Pergamon

Bioorganic & Medicinal Chemistry 10 (2002) 3339–3349

BIOORGANIC & MEDICINAL CHEMISTRY

Search for the Optimal Linker in Tandem Hairpin Polyamides

Inger Kers and Peter B. Dervan*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Received 11 February 2002; accepted 12 April 2002

Abstract—In order to target specific DNA sequences ≥ 10 base pairs in size by minor groove binding ligands, a search for the optimal linker in dimers of hairpin polyamides was initiated. Two series of tandem polyamides ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-L]^{HN}γ-PyPyPy-β-Dp (**1a**–**e**), where L represents a series of 4–8 carbon long aliphatic amino acid linkers, and ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-L]^{HN}γ-PyPyPy-β-Dp (**2a**–**e**), where L represents a series of 2–6 carbon long aliphatic amino acid linkers, were synthesized and characterized by quantitative DNase I footprinting. β, γ and Dp represents β-alanine, γ-aminobutyric acid, and 3-(dimethylamino)propylamine, respectively. It was found that the five-carbon 5-aminovaleric acid (δ), is suitable to span one base-pair (bp) of DNA when incorporated into a tandem polyamide. ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-δ]^{HN}γ-PyPyPy-β-Dp (**1b**) binds the 10 bp binding-site 5′-AGTGAAGTGA-3′ with equilibrium association constant $K_a = 3.2\times10^{10}$ M^{-1} and ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-δ]^{HN}γ-PyPyPy-β-Dp (**2d**) binds the 11 bp binding-site 5′-AGTGATAGTGA-3′ with $K_a = 9.7\times10^{9}$ M^{-1}. Tandem **1b** also bind the 11 bp site but with lower affinity affording a 15-fold specificity for the shorter binding site. Replacing a methylene group in the amino acid linker with an oxygen atom to form tandem polyamide ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-E]^{HN}γ-PyPyPy-β-Dp (**4**) where E represents the ether linker, resulted in that an 80-fold specificity for the 10 bp binding site over the 11 bp site.

Introduction

Small molecules with the ability to target predetermined sequences of DNA would be valuable tools in molecular biology and potentially in human medicine.[1–3] The natural products netropsin[4] and distamycin,[5] a dimer and a trimer, respectively, of *N*-methylpyrrole amino acid, have inspired the development of a class of molecules that bind in the minor groove of DNA with an affinity and specificity comparable with DNA binding proteins.[6,7] DNA sequence recognition of these polyamides, containing the three aromatic amino acids *N*-methylpyrrole (Py), *N*-methylimidazole (Im) and *N*-methyl-3-hydroxypyrrole (Hp), depend on the side-by-side amino acid pairings oriented N-$>$C with respect to the 5′-$>$3′ direction of the DNA duplex.[6,7] Antiparallel pairs Im/Py distinguish G·C from C·G and both of these from A·T/T·A base pairs. A Py/Py pair binds both A·T and T·A in preference to G·C/C·G.[6] The discrimination of T·A from A·T using Hp/Py pairs completes the four base pair (bp) code.[8–10] Connecting two antiparallel polyamide strands with γ-aminobutyric acid (γ) forms a hairpin structure and results in a $>$100-fold increase in affinity to a match-site relative to the unlinked dimers.[11–14]

Five contiguous aromatic ring pairings represent the upper limit for binding of a hairpin polyamide to a DNA duplex without an energetic penalty due to overcurving of the polyamide structure relative to the minor groove of DNA.[15,16] The use of eight-ring hairpin polyamides to modulate transcription has provided impetus for the design of polyamide motifs with a larger binding site.[17] Different approaches have been explored. Hairpin polyamides were extended by inserting β-alanine (β) pairs to relax the curvature of the ligand[18] and cooperative binding of extended hairpins have been used to target larger sites 10 or 12 bp in size.[19] Another approach to target larger DNA sites is to create covalent dimers of hairpin polyamides. A linked 'turn to tail' hairpin forms a tandem dimer. The feasibility of this approach was demonstrated for the tandem polyamide ImPyPy-(*R*)[ImPyPy-(R)^{H2N}γ-PyPyPy-δ]^{HN}γ-PyPyPy-β-Dp, formed by connecting two six-ring hairpins turn-to-tail via a 5-aminovaleric acid unit.[20] This molecule was shown to bind the designed 11-bp binding site 5′-TGTTATTGTTA-3′ with very high affinity ($K_a > 1\times10^{12}$ M^{-1}) and specificity. Unfortunately, attempts to apply

*Corresponding author. Tel.: +1-626-395-6002; fax: +1-626-683-8753; e-mail: dervan@caltech.edu

0968-0896/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0968-0896(02)00221-3

analogous tandem polyamide motifs to other 11-bp DNA sequences have been less successful in yielding ligands that bind the DNA match-site with very high affinity. This finding may be due to a lack of generality of 5-aminovaleric acid, the original linking unit, or an anomalous binding site size (11 bp) related to the microstructure of the DNA sequence studied in this case. For this reason we decided to undertake a systematic study on the effect on DNA binding by varying the length and composition of the linker connecting two six-ring hairpin polyamides into a tandem polyamide, with the aim to find a more general tandem hairpin polyamide motif. Tandem polyamide motifs were recently utilized by Laemmli and co-workers to specifically target vertebrate teleomers.[21] These polyamide dimers were formed by connecting two hairpins with 8-amino-3,6-dioxaoctanoic acid, a linker that is significantly longer than 5-aminovaleric acid used previously.[20]

We chose to study two key variables of the DNA:-tandem polyamide complex (i.e., increasing the length of the amino acid linker by incremental steps of one methylene unit or by varying the DNA binding-site from 10 to 11 bp by insertion of one bp at the linker position). These modifications leaves the basic motif for DNA–polyamide interactions unaltered while at the same time they should provide information about the required linker length relative to the binding site size. As in the previous study on tandem polyamides,[20] two hairpin polyamides are connected from the 2-amino-group on the turn of one hairpin to the C-terminus of the other hairpin, via an amino acid linker. The C-terminus of the tandem polyamide is the standard β-alanine 3-(dimethylamino)propylamine (Dp) tail (Fig. 1).

Two series of tandem polyamides were synthesized, both are based on the six-ring hairpin ImPyIm-(*R*)^{H_2N}γ-PyPyPy-β-Dp (**5**, Fig. 2). The first series ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-L]^{HN}γ-PyPyPy-β-Dp (**1a**–**e**) contains the straight-chain aliphatic amino acids (L): γ-aminobutyric acid, 5-aminovaleric acid (δ), 6-aminocaproic acid (ε), 7-aminoheptanoic acid (ζ) or 8-aminocaprylic acid (η), linkers ranging from four to eight carbons in size (Fig. 2). These linkers were chosen based on previous findings that the five carbon 5-aminovaleric acid linker was optimal for forming a tandem polyamide targeting a 11 bp binding-site.[20] A second series of tandem polyamides ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPyIm-L]^{HN}γ-PyPyPy-β-Dp (**2a**–**e**), containing a shorter amino acid and an unpaired Im as the linking motif, was also prepared (Fig. 2). The amino acid linkers (L) used for this series were glycine (G), β-alanine, γ-aminobutyric acid, 5-aminovaleric acid and 6-aminocaproic acid (i.e., their length range from two to six carbons). This second series containing an unpaired Im may afford a new class of tandem polyamides with modified characteristics and may also allow a G/C or C/G base-pair in the DNA binding-site corresponding to the linker region of the tandem polyamide. The DNA binding properties of these molecules (**1a**–**e** and **2a**–**e**) were characterized using quantitative DNase I footprint titrations on a DNA fragment from pIK2 containing

