HHS Public Access

Author manuscript

Org Biomol Chem. Author manuscript; available in PMC 2018 November 29.

Published in final edited form as:

Org Biomol Chem. 2017 November 29; 15(46): 9880–9888. doi:10.1039/c7ob02513k.

β-alanine and *N*-terminal cationic substituents affect polyamide-DNA binding

Beibei Liu[a,†], **Shuo Wang**[a,†], **Karl Aston**[b], **Kevin J. Koeller**[b], **Shahrzad Fanny Hakami Kermani**[b], **Carlos H. Castañeda**[b], **M. Jose' Scuderi**[b], **Rensheng Luo**[b], **James K. Bashkin**[*,b], and **W. David Wilson**[*,a]

[a]Department of Chemistry, Georgia State University, Atlanta, GA 30303, USA

[b]Department of Chemistry & Biochemistry, Center for Nanoscience, University of Missouri-St. Louis, St. Louis, MO 63121, USA

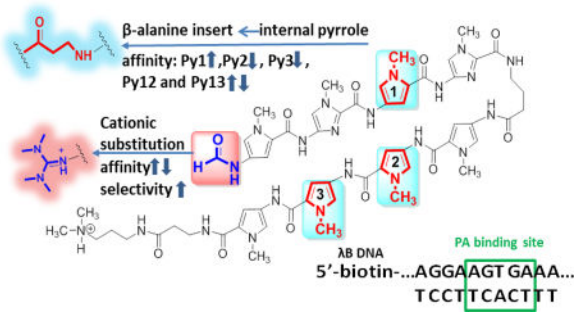
Abstract

Minor-groove binding hairpin polyamides (PAs) bind specific DNA sequences. Synthetic modifications can improve PA-DNA binding affinity and include flexible modules, such as β-alanine (β) motifs to replace pyrroles (Py), and increasing compound charge using *N*-terminal cationic substituents. To better understand the variations in kinetics and affinities caused by these modifications on PA-DNA interactions, a comprehensive set of PAs with different numbers and positions of β and different types of *N*-cationic groups was systematically designed and synthesized to bind their cognate sequence, the λB motif. The λB motif is also a strong binding promoter site of the major groove targeting transcription factor PU.1. The PA binding affinities and kinetics were evaluated using a spectrum of powerful biophysical methods: thermal melting, biosensor surface plasmon resonance and circular dichroism. The results show that β inserts affect PA-DNA interactions in a number and position dependent manner. Specifically, a β replacement between two imidazole heterocycles (ImβIm) generally strengthens binding. In addition, *N*-terminal cationic groups can accelerate the association between PA and DNA, but the bulky size of TMG can cause steric hindrance and unfavourable repulsive electrostatic interactions in some PAs. The future design of stronger binding PA requires careful combination of βs and cationic substituents.

Graphical abstract

[*] wdw@gsu.edu; Fax: +1 404-413-5505; Tel: +1 404-413-5503. [*] bashkinj@umsl.edu; Fax: +1 314-516-5342; Tel: +1 314-516-7352.
[†]B. L and S. W contributed equally to this paper.

Electronic Supplementary Information (ESI) available: Compound synthesis details, experimental details and supporting figures (Fig. S1–S2). See DOI: 10.1039/x0xx00000x



The binding preference of a systematic set of designed polyamides with β-inserts and cationic-substitutions with cognate DNA

