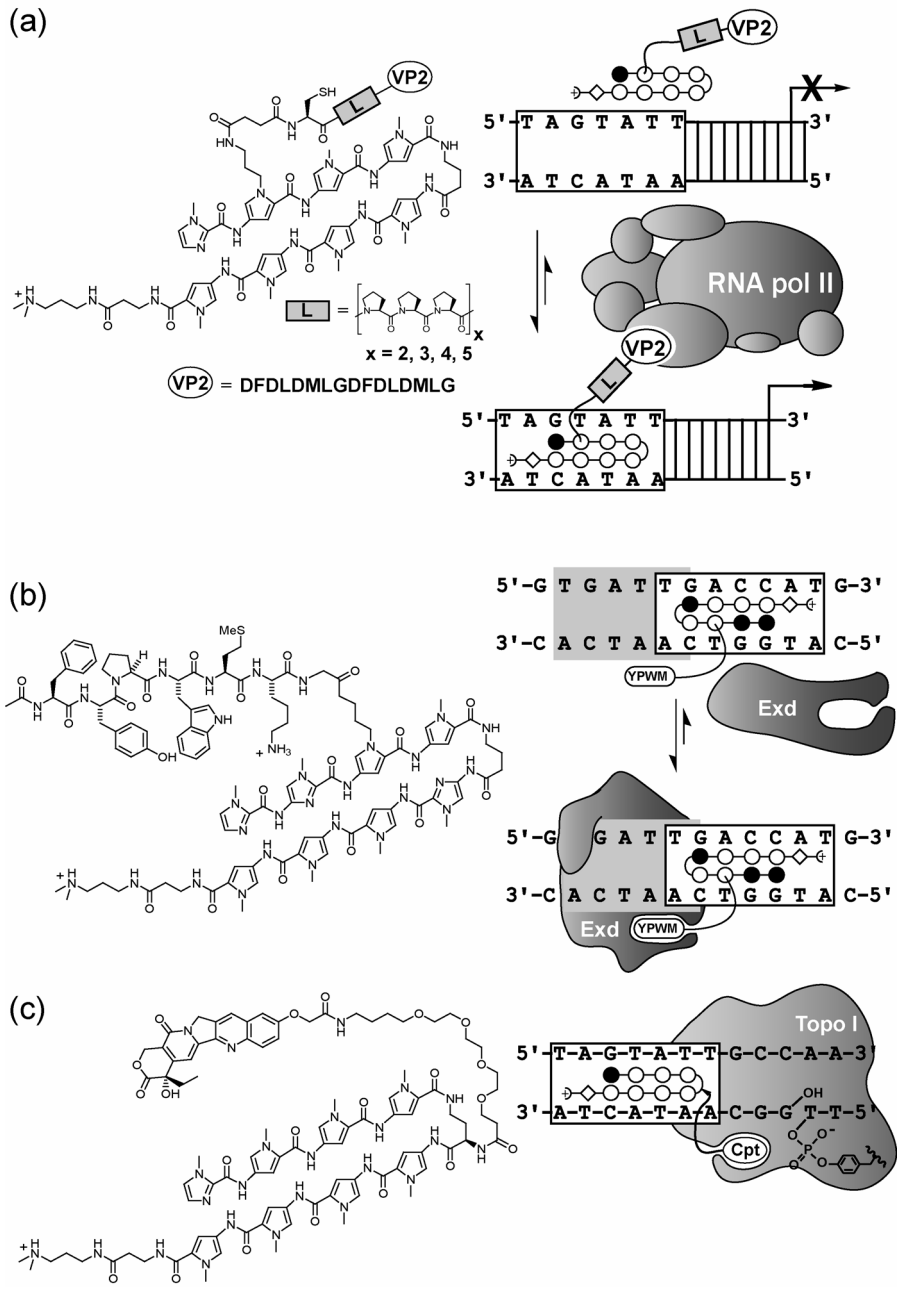


Figure 1.12. Recruitment of cellular proteins to DNA by polyamides. (a) Polyamides conjugated to the VP2 activation domain via a rigid poly-proline linker recruit the transcriptional machinery to a targeted site. (b) A polyamide functionalized with a short, YPWM peptide recruits Exd at nanomolar concentrations, changing the protein from a non-DNA-binding conformation (top) to one that binds the DNA•polyamide complex with high affinity (bottom). (c) Polyamide-camptothecin conjugates recruit topoisomerase I, inducing specific, targetable DNA strand breaks. Protein binding sites are shaded, and polyamide binding sites are boxed. All other symbols are labeled or defined in figure 1.4.

TALE (Three Amino Acid Loop Extension) class of homeodomain proteins. Recent crystal structures of one such ternary complex shows that the Hox protein Ultrabithorax

(Ubx) interacts with the *Drosophila* TALE protein extradenticle (Exd) via a short docking YPWM peptide.[75] A polyamide functionalized with this YPWM peptide successfully recruited Exd at nanomolar concentrations, outperforming the natural Ubx protein (Figure 1.12b).[76] This demonstrates that cooperative interactions among functionalized polyamides, DNA, and a protein can stabilize the formation of a ternary complex on a composite DNA site *in vitro*.

Polyamide camptothecin conjugates specifically recruited DNA topoisomerase I (Topo I) and induced single-strand cleavage.[77] Camptothecin is known to stabilize the cleavage complex formed between a tyrosine residue on Topo I and the 3'-phosphoryl end of the DNA backbone.[78] Using polyamide-camptothecin conjugates, this cleavage complex could be generated sequence-specifically at sites adjacent to the polyamide binding site (Figure 1.12c). Since camptothecin-Topo I–DNA complexes have been shown to arrest transcription elongation,[79] polyamide-camptothecin conjugates may function as sequence-specific transcription terminators.

Recognition of Nucleosomes.