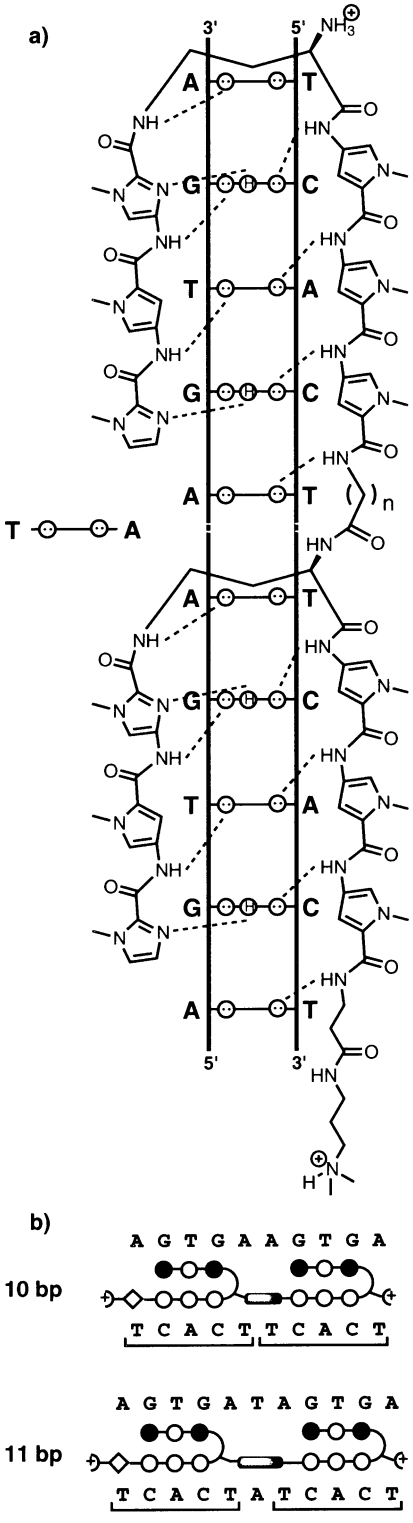

Figure 1. (a) Model of the polyamide/DNA complex formed between tandem polyamides **1a**–**e** and the 10 (or 11) bp match site 5′-AGTGA(T)AGTGA-3′. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. (b) A schematic ball and stick model of tandem polyamides **1a**–**e** to 10 or 11 bp DNA binding sites. Im and Py rings are represented as shaded and unshaded circles, respectively. The β-residue is illustrated by a diamond and the aminoacid linkers by a rectangle.

the 10 and 11 bp binding-sites 5′-AGTGAAGTGA-3′ and 5′-AGTGATAGTGA-3′ (Fig. 4). In attempts to address the generality of the optimal tandem polyamide motif ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-δ]^{HN}γ-PyPyPy-β-Dp (**1b**) was also characterized on a new DNA sequence from pIK3 with 10 and 11 bp binding-sites 5′-AGTGTTGTGA-3′ and 5′-AGTGTTTGTGA-3′. To address the issue of generality the tandem polyamide sequence was changed to ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyImPy-δ]^{HN}γ-PyImPy-β-Dp (**3**) with an Py to Im replacement (Fig. 2), derived from six-ring hairpin polyamide ImPyIm-(*R*)^{H_2N}γ-PyImPy-β-Dp (**6**), and characterized on a DNA fragment from pIK4 with the 10 and 11 bp binding-sites 5′-AGCGTTGCGA-3′ and 5′-AGCGTTTGCGA-3′. Tandem polyamide **3** also functions as a control for tandem polyamide **1b** in studies on mismatch binding. To analyze the effect of chemical composition of the linker in the tandem polyamide, 2-(2-aminoethoxy)acetic acid (E) was incorporated as a linker into a dimer producing the ether analogue ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-E]^{HN}γ-PyPyPy-β-Dp (**3**, Fig. 2) of tandem **1b**. Binding characteristics of **4** were studied on the DNA restriction fragment from pIK2.

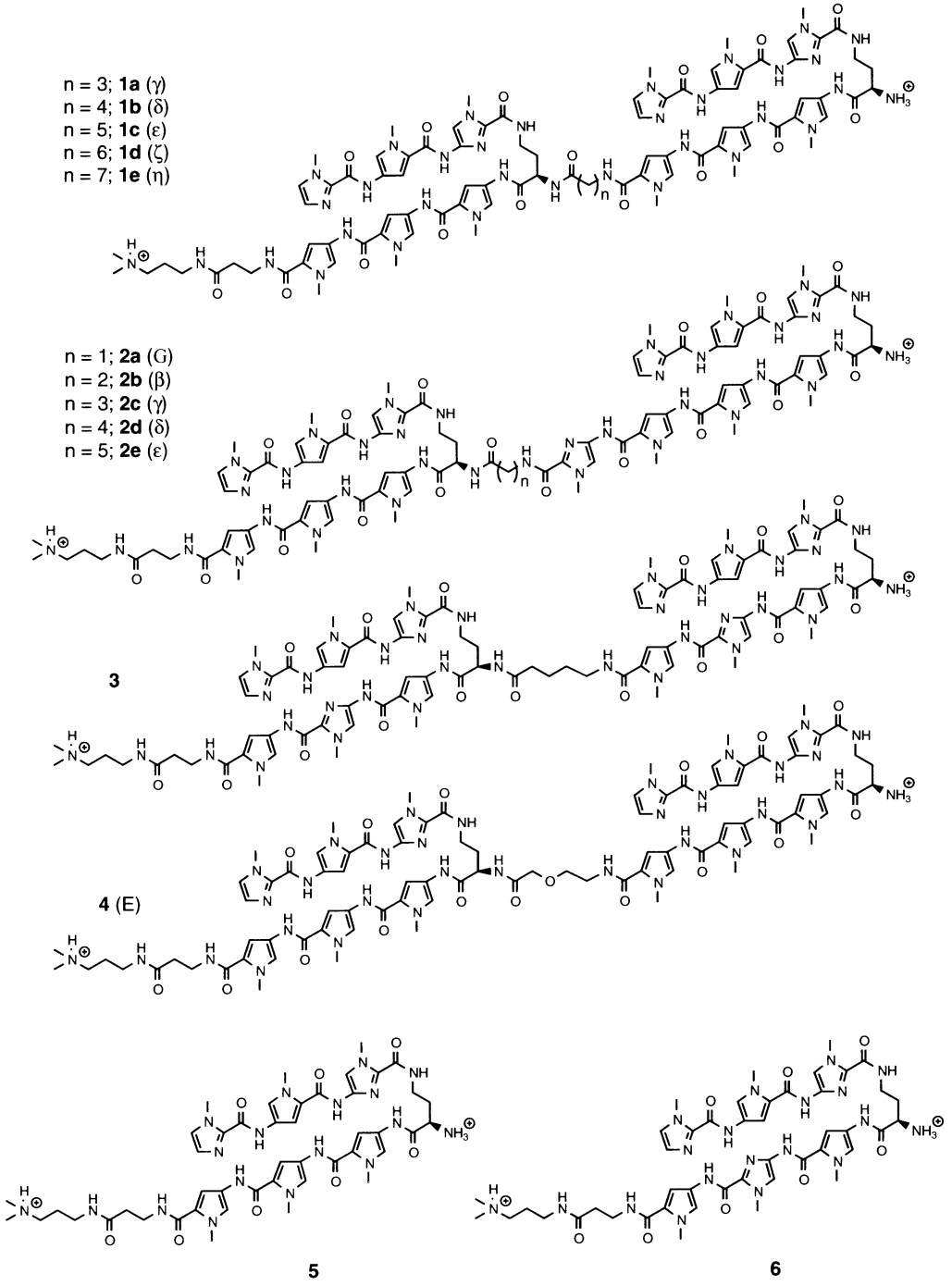

Figure 2. Structure of tandem hairpin polyamides **1a**–**e**, **2a**–**e**, **3**, and **4** and six-ring hairpin polyamides **5** and **6**.