Introduction

Polyamides (PA) are heterocyclic cations that consist of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) aromatic rings linked by amide bonds.[1–4] They are potentially useful in applications such as gene therapeutics and inhibiting protein-DNA interactions.[5–7] Derived from the natural DNA minor groove binding agents, netropsin and distamycin, PAs have been developed extensively over the years to enhance their DNA-binding affinity and specificity. Even though targeting the relatively narrow DNA minor groove, distamycin was found to form 2:1 antiparallel stacked dimer complexes with AT-rich DNA.[8] The finding led to the concept of covalently linking two PA monomers by a flexible linker that can mimic the bend of a hairpin. Therefore, hairpin PAs can form a 1:1 side-by-side complex with the two strands of DNA minor groove at higher affinity and specificity than traditional PAs.[1,9–11] Based on the hydrogen bonding interactions of the complex and its pairwise heterocycle stacking properties, general PA-DNA recognition rules were developed and have advanced molecular design.[3,10,12] However, challenges still exist, since elongating hairpin PAs to eight or more rings leads to an over-curved structure which no longer matches the minor groove shape of B-form DNA. The resulting shape mismatch for long PAs causes a loss of hydrogen bonding strength that erodes both binding affinity and specificity.[13,14] Incorporating a flexible motif such as β-alanine (β) to serve as a replacement of internal Py moieties has been reported to help compensate for the structural incompatibility and improve binding.[15–18] However, some literature results show a variety of binding affinity effects of β-inserts, including a decrease in binding affinity of some PAs by the introduction of βs. It is quite important to notice that the insertion of β at different positions of the same molecule has been shown to have very diverse effects on binding affinity.[17,19–21] It is essential, therefore, to understand the complex role of β insertion in PA-DNA interactions to guide the molecular design of better PAs.

Another challenge of large size hairpin PAs which contain more than six heterocyclic rings is aggregation and low solubility at high concentrations.[20,22,23] The natural product, netropsin, contains a cationic guanidine moiety at its *N*-terminus, which presumably enhances both the solubility and its affinity to anionic DNA sequences.[24] A previous study has demonstrated that the replacement of the monocationic Dp group [3-(dimethylamino)propylamine] by a dicationic Ta group (3,3'-diamino-N-

methyldipropylamine) could maintain the DNA binding mode and affinity of eight-ring hairpin PAs as well as significantly reduce PA aggregation.[20] Bashkin and coworkers have also reported that guanidinyl-substituted PAs have better antiviral activities against human papillomavirus (HPV) than traditional PAs.[25] Thus, it is important to incorporate cationic substituents into PA, explore their effects on PA-DNA binding interactions, and better understand how to modulate the solution properties of PAs.

Compound Design

In order to evaluate the effects of β inserts and cationic substituents on PA-DNA interactions, a series of hairpin PAs was systematically designed that covered reasonable internal exchanges of Py to β (Fig. 1A). In parallel, cationic substituents were also added to each counterpart with the exception of two molecules that have β inserts close to the *C*-terminus (Fig. 1A, KA2114 and KA2115). This is because β-alanine confers more flexibility to PAs when inserted close to a terminus than when positioned internally. With one positive charge at the *N*-terminus and two positive charges at the *C*-terminus, strong electrostatic repulsions would be expected to disrupt stacking of PA heterocycles and interfere with binding to the targeted DNA site. For this reason, KA2114 and KA2115 do not have cationic substituted counterparts.

The molecules in Fig. 1A are categorized into two groups: β-alanine inserts and cationic substituents, based on their modifications. To further assist data interpretation, all the heterocycles were numbered from 1 to 8 starting from the *N*-terminus. The parent PA, KA2035, contains eight heterocyclic rings and an *N*-terminal formamido group that is identical to that found in natural product Distamycin A. In the β-alanine inserts group, analogs of KA2035 that contain a single β replacement (KA2034 at ring 3, KA2041 at ring 6, and KA2114 at ring 7) and two βs (KA2040 at rings 3 and 6, and KA2115 at rings 3 and 7) at three different Py positions have been developed and synthesized. In the cationic substituents group: KJK6162 has the same heterocyclic rings as KA2035 but incorporates tetramethylguanidine (TMG) at the *N*-terminus. The TMG cationic *N*-terminal containing molecules also include FH1026 and FH1028, both of which have a single β insert, but at rings number 6 and number 3, respectively. Another member in this group is FH1024, which possesses the same composition with KA2040 and a TMG at the *N*-terminus.

The designed PA-DNA binding site is AGTGA. The binding site is contained in the λB promoter region, to which transcription factor PU.1 binds. Therefore, λB sequence is chosen as the target DNA sequence in this study to screen potential PU.1-DNA inhibitors in the future. This is important in that PU.1 is involved in many physiological diseases such as auto-immune diseases and acute myeloid leukemia (AML).[26,27] Extensive biophysical studies were carried out with the systematically designed compounds and their cognate and mutant DNAs. The results showed that β-inserts affect PA-DNA binding affinity in a manner that is dependent on PA composition and β position. For the current set of PA, a clear binding pattern between β-inserted PA and DNA was developed. The effects of flanking base pairs on PA-DNA binding were also investigated.