In eukaryotic cells, DNA is tightly packaged by compaction into chromatin, and changes in chromatin structure can alter the accessibility of specific sequences and affect components of the molecular machinery in the nucleus. The fundamental repeating unit of chromatin is the nucleosome, comprising a 20–80 bp DNA linker region and the nucleosome core particle (NCP)– roughly two tight superhelical turns of DNA (147 bp in length) wrapped around a disk of eight histone proteins. The ability of DNA-binding proteins to recognize their cognate sites in chromatin is restricted by the structure and

dynamics of nucleosomal DNA, and by the translational and rotational positioning of the histone octamer. Using six different hairpin polyamides, it was shown that sites on nucleosomal DNA facing away from the histone octamer, or even partially facing the octamers, are fully accessible.[80] Remarkably, one section of 14 consecutive base pairs—more than a full turn of the DNA helix—was accessible for high affinity polyamide binding. The only positions very poorly bound by polyamides were sites near the amino-terminal tails of histone H3 or histone H4. Removal of either tail allowed polyamides to bind, suggesting that the structure of the DNA and perhaps its rotational position are strongly influenced by the N-terminal tails of histone H3 and H4.[80]

Subsequently, the structures of three of these polyamide-NCP complexes were determined by X ray crystallography.[81] The histone octamer is unaffected by polyamide binding, but the nucleosomal DNA undergoes significant structural changes at the ligand binding sites and the adjacent regions. Significantly, distortions in DNA twist can propagate over long distances without disrupting histone-DNA contacts, giving a potential mechanistic rationale for the role of twist diffusion in nucleosome translocation. Although the three polyamides display very similar affinities for their binding sites in the α-satellite nucleosome particle, only the relatively non-specific polyamide ImPyPyPy-γ-PyPyPyPy-β-Dp inhibits temperature-induced nucleosome translocation.[81] This may indicate that ligand positioning is critical, such that a single properly placed polyamide would effectively block translocation; or that the small effects of a single bound ligand can be amplified, such that a combination of several different polyamides would block translocation.