Results

Syntheses

Tandem polyamides **1–4** were synthesized by solid-phase methods in 11 steps starting with β-alanine substituted Pam-resin using DCC/HOBt activated Boc-protected amino acid monomers or dimers according to previously reported procedures[20,22] (Fig. 3). The polyamides were cleaved from the resin by aminolysis with 3-(dimethylamino)propylamine and purified by reversed-phase HPLC to provide polyamides **1–4**. Two six-ring hairpin polyamides were also prepared by cleaving a sample of the resin with 3-(dimethylamino)-propylamine, after the first five couplings were performed, affording polyamides **5** and **6** (Fig. 2) after HPLC purification. Purity and identity of the polyamides were confirmed by analytical HPLC and MALDI-TOF MS.

DNase I footprint titration

Quantitative DNase I footprint titration experiments[23–25] (10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, 22 °C) were performed (Figs. 5–7) on a 3′-^{32}P end labeled 268 bp *Eco*RI/*Pvu*II restriction fragment from plasmids pIK2, pIK3 and pIK4 (Fig. 4) to determine the equilibrium association constants (Tables 1–5) for polyamides **1–6** on the designed binding-sites by calculating a fractional saturation value at the site, for each polyamide concentration, and fitting the data to a modified form of the Hill equation.

Aliphatic linked hairpin dimers

For the first series of tandem polyamides ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-L]^{HN}γ-PyPyPy-β-Dp (Fig. 2, **1a–e**), containing aliphatic amino acid linkers of varying length, a clear dependence on linker length for the binding of the polyamides to the designed 10 bp 5′-AGTGAAGTGA-3′ and 11 bp 5′-AGTGATAGTGA-3′ sites (Table 1) in the restriction fragment from pIK2 (Fig. 5) could be observed. All polyamides **1a–e** bound the 10 bp binding site, polyamides **1a** and **1b** containing the shortest linkers (γ and δ respectively) have slightly higher affinity ($K_a = 3.2\times10^{10}$ M^{-1}) compared to **1c**, **1d**, and **1e** ($K_a = 1.3\times10^{10}$ M^{-1}, $K_a = 1.8\times10^{10}$ M^{-1} and $K_a = 1.3\times10^{10}$ M^{-1}) with the longer linkers ε, ζ and η (Table 1). The size of the footprints indicates that the polyamides cover the whole 10 bp binding site. For the 11 bp binding site a more pronounced difference in DNA-binding could be observed. Tandem polyamide **1a** and **1b** bind the 11 bp binding site with significantly lower binding affinity ($K_a = 1.8\times10^{9}$ M^{-1} and $K_a = 2.1\times10^{9}$ M^{-1})

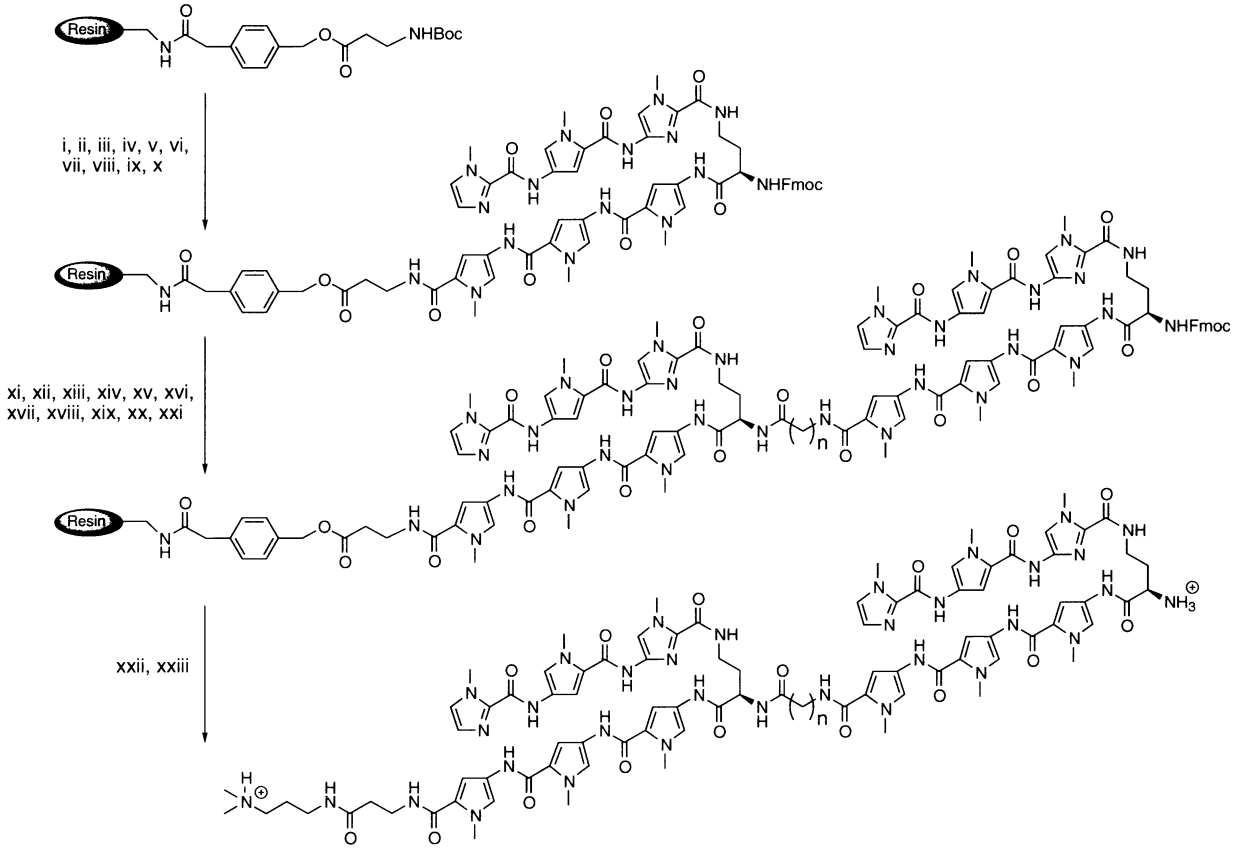

Figure 3. (i) 80% TFA/DCM; (ii) Boc-Py-Py-OH, DIEA, HOBt, DCC; (iii) 80% TFA/DCM; (iv) Boc-Py-OBt, DIEA; (v) 80% TFA/DCM; (vi) Fmoc-α-Boc-γ-diaminobutyric acid, HBTU, DIEA; (vii) 80% TFA/DCM; (viii) Boc-Py-Im-OH, DIEA, HOBt, DCC; (ix) 80% TFA/DCM; (x) Im-OH, DIEA, HOBt, DCC; (xi) 80% piperidine/DMF; (xii) Boc-amino acid linker; (xiii) 80% TFA/DCM; (xiv) Boc-Py-Py-OH, DIEA, HOBt, DCC; (xv) 80% TFA/DCM; (xvi) Boc-Py-OBt, DIEA; (xvii) Fmoc-α-Boc-γ-diaminobutyric acid, HBTU, DIEA; (xviii) 80% TFA/DCM; (xix) Boc-Py-Im-OH, DIEA, HOBt, DCC; (xx) 80% TFA/DCM; (xxi) Im-OH, DIEA, HOBt, DCC; (xxii) 80% Piperidine/DMF; (xxiii) 3-(dimethylamino)propylamine.