Results

Quantitative evaluation of PA-DNA binding kinetics and affinity

β-inserts—Biosensor-Surface Plasmon Resonance (SPR) is a label-free technique with excellent sensitivity that can quantitatively monitor biomolecular binding interactions in real time. Representative SPR sensorgrams are shown in Fig 2. It is noticeable that the dissociation rate is getting faster from strong to weak binders. The other sensorgrams are listed in Supplementary, Fig S1. The kinetic rate constants and equilibrium binding affinities of eleven designed PAs were determined using SPR and summarized in Table 1. The six PAs in the β-insert group have a formamido (F) group at the *N*-terminus and β at different positions. KA2115, with two βs replacing Py rings at positions 3 and 7, leads the group with the highest binding affinity (0.49 ± 0.01 nM). KA2040, another PA with two βs but at rings 3 and 6, binds to the cognate sequence essentially equally as strong (0.54 ± 0.01 nM). A single β inserted PA at ring 3, KA2034, displayed a slightly weaker but quite comparable binding affinity (0.67 ± 0.01 nM) to KA2115 and KA2040. These results suggest that the position and number of β inserts in the lower strand (rings 5–8) is not a dominant factor for this hairpin PA-DNA interaction, and do not contribute significantly to binding affinity enhancement. The unmodified parent molecule, KA2035, exhibits a moderate affinity (1.30 ± 0.01 nM) with regard to the last three molecules. This is primarily caused by its relatively lower association rate (k_a). KA2035, however, has the second lowest dissociation rate (k_d) among all the PAs tested, which compensated for its low k_a and rendered KA2035 a relatively strong binder. KA2041, with only one β-alanine at position 6, binds the cognate DNA with a higher k_a but a much higher k_d compared to KA2035, resulting in weakened overall interactions. Nonetheless, this situation worsened when the replacement was moved to position 7, as in KA2114. The k_d of KA2114 is so high, around 20 times higher than the parent molecule KA2035, that it cannot obtain high binding affinity even with a higher k_a value than KA2035.

Cationic substitutions—Relative to the β insertion group, in which all molecules conserve an unmodified *N*-terminus, four PAs were designed with further modification at the *N*-terminus with the specific cationic substituent, tetramethylguanidinyl (TMG). The binding affinities and kinetics of these PAs have been evaluated (Table.1). FH1028, the PA with a single β insert at ring 3 and a TMG at the *N*-terminus leads all the other PAs in binding affinity to the cognate DNA (0.16 ± 0.02 nM). The high affinity is attributed to its high k_a and relatively low k_d. Compared to its analog KA2034 which has the identical β modification but different *N*-terminal substituents, FH1028 has a five-fold higher k_a which contributes to the strengthened binding affinity. This result indicates that the TMG group can accelerate PA-DNA binding. Similarly, FH1024 (*N*-TMG) and KA2040 (*N*-F) were compared to each other because they have identical β-inserts at positions 3 and 6. The *N*-TMG attached FH1024 has a k_a that is around two times higher than KA2040 (*N*-F), but its drastically increased k_d makes it non-competitive with KA2040. Interestingly, KJK6162, modified directly from the parent, KA2035, by substituting F with TMG at the *N*-terminus, binds the λB sequence so weakly (60 times weaker than KA2035) that its kinetic rate constant cannot be determined. Similarly, FH1026, the *N*-TMG analog of KA2041, has a 20 times weaker binding affinity than KA2041, well out-of-range kinetic constants by SPR.

Screening of relative binding affinities with cognate and mutant DNAs by Thermal Melting (Tm)

Thermal melting is an effective way to evaluate the relative binding affinity and the selectivity of a particular PA among several DNA sequences. Tm of the cognate DNA and DNA mutants and their complexes with PAs are determined at 1:1 molar ratio. The difference between Tm values of the complex and DNA, ΔTm, reflects the thermal stability of the complex, thus the binding affinity of PA. The values are listed in Table 2.