Although polyamides can block transcription by targeting promoter elements,

they do not affect transcription when bound in the RNA coding regions of DNA.[82] Presumably, the strand melting required for RNA polymerase progression disrupts the minor groove and displaces polyamides. 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.[82] The two molecules that blocked heat-induced nucleosomal translocation also blocked transcription by T7 RNA polymerase. Each of these polyamides binds with high affinity 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 ImPyPyPy-γ-PyPyPyPy-β-Dp in the crystal structure.[81] Although the precise mechanisms involved in nucleosome repositioning remain in question, it appears that, in some cases, DNA can "roll" over the histones, and certain polyamides can act as chocks to prevent the DNA from moving.

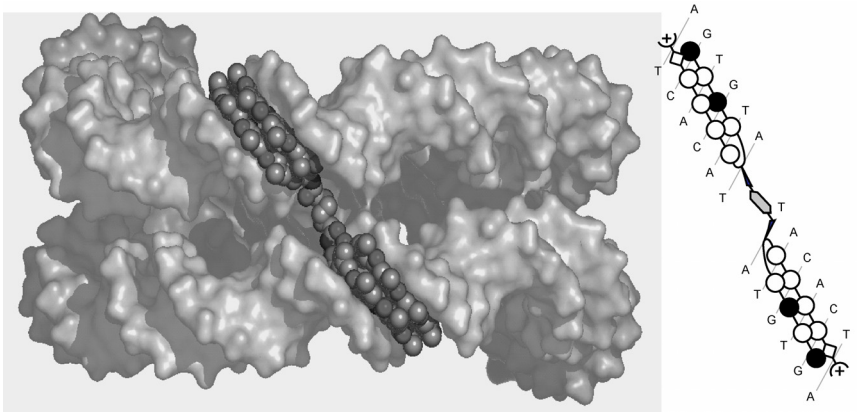


Figure 1.13. Left: Detail of the X ray crystal structure of the nucleosome core particle (NCP) with turn-to-turn dimer clamp **PW2** bound to a single supergroove. DNA helices run horizontally and are colored white. Polyamide bound in proximal supergroove is dark grey. At right is a schematic representation of binding, highlighting the DNA sequence.

One of the polyamide-NCP crystal structures showed a striking alignment of polyamides bound to sites almost 80 linear base pairs apart. Polyamides bound to these adjacent minor groove sites have their γ-turn moieties juxtaposed. Dimer polyamides

linked turn-to-turn bind to such "super groove" sites and serve as clamps, locking the DNA onto the nucleosome. (Figure 1.13).[83] This feature makes it possible to explore chromatin superhelix recognition. Such nucleosome targeting becomes relevant to biological systems, as the majority of cellular DNA is nucleosome-bound.

Nuclear Uptake.

DNA-binding polyamides can inhibit and influence a wide variety of protein-DNA interactions in solution, yet effectiveness in cell culture has proved to be dependent on cell type. A series of fluorescently labeled polyamides was prepared to analyze the intracellular distribution of these molecules in a panel of cell lines.[84] In cell types that had shown robust responses to polyamides, such as primary human T-cells, fluorescent polyamide-bodipy conjugates were observed to enter the nuclei of live cells.[84] However, in the majority of cell lines, polyamide-bodipy conjugates were excluded from the nucleus. Costaining with organelle-specific dyes indicates that polyamide-bodipy conjugates are often trapped in lysosomes and other cytoplasmic vesicles,[84] such that cells treated with polyamides can give a false nuclear signal upon fixing, even if they are washed extensively. Bashkin and coworkers have demonstrated that a polyamide-bodipy conjugate will traffic to the nucleus of a human cell line in the presence of verapamil, a p-glycoprotein inhibitor.[85]