than the 10 bp one. This affinity is close to that of the 6-ring hairpin ImPyIm-(*R*)^{H_2N}γ-PyPyPy-β-Dp (**5**) ($K_a = 1.2\times10^9$ M^{-1}) on the binding-site 5′-AGTGA-3′. Also, the size of the footprint shows that the tandem polyamides **1a** and **1b** are unable to cover the whole 11 bp binding-site upon binding to DNA (Fig. 5). The binding of **1c** to this site could not be characterized due to a strong mismatch binding to the 3′-end that is partially overlapping with the 11 bp site. Increasing the linker length further permits a more optimal binding of the polyamides **1d** ($K_a = 4.7\times10^9$ M^{-1}) and **1e** ($K_a = 1.4\times10^{10}$ M^{-1}) to the 11 bp site. The footprint of **1e** shows complete blocking of the 11 bp site to cleavage by DNase I. Generally, the tandem polyamides with longer linkers show a higher degree of non-specific DNA-binding. Tandem polyamide **1b** binds with the highest specificity in the series, to the designed 10 bp binding-site on the DNA fragment.

Tandem polyamide **1b** was also characterized (Table 4) on a restriction fragment from pIK3 with the 10 bp 5′-AGTGTTGTGA-3′ and 11 bp 5′-AGTGTTTGTGA-3′ sites, which contain T(T)T as central base-pairs instead of the central A(T)A of pIK2 (Fig. 6). Tandem polyamide ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyImPy-δ]^{HN}γ-PyImPy-β-Dp (**3**), with the replacement of one of the Py to Im in each hair-pin compared to tandem polyamide **1b**, was characterized on a restriction fragment from pIK4 that contains two designed binding-sites, 10 bp 5′-AGCGTTGCGA-3′ and 11 bp 5′-AGCGTTTGCGA-3′, in analogy to the initial studies (Fig. 6). DNase I footprint studies of **1b** on pIK4 and of **3** on pIK3 allowed studies on the mismatch binding of those compounds. **1b** bind to the 10 bp ($K_a = 4.7\times10^{10}$ M^{-1}) and the 11 bp ($K_a = 2.4\times10^9$ M^{-1}) binding-sites in pIK3 with affinities similar as for pIK2. **1b** bind with lower affinity ($K_a = 1.3\times10^9$ M^{-1}) the 10 bp site in pIK4, which contain one base-pair mismatch per hairpin module. Binding to the 11 bp mismatch site in pIK4 could not be observed for **1b** in the concentration range used ($K_a < 1\times10^8$ M^{-1}). **3** shows the same binding tendencies as **1b** but binds with lower affinity compared to **1b**, both to the 10 and 11 bp match-sites in pIK4 ($K_a = 8.3\times10^8$ M^{-1} and $K_a \sim 2\times10^7$ M^{-1}, respectively) and the 10 bp mismatch-site in pIK3 ($K_a \sim 2\times10^7$ M^{-1}). No binding for **3** at the 11 bp mismatch binding-site of pIK3 ($K_a < 5\times10^6$ M^{-1}) was observed. Six-ring hairpin **6** with the same polyamide sequence as tandem polyamide **3** was characterized on a match-site, 5′-TGCGT-3′, ($K_a = 5.2\times10^8$ M^{-1}) (Table 5).

Aliphatic-imidazole linked hairpin dimers

The general trend for the binding of polyamides ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPyIm-L]^{HN}γ-PyPyPy-β-Dp **2a**–**e** to the restriction fragment from pIK2 is similar to that of **1a**–**e**, but the two series show a somewhat different behavior. Dimer **2a** containing the shortest linker G was the optimal binder in the series ($K_a = 1.0\times10^{10}$ M^{-1}) for the 10 bp binding-site. The binding constants decrease in the order **2b**, **2c**, and **2e** with linkers b, g and e respectively ($K_a = 3.4\times10^9$ M^{-1}, $K_a = 2.5\times10^9$ M^{-1} and $K_a = 1.8\times10^9$ M^{-1}). Dimer **2d** ($K_a = 5.0\times10^9$ M^{-1}) with linker d is an exception from this trend. Tandem polyamides **2a** ($K_a = 8.7\times10^8$ M^{-1}) and **2b** ($K_a \sim 4\times10^8$ M^{-1}) containing the shortest linkers cannot bind to the whole 11 bp binding-site. Tandem **2c** ($K_a = 6.3\times10^8$ M^{-1}) and **2d** ($K_a = 9.7\times10^9$ M^{-1}) both manage to protect the full binding-site from DNase I cleavage, with **2d** having the highest affinity to the 11 bp site in the series. A drop in affinity and lower efficiency in the protection against DNase I cleavage is observed when increasing the linker length with one methylene unit (**2e**, $K_a = 7.5\times10^9$ M^{-1}).

Ether linked hairpin dimers

The ether modified tandem polyamide ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-E]^{HN}γ-PyPyPy-β-Dp (**4**) was characterized on pIK2 and was found to bind the 10 bp binding-site with full protection and similar affinity ($K_a = 3.5\times10^{10}$ M^{-1}) as 1b (Fig. 7). Tandem polyamide 4, analogously to **1b**, did not cover the whole 11 bp binding-site but surprisingly did bind the 11 bp binding-site with lower affinity ($K_a = 4.4\times10^8$ M^{-1}) than **1b**.

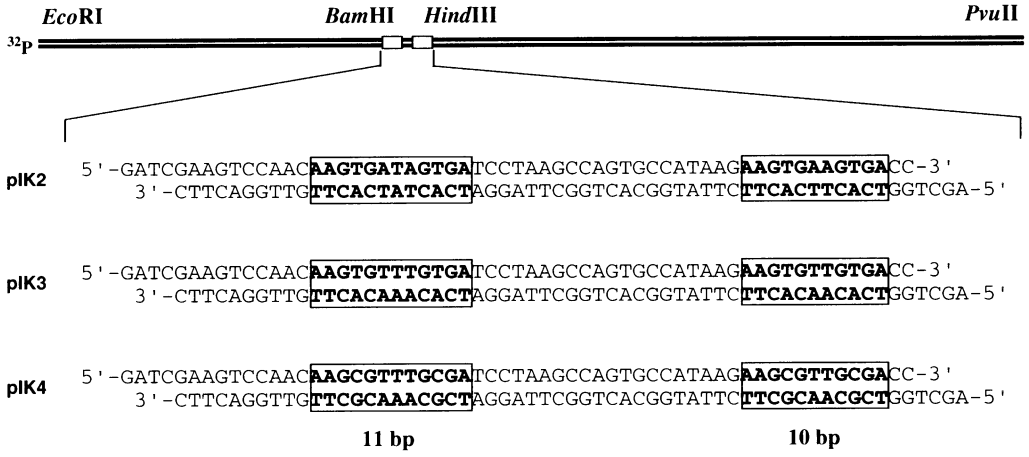

Figure 4. Illustration of the *Eco*RI/*Pvu*II restriction fragment of pIK2, pIK3, and pIK4 plasmids with the *Bam*HI/*Hind*III insertion site indicated. Each synthesized insert contain two binding sites, 10 respectively 11 bp, marked with boxes.