The comparison of mutant DNAs are made with each PA-cognate DNA complex and are categorized according to their extent of binding. ΔTm values in ±5% relative to the ΔTm of cognate DNA are designated as equal and uncolored, dark gray means slightly stronger (≤120%, >105%), gray means slightly weaker (<95%, ≥80%), light gray is categorized as weaker (<80%, ≥50%) and lightest gray is much weaker (<50%). The determined ΔTm values of PAs with cognate DNA are in good agreement with the SPR results.

The first mutant sequence evaluated is a Py recognizing T that is mutated to A. Both in β-insert and the cationic substituent group, the affinity is not affected significantly and is either equal or slightly higher or lower, even for the molecules where Py is replaced by β (single-β or double-β replacement). Particularly, KA2034 (f-PyIm βIm-γ-PyPyPyPy-β-Dp) showed slightly weaker binding when β/Py targeting T·A is switched to A·T. In a similar case, KA2041 (f-PyImPyIm-γ-Py βPyPy-β-Dp) who has Py/β targeting T·A exhibited slightly higher affinity to the A·T mutant. These results indicate that T·A prefers β/Py to Py/β, and vice versa. The second mutant is made by switching the terminal amine recognizing A to T. In both groups, most molecules are not affected by this change and the binding affinity stays comparable to that of the cognate DNA. What is interesting is, KA2040 and FH1024, both of which have double-β replacement where Py/Py is substituted by β/β, have slightly elevated binding affinities. The third mutant is considered as a more dramatic mutation, in which the Im recognizing G is mutated to T. As a result, an apparent drop of binding affinity is observed for most PAs. However, as a strong binder to cognate DNA, FH1028 and KA2115 still binds to mutant 3 quite strongly, ΔTm = 14.8, and 10.5, respectively. Interestingly, the binding of KA2041 and KA2114 to mutant 3 is comparable to that of cognate DNA, especially KA2114. This result indicates that even though Im recognizes G specifically, it can still tolerate T to some extent, so that the complex can still be formed. Mutants 4 and 5 are made by switching G to C. Much weaker binding appeared at this level of mutation. Since in this case, Py is designated to target G, which is known with either one-β or two-β replacement, the Im/Py or Im/β combination is not very tolerant with C·G base pairs.

Effect of *N*-terminal cationic group on PA selectivity of DNA flanking sequences

To further evaluate the effect of the *N*-terminal cationic group on PA-DNA interactions, binding affinities of the same set of PAs to DNAs that have different mutated flanking sequences were measured and compared. The three DNAs used here are λB DNA (5'-Biotin-CCAAATAAAAGG **AAGTGA** AACCAAG*CTCT*CTTGGTTTCACTT CCTTTTATTTGG-3'), SC1 DNA (5'-Biotin-CGGCCAAGCCGG **AAGTGA** G TGCC*TCTC*GGCACTCACTTCCGGCTTGGCCG-3'), and GAGA mutant DNA (5'-

Biotin- CCAAATAAAAGA **G**<u>**AGTGA**</u> **A**ACCAAG*CTCT*CTTGGTTTC ACTCTCTTTTATTTGG-3'). In λB DNA, the PA binding site -AGTGA- is surrounded by A on both 5' and 3' ends. The SC1 sequence has a G flanking at the 3' side, while GAGA sequence has a G at the 5' end. The binding to the DNAs was evaluated by SPR as described above and the SPR sensorgrams and binding results are shown in Supplementary Fig. S2 and Table. 3. By comparing binding of λB and SC1 DNA, the *N*-F molecules (KA2035 plus β-inserts group) have very similar binding affinities to both sequences. For the *N*-TMG molecules, the binding affinities to λB DNA are fairly strong. However, when tested against the SC1 DNA, their binding affinities significantly decreased, even to the point of no detectable binding. Interestingly, when we compare the binding to GAGA sequence with λB DNA, the *N*-TMG group of PAs show strong and comparable binding in general, and binding of FH1024 is even stronger to GAGA than λB DNA. Fig. 3B shows the binding affinity of PAs to λB, SC1 and GAGA mutant sequences. It is clear that each PA in the β-inserts group has very close binding strength to all three DNAs, except that KA2115 has a slightly larger deviation. But in the cationic substitution group, hardly any traces of PA-SC1 binding (black column) can be seen. These results suggest that *N*-TMG PAs are quite sensitive and selective to the 3' flanking base pairs of -AGTGA-, but not 5', and they prefer A to G.