Recently, a series of fluorescein-labeled polyamides were assayed for nuclear uptake against a panel of live mammalian cells.[86, 87] In some cases, small changes, such as the removal of a β-Ala residue at the C-terminus of a polyamide dramatically enhanced nuclear localization (Figure 1.14). Nuclear uptake of tested polyamide-fluorescein

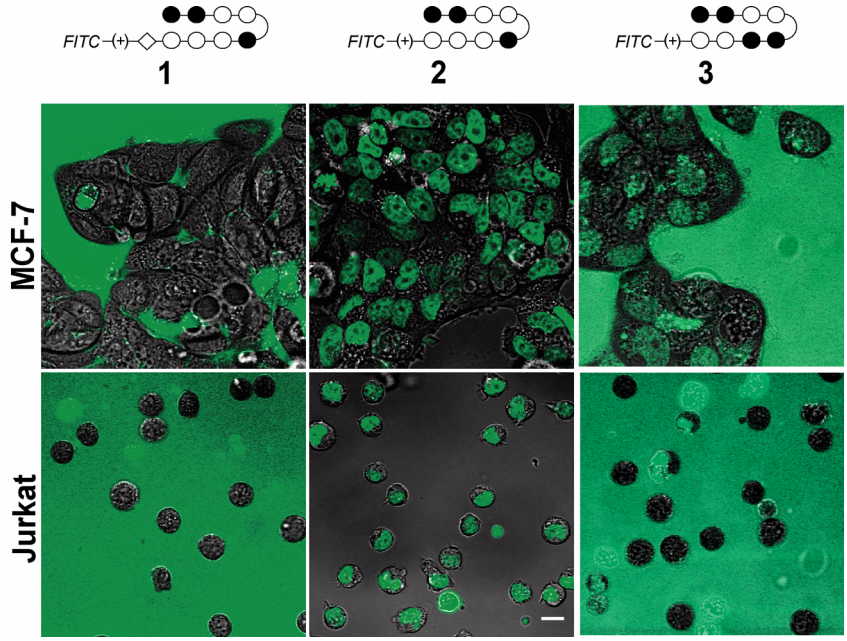


Figure 1.14. Nuclear uptake of hairpin polyamides. Representative confocal microscopy images of fluorescein-labeled polyamides in MCF-7 and Jurkat cells. Polyamide **1** exhibits poor uptake and is excluded from the nuclei of both cell types. Removal of the C-terminal *β*-Ala residue results in a polyamide (**2**) with excellent uptake properties, localizing to the nuclei of both cell types. Polyamide **3** differing from **2** by a single pyrrole to imidazole substitution, localizes to MCF-7 nuclei less strongly than does **2**, and **3** is completely excluded from Jurkat cells.

conjugates is an energy-dependent process. HeLa cells grown in energy inhibitory medium (supplemented with 2-deoxyglucose and sodium azide) displayed little to no discernable nuclear staining when treated with a fluorescein-labeled polyamide, while the same cells grown in normal medium showed clear nuclear staining. Washing of the inhibitory medium, and replacement with normal medium (supplemented with additional polyamide) resulted in nuclear staining.

While there are currently no general rules for cellular uptake of polyamides, determinants such as polyamide size, imidazole content, structure and attachment point of the fluorescent dye, and structure of the "tail" are important for nuclear localization. Each cell line possess a unique uptake profile such that choices of specific cell lines and compound architectures will be critical for future biological experiments.

Scope of This Work.

The work presented in this thesis is mainly focused towards using the target DNA to template the formation of higher-order polyamide structures with improved binding properties. As has been noted in this introduction, hairpin dimers, while targeting long sequences of DNA with high affinity and specificity, are unable to translocate to the nuclei of living cells, and thus cannot be used for experiments in organisms. We report our efforts towards forming these higher-order dimeric polyamides from smaller, potentially cell-permeable starting materials. Chapter 2 of this thesis is concerned with turn-to-tail templated reactions of duplex DNA. Chapter 3 of this thesis is concerned with turn-to-turn templated reactions both on duplex DNA and across NCP supergrooves. Chapter 4 of this thesis presents our efforts towards non-covalent interactions between polyamide strands bound in the minor groove.