Discussion

Two series of tandem polyamides, with either an aliphatic amino acid linker (**1a**–**e**) or a short amino acid linker together with an unpaired Im (**2a**–**e**), were synthesized and characterized by means of DNase I footprint titration, in order to correlate the proper linker length with selective targeting of either a 10 or 11 bp binding-site. These studies showed that a five-carbon linker, when connecting two six-ring hairpin polyamides, has optimal length to span one base-pair in the minor groove of DNA. This conclusion could be drawn from the finding that polyamide ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-δ]^{HN}γ-PyPyPy-β-Dp (**1b**) containing a 5-aminovaleric acid linker was found to be the *optimal ligand for the 10 bp binding* site 5′-AGTGAAGTGA-3′

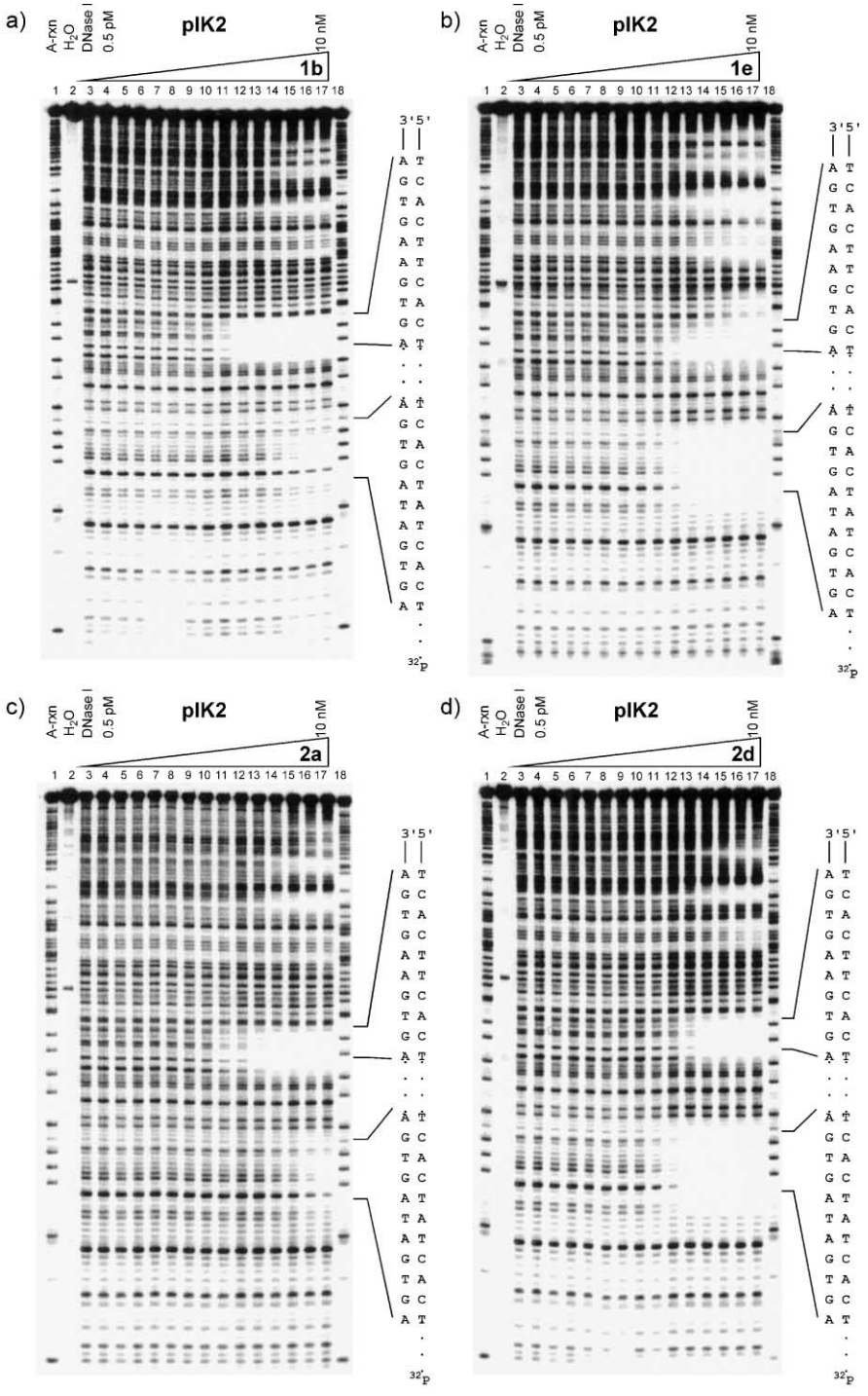

Figure 5. Quantitative DNase I footprint titration experiment with tandem polyamides (a) **1b**, (b) **1e**, (c) **2a** and (d) **2d** on the 3′-end labeled 268 bp restriction fragment of pIK2: lanes 1 and 18, A reaction; lane 2, intact DNA; lane 3, DNase I standard; lanes 4–17, DNase I digestion products in the presence of 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500 pM, 1, 2, 5, 10 nM polyamide respectively. The 10 bp (high) and the 11 bp (low) binding-sites were analyzed and are shown on the right side of the autoradiogram. All rections contain 15 kcpm restriction fragment, 10 mM Tris–HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$.

(Table 1). This was further shown for ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPyIm-δ]^{HN}γ-PyPyPy-β-Dp (**2d**) with δ-Im linker which had the highest affinity for the 11 bp binding site 5′-AGTGATAGTGA-3′ out of polyamides **2a**–**e**. For dimer **2d** the unpaired Im covers 1 bp, leaving 1 bp to bridge for the 5-aminovaleric acid linker at the 11 bp site. Tandem polyamide **1b** shows good selectivity (15-fold) for the 10 bp binding site over the 11 bp binding site. Dimer **2d**, on the other hand, is modestly selective for the 11 bp binding site (2-fold). When targeting the 11 bp binding-site using only an aliphatic amino acid linker to span the intervening two bps it is necessary to use a linker eight carbons in length. This tandem polyamide ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPy-η]^{HN}γ-PyPyPy-β-Dp (**1e**) shows no selectivity for the 11 bp site over the 10 bp site, both sites are bound with the same affinity. By using a two carbon linker together with an unpaired imidazole ImPyIm-