Evaluation of the PA binding mode by CD

In Figure 4, the large positively induced CD signals upon PA-DNA binding at around 300 nm to 370 nm, where PAs absorb while DNA signals do not interfere indicate a DNA minor groove binding mode as expected for hairpin PAs.[28]

Discussion

Quantitative evaluation of PA-DNA binding kinetics and affinity

SPR has greatly facilitated the quantification of kinetics between small molecules and DNA. As important, sequence-specific DNA minor groove binders, polyamides have been extensively studied over the decades and are seen as potential drug targets for antiviral, antitumor and antibacterial therapeutics. Yet the position and number dependence of β-modified PA-DNA interaction still have many uncertainties and their features play a critical role in designing more specific and stronger binding PAs.

β-inserts—As shown in Fig. 1B, along the λB sequence, the 5' side of PU.1 binding site is an AT rich region. Previous literature has employed the A·T specific minor groove targeting compounds to bind to the 5' side A·T base pairs in order to inhibit PU.1 binding.[29] The efficient allosteric inhibition activity demonstrated by these compounds has encouraged us to further extend our target binding site to the 3' side. A combination of 5' side binding minor groove heterocyclic cations and 3' targeting PAs should provide a very powerful inhibition potential.

The parent PA, KA2035 was thus designed, according to the recognition rules, to target the 3' side AGTGA sequence. The PA has a strong binding affinity of 1.30 nM, and was able to successfully and precisely recognize the target sequence. β-alanine is a building block that is

incorporated into PAs to increase their flexibility. In a successful effort to improve binding affinity, KA2035 was modified by replacing internal Py at different positions with β inserts. The SPR binding affinities of modified PAs vary in a β insert number-and-position dependent manner.

Specifically, when compared to KA2035, three PAs (KA2115, KA2040 and KA2034) have strengthened binding affinities. A common modification strategy adopted by all the three PAs is that Py heterocycle 3 is substituted by one β insert. It is highly likely that this position 3 β insert in the upper strand (rings 1–4) added the flexibility needed to the molecule to adjust to the cognate DNA minor groove curvature. However, this particular β also allows for relaxation and realignment of the PA to compensate for the different positions of H-bond donors and acceptors in Py and Im components of PAs, where Py uses an exocyclic amide NH as an H-bond donor to DNA (for A, T and C) and Im uses a cyclic N as an H-bond acceptor from G. The H-bonding locations on DNA are in register with each other, but those on the PA are not. Therefore, the locations of hydrogen bond donors and acceptors between PA and DNA are better oriented at more optimized distances and angles, resulting in strengthened hydrogen bonding with higher PA-DNA binding affinity. On the other hand, a single β insert on the lower strand (rings 5–8), KA2041 and KA2114, gave no enhanced binding. This is due to the possibility that replacement of Py by a β at position 6 or 7 does not help the PA overcome its overall rigidity. Instead, a reduction of van der Waals contacts occurred when substituting Py 6 or 7 with a β. Interestingly, the β position in double β-substituted KA2115 (K_D = 0.49 nM) is a combination of βs in the single β-substituted KA2034 (K_D = 0.67 nM) and KA2114 (K_D = 4.1 nM), but the binding affinity of KA2115 is higher than either of the other two. Very similarly, the double β-substituted KA2040 (K_D = 0.54 nM) also has higher binding affinities than the single β-substituted KA2034 (K_D = 0.67 nM) and KA2041 (K_D = 1.62 nM). This is probably because that the combination of upper strand and lower strand β inserts gives PA even better flexibility to arrange itself to the optimal orientation for hydrogen bonding upon interacting with DNA, and because that optimal β-β stacking can occur when binding to certain DNA sequence.