This thesis also contains three appendices. While much of polyamide research has focused on biological applications such as gene regulation, polyamides possess amazing DNA-binding properties that make them well suited for use as molecular biology tools. Appendix I of this thesis details our efforts towards the use of polyamide-biotin conjugates for the capture and purification of DNA sequences from fragmented genomes. Appendix III of this thesis details our efforts towards the use of polyamides as specificity agents for structure determination of non-sequence-specific DNA enzymes.

Finally, Appendix II of this thesis examines the DNA-binding properties of tail-less polyamides synthesized on oxime resin.

References

1. Arcamone, F.; Nicolell, V; Penco, S.; Orezzi, P.; Pirelli, A., *Nature*, **1964,** *203*, 1064–1066.

2. Zimmer, C.; Wahnert, U., *Progress in Biophysics & Molecular Biology*, **1986,** *47*, 31–112.

3. Schuhman, E.; Haupt, I.; Thrum, H.; Taubenec.U; May, U., *Zeitschrift Fur Allgemeine Mikrobiologie*, **1974,** *14*, 321–327.

4. Ginsburg, H.; Nissani, E.; Krugliak, M.; Williamson, D. H., *Molecular and Biochemical Parasitology*, **1993,** *58*, 7–15.

5. Thrum, H.; Haupt, I.; Bradler, G.; Zimmer, C. G.; Reinert, K. E., *Antimicrobial and Antineoplastic Chemotherapy*. Czech. Med. Press: Prague, 1972; Vol. 1, p. 819.

6. Verini, M. A.; Ghione, M., *Chemotherapy*, **1964,** *9*, 144–148.

7. Mars, G.; Regoli, U., *Clin. Ter.*, **1968,** *30*, 573–580.

8. 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*, 1376–1380.

9. Pelton, J. G.; Wemmer, D. E., *Proceedings of the National Academy of Sciences of the United States of America*, **1989,** *86*, 5723–5727.

10. Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A., *Proceedings of the National Academy of Sciences of the United States of America*, **1987,** *84*, 8385–8389.

11. Dervan, P. B.; Edelson, B. S., *Current Opinion in Structural Biology*, **2003,** *13*, 284–299.

12. Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2001,** *9*, 2215–2235.

13. Wade, W. S.; Mrksich, M.; Dervan, P. B., *Journal of the American Chemical Society*, **1992,** *114*, 8783–8794.

14. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C., *Nature Structural Biology*, **1998,** *5*, 104–109.

15. Pilch, D. S.; Poklar, N.; Baird, E. E.; Dervan, P. B.; Breslauer, K. J., *Biochemistry*, **1999,** *38*, 2143–2151.

16. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Nature*, **1998,** *391*, 468–471.

17. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Science*, **1998,** *282*, 111–115.

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*, 557–567.

19. Wellenzohn, B.; Loferer, M. J.; Trieb, M.; Rauch, C.; Winger, R. H.; Mayer, E.; Liedl, K. R., *Journal of the American Chemical Society*, **2003,** *125*, 1088–1095.

20. Mrksich, M.; Parks, M. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1994,** *116*, 7983–7988.

21. Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Nature*, **1996,** *382*, 559–561.

22. deClairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E., *Journal of the American Chemical Society*, **1997,** *119*, 7909–7916.

23. White, S.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1997,** *119*, 8756–8765.

24. Swalley, S. E.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1999,** *121*, 1113–1120.

25. Hawkins, C. A.; de Clairac, R. P.; Dominey, R. N.; Baird, E. E.; White, S.; Dervan, P. B.; Wemmer, D. E., *Journal of the American Chemical Society*, **2000,** *122*, 5235–5243.