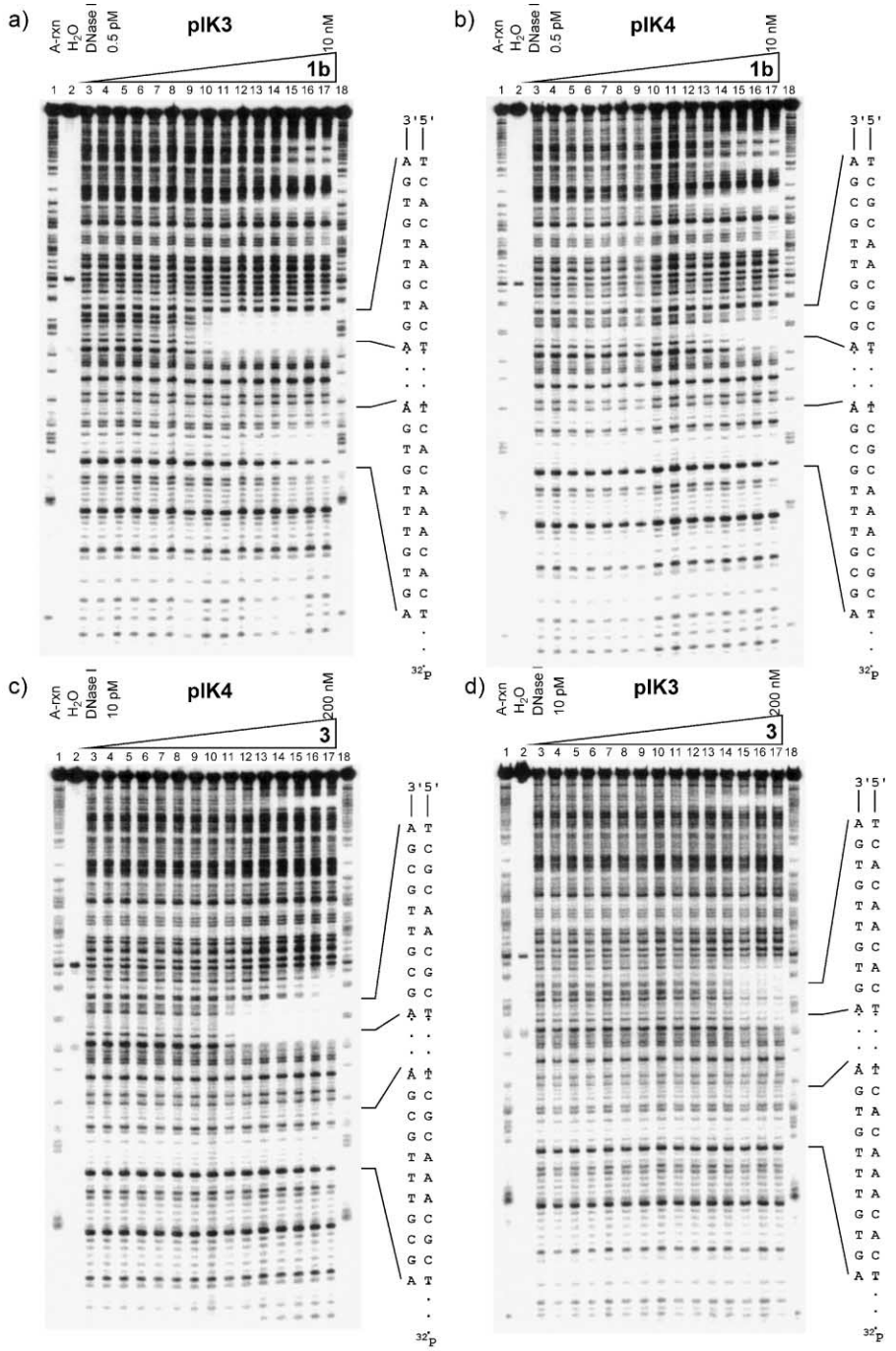

Figure 6. Quantitative DNase I footprint titration experiment with tandem polyamide **1b** on the 3′-end labeled 268 bp restriction fragment of (a) pIK3 and (b) pIK4 and tandem polyamide **3** on on the 3′-end labeled 268 bp restriction fragment of (c) pIK4 and (d) pIK3: lane 1 and 18, A reaction; lane 2, intact DNA; lane 3, DNase I standard; lanes 4–17, DNase I digestion products in the presence of 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500 pM, 1, 2, 5, 10 nM polyamide, respectively, for (a) and (b) and 10, 20, 50, 100, 200, 500 pM, 1, 2, 5, 10, 20, 50, 100, 200 nM polyamide, respectively, for (c) and (d). The 10 bp (high) and the 11 bp (low) binding sites were analyzed and are shown on the right side of the autoradiogram. All rections contain 15 kcpm restriction fragment, 10 mM Tris–HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$.

(*R*)[ImPyIm-(R)^{H_2N}γ-PyPyPyIm-G]^{HN}γ-PyPyPy-β-Dp (**2a**) it is also possible to target the 10 bp site with selectivity (11-fold) but with lower affinity (3-fold) than when using **1b** (Table 2). A drop in affinity for the 10 bp site, compared to **2a**, was observed for the molecules containing an aliphatic amino acid (longer than two carbons)-Im linker (**2b**–**e**). This is probably due to a limited space available for the linker between the two hairpins in the minor groove of the DNA. This observation is not true for the series of tandem polyamides **1a**–**e** when binding to the 10 bp binding-site, only a modest drop in affinity is observed for the molecules containing longer linkers.

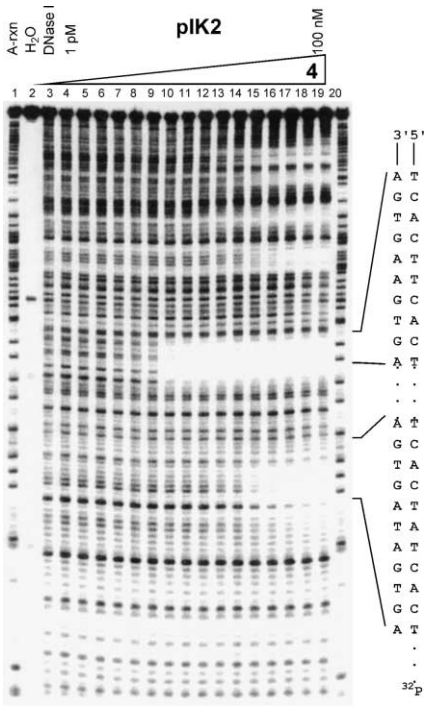

Figure 7. Quantitative DNase I footprint titration experiment with tandem polyamide **4** on the 3′-end labeled 268 bp restriction fragment of pIK2: lane 1 and 20, A reaction; lane 2, intact DNA; lane 3, DNase I standard; lanes 4–19, DNase I digestion products in the presence of 1, 2, 5, 10, 20, 50, 100, 200, 500 pM, 1, 2, 5, 10, 20, 50, 100 nM polyamide, respectively. The 10 bp (high) and the 11 bp (low) binding sites were analyzed and are shown on the right side of the autoradiogram. All reactions contain 15 kcpm.

The question of *sequence generality* when using 5-aminovaleric acid linked tandem polyamides, for targeting a 10 bp DNA binding-site was also addressed. Two modifications of the system was investigated. The DNA sequence covered by the linker part of the tandem polyamide was changed and the DNA sequence where the hairpin moiety binds was altered, which required the use of a new tandem polyamide with a different sequence. Tandem polyamide ImPyIm-(*R*)[ImPyIm-(R)^{H_2N}γ-PyImPy-δ]^{HN}γ-PyImPy-β-Dp (**3**), with an Py to Im replacement and containing 5-aminovaleric acid, showed a preference for the 10 bp binding-site 5′-AGCGTTGCGA-3′ over a 11 bp site, further supporting the finding that a five-carbon linker is suited to span 1 bp in the minor groove of DNA. DNase I footprint titration of dimer **1b** on a new DNA sequence with the 10 bp 5′-AGTGAAGTGA-3′ and the 11 bp 5′-AGTGATAGTGA-3′ binding sites showed similar binding affinity as previously, indicating a tolerance for different DNA sequences.

The key finding in this study is that a 5-aminovaleric acid linker in a tandem six-ring/six-ring polyamide (**1b**, **2d**, or **3**) optimally binds 10 bp, not 11 bp as reported in the first case study. Perhaps this difference can be explained by a variation in the DNA microstructure of the DNA sequences used in the first study.[26–32] Based on this more comprehensive data base, we believe that we have found a general solution for using tandem

Table 1. Equilibrium association constants[a] (M^{-1}) for aliphatic linked tandem polyamides **1a**–**1e** on a restriction fragment from pIK2

Polyamide	10 bp 5′-gaAGTGAAGTGAcc-3′	11 bp 5′-caAGTGATAGTGAtc-3′	Specificity, K_A(10 bp)/K_A(11 bp)
1a	3.2×10^{10}	1.8×10^{9}	18
1b	3.2×10^{10}	2.1×10^{9}	15
1c	1.3×10^{10}	N.A.[b]	–
1d	1.8×10^{10}	4.7×10^{9}	4
1e	1.3×10^{10}	1.4×10^{10}	1

[a]The reported equilibrium association constants are the mean values obtained from three DNase I footprint titration experiments. Assays were carried out in the presence of 10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22 °C.
[b]Equilibrium association constant could not be measured due to an overlapping mismatch binding site.

Table 2. Equilibrium association constants[a] for tandem polyamides **2a**–**2e** with aliphatic amino acid-imidazole linker on a restriction fragment from pIK2

Polyamide	10 bp 5′-gaAGTGAAGTGAcc-3′	11 bp 5′-caAGTGATAGTGAtc-3′	Specificity, K_A(10 bp)/K_A(11 bp)
2a	1.0×10^{10}	8.7×10^{8}	11
2b	3.4×10^{9}	$\sim4\times10^{8}$	9
2c	2.5×10^{9}	6.3×10^{9}	0.4
2d	5.0×10^{9}	9.7×10^{9}	0.5
2e	1.8×10^{9}	7.5×10^{9}	0.2

[a]The reported equilibrium association constants are the mean values obtained from three DNase I footprint titration experiments. Assays were carried out in the presence of 10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22 °C.

polyamides to target 10 (and 11) bp binding sites. At the 10 bp site tandems **1b** and **3** shows a preference for the match over the double mismatch in the range of 36- and 40-fold, respectively.

Comparison of tandem polyamide **3**, which contain three imidazole rings per hairpin module, with tandem polyamide **1b**, containing two imidazole rings in each hairpin, reveals for **3** a lower enhancement in binding affinity when forming the tandem polyamide from the hairpin polyamide. Most likely, the precise placement of a polyamide in the minor groove of DNA is highly dependent on the correct positioning of the N-3 of each imidazole ring to the exocyclic amino group of G. A higher imidazole content might put further constraints on the structure of a tandem polyamide resulting in lower energetic gain for dimerization. Finally, the chemical composition of the linker unit has the potential to affect the binding of a tandem polyamide to DNA. One methylene unit in the 5-aminovaleric acid was replaced with an oxygen atom to form tandem polyamide ImPyIm-(*R*)[ImPyIm-(*R*)^{H_2N}γ-PyPyPy-E]^{HN}γ-PyPyPy-β-Dp (**4**). This replacement was expected to conserve the approximate length of the aliphatic linker but may change the binding properties of the DNA ligand and increase the water solubility. Polyamide **4** containing an ether linker shows a decreased affinity compared to **1b** only for the 11 bp binding site, but not the 10 bp site, an

Table 3. Equilibrium association constants[a] for ether linked tandem polyamide **4** on a restriction fragment from pIK2

Polyamide	10 bp 5′-gaAGTGAAGTGAcc-3′	11 bp 5′-caAGTGATAGTGAtc-3′	Specificity, K_A(10 bp)/K_A(11 bp)
4	3.5×10^{10}	4.4×10^{8}	80

[a]The reported equilibrium association constants are the mean values obtained from three DNase I footprint titration experiments. Assays were carried out in the presence of 10mM Tris–HCl, 10mM KCl, 10 mM $MgCl_2$, and 5mM $CaCl_2$ at pH 7.0 and 22 °C.

Table 4. Match and mismatch equilibrium association constants[a] for **1b** and **3**

Polyamide and plasmid	10 bp	11 bp	Specificity, K_A(10 bp)/K_A(11 bp)
1b pIK3	C A **A G T G T T G T G A** T C G T **T C A C A A C A C T** A G 4.7×10^{10}	C A **A G T G T T T G T G A** T C G T **T C A C A A A C A C T** A G 2.4×10^{9}	20
1b pIK4	C A **A G C G T T G C G A** T C G T **T C G C A A C G C T** A G 1.3×10^{9}	C A **A G C G T T T G C G A** T C G T **T C G C A A A C G C T** A G $<1\times10^{8}$	>13
3 pIK4	C A **A G T G T T G T G A** T C G T **T C A C A A C A C T** A G 8.3×10^{8}	C A **A G T G T T T G T G A** T C G T **T C A C A A A C A C T** A G $\sim2\times10^{7}$	42
3 pIK3	C A **A G T G T T G T G A** T C G T **T C A C A A C A C T** A G $\sim2\times10^{7}$	C A **A G T G T T G T G A** T C G T **T C A C A A C A C T** A G $<5\times10^{6}$	>4

[a]The reported equilibrium association constants are the mean values obtained from three DNase I footprint titration experiments. Assays were carried out in the presence of 10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22 °C.

80-fold preference for the 10 bp binding-site over 11 bp (Table 3). However, a general trend emerging in this study is that tandem polyamides with longer linker motifs show a larger tendency to mismatch binding and it may be important not to choose a linker longer than necessary.

Conclusion

The present study was aimed at finding a general motif for tandem hairpin polyamides linked 'turn-to-tail'. The studies on the DNA binding affinity of tandem polyamides with varying linker length and composition, and different DNA and polyamide sequences presented herein have achieved this. Using 5-aminovaleric acid or 2-(2-aminoethoxy)acetic acid as a linker, it is possible to target a 10 bp DNA binding-site. An 11 bp binding site can be targeted with a low selectivity over the 10 bp site using 5-aminovaleric acid with an unpaired Im unit. The finding that a high imidazole content in a tandem hairpin results in a lower increase in affinity upon forming a dimer might possibly be amended by including β/ring pairings at appropriate positions in the hairpin. This is based on earlier observations that indicate a more optimal positioning of imidazole units upon incorporation of the β modification.[18] Although dimers of hairpin polyamides target larger binding-sites on DNA, it remains to be seen whether a higher selectivity is realized when targeting selected sites in a gigabase size genomic context.