26. Herman, D. M.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1998,** *120*, 1382–1391.

27. Dervan, P. B.; Urbach, A. R., In *Essays in Contemporary Chemistry*, Quinkert, G.; Kisakurek, M. V., Eds. Verlag Helv. Chim. Acta: 2001; p. 327–339.

28. Urbach, A. R. 1:1 Motif for DNA REcognition by β–Alanine-Linked Polyamides. California Institute of Technology, Pasadena, 2002.

29. Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L., *Journal of the American Chemical Society*, **2002,** *124*, 2148–2152.

30. Kelly, J. J.; Baird, E. E.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America*, **1996,** *93*, 6981–6985.

31. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1998,** *120*, 6219–6226.

32. Wang, C. C. C.; Ellervik, U.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2001,** *9*, 653–657.

33. Herman, D. M.; Baird, E. E.; Dervan, P. B., *Chemistry-a European Journal*, **1999,** *5*, 975–983.

34. Kers, I.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2002,** *10*, 3339–3349.

35. Weyermann, P.; Dervan, P. B., *Journal of the American Chemical Society*, **2002,** *124*, 6872–6878.

36. Maeshima, K.; Janssen, S.; Laemmli, U. K., *Embo Journal*, **2001,** *20*, 3218–3228.

37. Greenberg, W. A.; Baird, E. E.; Dervan, P. B., *Chemistry-a European Journal*, **1998,** *4*, 796–805.

38. Olenyuk, B.; Jitianu, C.; Dervan, P. B., *Journal of the American Chemical Society*, **2003,** *125*, 4741–4751.

39. Heckel, A.; Dervan, P. B., *Chemistry-a European Journal*, **2003,** *9*, 3353–3366.

40. Hunter, C. A., *Journal of Molecular Biology*, **1993,** *230*, 1025–1054.

41. Urbach, A. R.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America*, **2001,** *98*, 4343–4348.

42. Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B., *Journal of Molecular Biology*, **2002,** *320*, 55–71.

43. Marques, M. A.; Doss, R. M.; Urbach, A. R.; Dervan, P. B., *Helvetica Chimica Acta*, **2002,** *85*, 4485–4517.

44. Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U. K., *Molecular Cell*, **2000,** *6*, 1013–1024.

45. Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society*, **1996,** *118*, 6141–6146.

46. Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Organic Letters*, **2001,** *3*, 1201–1203.

47. Fattori, D.; Kinzel, O.; Ingallinella, P.; Bianchi, E.; Pessi, A., *Bioorganic & Medicinal Chemistry Letters*, **2002,** *12*, 1143–1147.

48. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2002,** *10*, 2767–2774.

49. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2003,** *11*, 4333–4340.

50. Minehan, T. G.; Gottwald, K.; Dervan, P. B., *Helvetica Chimica Acta*, **2000,** *83*, 2197–2213.

51. Briehn, C. A.; Weyermann, P.; Dervan, P. B., *Chemistry-a European Journal*, **2003,** *9*, 2110–2122.

52. Renneberg, D.; Dervan, P. B., *Journal of the American Chemical Society*, **2003,** *125*, 5707–5716.

53. Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Nature*, **1997,** *387*, 202–205.

54. Darnell, J. E., *Nature Reviews Cancer*, **2002,** *2*, 740–749.

55. Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B.; Gottesfeld, J. M., *Journal of Molecular Biology*, **1997,** *274*, 439–445.

56. 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*, 12890–12895.

57. 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*, 1723–1733.

58. McBryant, S. J.; Baird, E. E.; Trauger, J. W.; Dervan, P. B.; Gottesfeld, J. M., *Journal of Molecular Biology*, **1999,** *286*, 973–981.

59. Winston, R. L.; Ehley, J. A.; Baird, E. E.; Dervan, P. B.; Gottesfeld, J. M., *Biochemistry*, **2000,** *39*, 9092–9098.

60. Lenzmeier, B. A.; Baird, E. E.; Dervan, P. B.; Nyborg, J. K., *Journal of Molecular Biology*, **1999,** *291*, 731–744.