Experimental

Materials

Boc-β-Alanine-(4-carboxamidomethyl)-benzyl-ester-copoly (styrene-divinylbenzene) resin (Boc-β-PAM-resin) was purchased from Peptide International, Louisville, KY. *N*,*N*′-dicyclohexylcarbodiimide (DCC), *N*,*N*-diisopropylethyl amine (DIEA), 1-methyl-2-pyrrolidinone and 3-(dimethylamino)propylamine were purchased from Aldrich, Milwaukee, WI. (*R*)-2-Fmoc-4-Boc-diaminobutyric acid was from Bachem and *N*-hydroxybenzotriazole (HOBt), 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexa-fluorophosphate (HBTU) and Boc-β-alanine from Novabiochem, San Diego, CA. Dichloromethane (DCM) and *N*,*N*-dimetyhlformamide (DMF) were reagent grade from EM Science, Gibbstown, NJ and trifluoroacetic acid from Halocarbon, River Edge, NJ.

Table 5. Equilibrium association constants[a] for **5** and **6**

Polyamide	**5**	**6**
DNA binding site and K_a	A G T G A T C A C A 1.2×10^9	T G C G T A C G C A 5.2×10^8

[a]The reported equilibrium association constants are the mean values obtained from three DNase I footprint titration experiments. Assays were carried out in the presence of 10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22 °C.

Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF) was performed at the Protein and Peptide Microanalytical Facility and DNA sequencing was performed at the Sequence/Structure Analysis Facility (SAF) at the Californian Institute of Technology. UV spectra were measured in water on a Hewlett-Packard Model 8452A diode array spectrophotometer. Preparatory reversed-phase HPLC was performed on a Beckman Gold system with a Waters PrepPak 25×100 mm, 6 μm C18 column. Water was obtained from a Millipore MilliQ water purification system and all buffers were 0.2 μm filtered.

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-γ]^{HN}γ-PyPyPy-β-Dpm (1a). Maldi-TOF-MS (monoisotopic), 1897.9 (1897.88 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-δ]^{HN}γ-PyPyPy-β-Dp (1b). Maldi-TOF-MS (monoisotopic), 1912.0 (1911.89 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-ε]^{HN}γ-PyPyPy-β-Dp (1c). Maldi-TOF-MS (monoisotopic), 1926.1 (1925.91 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-ζ]^{HN}γ-PyPyPy-β-Dp (1d). Maldi-TOF-MS (monoisotopic), 1940.76 (1939.92 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-η]^{HN}γ-PyPyPy-β-Dp (1e). Maldi-TOF-MS (monoisotopic), 1955.52 (calcd 1953.94 for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-G]^{HN}γ-PyPyPy-β-Dp (2a). Maldi-TOF-MS (monoisotopic), 1994.16 (1992.89 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-β]^{HN}γ-PyPyPy-β-Dp (2b). Maldi-TOF-MS (monoisotopic), (2006.90 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-γ]^{HN}γ-PyPyPy-β-Dp (2c). Maldi-TOF-MS (monoisotopic), 2020.26 (2020.92 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-δ]^{HN}γ-PyPyPy-β-Dp (2d). Maldi-TOF-MS (monoisotopic), 2036.43 (2034.93 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPyIm-ε]^{HN}γ-PyPyPy-β-Dp (2e). Maldi-TOF-MS (monoisotopic), 2050.91 (2048.95 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyImPy-δ]^{HN}γ-PyImPy-β-Dp (3). Maldi-TOF-MS (monoisotopic), 1914.1 (1913.88 calcd for M + H).

ImPyIm-(*R*)[ImPyIm-(R)^{H2N}γ-PyPyPy-δ]^{HN}γ-PyPyPy-β-Dp (4). Maldi-TOF-MS (monoisotopic), (1913.87 calcd for M + H).

ImPyIm-(*R*) ^{H2N}γ-PyPyPy-β-Dp (5). Maldi-TOF-MS (monoisotopic), 993.51 (993.49 calcd for M + H).

ImPyIm-(*R*) ^{H2N}γ-PyImPy-β-Dp (6). Maldi-TOF-MS (monoisotopic), 994.47 (994.49 calcd for M + H).

Construction of plasmid DNA

The plasmids pIK2, pIK3 and pIK4 were prepared by hybridization of the complementary sets of synthetic oligonucleotides indicated for each plasmid in Figure. 6. Each hybridized insert was ligated individually into *Bam*HI/*Hind*III linearized pUC19 using T4 DNA ligase. *Escherichia coli* JM109 high efficiency competent cells were then transformed with the ligated plasmid, and plasmid DNA from ampicillin-resistant white colonies isolated using a Qiagen HiSpeed Plasmid Maxi Kit. The presence of the desired insert was determined by dideoxy sequencing. Concentration of the prepared plasmid was determined at 260 nm from the relationship of 1 OD unit = 50 μg mL^{-1} duplex DNA.

Preparation of ^{32}P-end-labeled restriction fragments

Plasmid pIK2, pIK3 or pIK4 was linearized with *Eco*RI and *Pvu*II restriction enzymes, then treated with Klenow enzyme, deoxyadenosine 5′-[α-^{32}P]triphosphate and thymidine 5′-[α-^{32}P]triphosphate for 3′-labeling. The labeled fragment was loaded onto a 7% nondenaturating polyacrylamide gel, and the desired 268 bp band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published methods.

Quantitative DNase I footprinting

DNase I footprinting reactions were carried out as previously described.[23–25] Photostimulable storage phosphorimaging plates (Storage Phosphor Screen from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark for 12–16 h. A molecular Dynamics 425E PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integrations of all bands using ImageQuant v. 3.2 software.

Acknowledgements

We are grateful to the National Institutes of Health (GM 27681) for research support, and Swedish Research Council for Engineering Sciences (TFR) and The Swedish Foundation for International Cooperation in Research and Higher Education (STINT) for postdoctoral fellowships to I.K.

References and Notes

1. Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12890.
2. Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202.
3. Mapp, A. K.; Anasari, A. Z.; Ptashne, M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3930.
4. Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X. *J. Am. Chem. Soc.* **1951**, *73*, 341.
5. Arcamone, F.; Penco, S.; Orezzi, P.; Nicolella, V.; Pirelli, A. *Nature* **1964**, *203*, 1064.
6. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
7. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559.
8. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111.
9. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468.
10. White, S.; Turner, J. M.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 260.
11. deClairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909.
12. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
13. Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147.
14. Pilch, D. S.; Poklar, N.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306.
15. Kelly, J. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6981.
16. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104.
17. Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 7636.
18. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 6219.
19. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Angew. Chem.-Int. Edit. Engl* **1998**, *37*, 1421.
20. Herman, D. M.; Baird, E. E.; Dervan, P. B. *Chem.-Eur. J.* **1999**, *5*, 975.
21. Maeshima, K.; Janssen, S.; Laemmli, U. K. *EMJO J.* **2001**, *20*, 3218.
22. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.
23. Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132.
24. Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8462.
25. Trauger, J. W.; Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450.
26. Geierstanger, B. H.; Wemmer, D. E. *Annu. Rev. Biochem.* **1995**, *24*, 463.
27. Goodsell, D. S.; Kopka, M. L.; Cascio, D.; Dickerson, R. E. *Proc. Natl. Sci. U.S.A* **1993**, *90*, 2930.
28. Hansen, M. R.; Hurley, L. H. *Accounts Chem. Res* **1996**, 249.
29. Kahn, J. D.; Yun, E.; Crothers, D. M. *Nature* **1994**, *368*, 163.
30. Paolella, D. N.; Palmer, R.; Schepartz, A. *Science* **1994**, *264*, 1130.
31. Steitz, T. A. *Annu. Rev. Biophys* **1990**, *23*, 205.
32. Wu, H.; Crothers, D. M. *Nature* **1984**, *308*, 509.