61. Simon, H.; Kittler, L.; Baird, E.; Dervan, P.; Zimmer, C., *Febs Letters*, **2000,** *471*, 173–176.

62. Wurtz, N. R.; Pomerantz, J. L.; Baltimore, D.; Dervan, P. B., *Biochemistry*, **2002,** *41*, 7604–7609.

63. Yang, F.; Belitsky, J. M.; Villanueva, R. A.; Dervan, P. B.; Roth, M. J., *Biochemistry*, **2003,** *42*, 6249–6258.

64. Dickinson, L. A.; Trauger, J. W.; Baird, E. E.; Dervano, P. B.; Graves, B. J.; Gottesfeld, J. M., *Journal of Biological Chemistry*, **1999,** *274*, 12765–12773.

65. 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*, 24246–24254.

66. Schaal, T. D.; Mallet, W. G.; McMinn, D. L.; Nam, V. N.; Sopko, M. M.; John, S.; Parekh, B. S., *Nucleic Acids Research*, **2003,** *31*, 1282–1291.

67. Bremer, R. E.; Baird, E. E.; Dervan, P. B., *Chemistry & Biology*, **1998,** *5*, 119–133.

68. Bremer, R. E.; Wurtz, N. R.; Szewczyk, J. W.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2001,** *9*, 2093–2103.

69. Fechter, E. J.; Dervan, P. B., *Journal of the American Chemical Society*, **2003,** *125*, 8476–8485.

70. Dickinson, L. A.; Trauger, J. W.; Baird, E. E.; Ghazal, P.; Dervan, P. B.; Gottesfeld, J. M., *Biochemistry*, **1999,** *38*, 10801–10807.

71. Coull, J. J.; He, G. C.; Melander, C.; Rucker, V. C.; Dervan, P. B.; Margolis, D. M., *Journal of Virology*, **2002,** *76*, 12349–12354.

72. 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*, 3930–3935.

73. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., *Chemistry & Biology*, **2001,** *8*, 583–592.

74. Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., *Journal of the American Chemical Society*, **2002,** *124*, 13067–13071.

75. Passner, J. M.; Ryoo, H. D.; Shen, L. Y.; Mann, R. S.; Aggarwal, A. K., *Nature*, **1999,** *397*, 714–719.

76. 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*, 13322–13323.

77. Wang, C. C. C.; Dervan, P. B., *Journal of the American Chemical Society*, **2001,** *123*, 8657–8661.

78. Hsiang, Y. H.; Hertzberg, R.; Hecht, S.; Liu, L. F., *Journal of Biological Chemistry*, **1985,** *260*, 4873–4878.

79. Bendixen, C.; Thomsen, B.; Alsner, J.; Westergaard, O., *Biochemistry*, **1990,** *29*, 5613–5619.

80. Gottesfeld, J. M.; Melander, C.; Suto, R. K.; Raviol, H.; Luger, K.; Dervan, P. B., *Journal of Molecular Biology*, **2001,** *309*, 615–629.

81. 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*, 371–380.

82. Gottesfeld, J. M.; Belitsky, J. M.; Melander, C.; Dervan, P. B.; Luger, K., *Journal of Molecular Biology*, **2002,** *321*, 249–263.

83. Edayathumangalam, R. S.; Weyermann, P.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., *Proceedings of the National Academy of Sciences of the United States of America*, **2004,** *101*, 6864–6869.

84. Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B., *Bioorganic & Medicinal Chemistry*, **2002,** *10*, 3313–3318.

85. Crowley, K. S.; Phillion, D. P.; Woodard, S. S.; Schweitzer, B. A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J. K., *Bioorganic & Medicinal Chemistry Letters*, **2003,** *13*, 1565–1570.

86. 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*, 12063–12068.

87. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., *Nucleic Acids Research*, **2004,** *32*, 2802–2818.