The binding affinity is determined by the association and dissociation rate constants. From Table 1, it is clear that all PAs in the β-inserts group have higher association rate constants than the parent PA, KA2035. This is in contrast to a previous study in which the association rate constants of modified PAs were generally lower than those of unmodified compounds.[20] Of note, the 2014 study focused on the dImImPyIm-γ-PyPyPyPy-β-Dp/Ta sequence and derivatives while the current work studies formamidoPyImPyIm-γ-PyPyPyPy-β-Dp/Ta and derivatives. The extra, in register H-bond provided by the potentially rotatable formyl group coupled with the disparate target sequences of the two studies and different minor groove widths may all contribute to fundamentally different binding kinetics. The opposite effects of β-alanine inserts on the association rate of these two (current and previous) sets of PAs indicate that changes in PA composition (Py and Im heterocycles arrangement and content) and targeted, cognate DNA sequence respond differently to β inserts. After inserting β, the previous PAs were too flexible and needed time to adjust themselves to DNA upon binding, while the current set of β-inserted PAs fit better to the curvature of λB sequence than the parent molecule, resulting in more rapid binding.

The dissociation rate constants of those PAs with increased affinity are quite comparable to the parent PA, KA2035. For these tighter-binding PAs, as mentioned above, the upper strand β modification has greatly optimized hydrogen bond donor and acceptor orientations and thus strengthened hydrogen bonding between PAs and DNA (KA2034). But for the weak binding KA2041 and KA2114, their dissociation is so much faster that they can hardly maintain strong binding. It is noticeable that the single lower strand β replacement resulted in higher dissociation rate constant. Bashkin *et al.* built a docking model to compare a single-β-inserted PA2 (ImImβPy-γ-PyImPyPy-β-Dp) with an all-ring eight-ring PA1 (ImImPyPy-γ-PyImPyPy-β-Dp), and found that Py is more effectively stacked with both the adjacent PA strand and the DNA backbone than β-alanine.[19] Higher dissociation rate constants for some β derivatives might be a result of loss of hydrophobic interactions.

Cationic substitutions—Since DNA is highly negatively charged, adding positive charges on PAs should help enhance the electrostatic interactions between the two. More charged groups can also improve the solubility of PAs. Therefore, in addition to β inserts, additional cationic groups are attached to the *N*-termini to facilitate PA-DNA binding. SPR results showed varied effects of cationic substitutions on PA-DNA interactions.

As shown in Fig. 3A, FH1024 and FH1028, compared to KA2040 and KA2034, their analogs with formamido *N*-termini (*N*-F), exhibited slightly different binding affinities. FH1024 binds about three times weaker than KA2040 and FH1028 binds around three times tighter than KA2034. Interestingly, both TMG derivatives have higher association rate constants than their formamido analogs. The results implied that *N*-TMG can help promote PA-DNA binding by adding extra charge-charge interactions. On the other hand, the binding affinities of *N*-TMG KJK6162 and FH1026 dropped to dramatically low values for eight-ring PAs and cannot compete with their *N*-F analogs. This is probably because when the formamido group is substituted by the bulky TMG at the *N*-terminus while not having β inserts on the top strand, molecules like FH1026 and KJK6162 are not flexible enough to accommodate either the repulsive electronic interactions or the steric hindrance between the two termini. As a result, the free molecules tend to be less stacked, making it harder to form hairpin and conform ideally to the minor groove of cognate DNA upon binding. Interestingly, unlike KA2040 or KA2115, the affinity of the double β-substituted TMG compound FH1024 (K_D = 1.65 nM) is between the two single-β-substituted FH1028 (K_D = 0.16 nM) and FH1026 (K_D = 35 nM). The much higher dissociation constants of FH1024 than KA2040 plays an important role in lowering its binding affinity, indicating that the *N*-TMG can affect optimal β-β stacking and that the extra flexibility does not always optimally orient the PA upon binding to DNA, while the loss of hydrophobic interactions can overshadow the effort to enhancing hydrogen bonding strength.

Effect of *N*-terminal cationic group on PA selectivity of DNA flanking sequences

The selective properties of the cationic TMG group over the immediate flanking sequence of DNA binding site can be utilized as an important tool for drug design. The binding of the β-insert group molecules to SC1 sequence are in good agreement with their binding to λB sequence, meaning that β inserts have similar effects on the binding to both DNA sequences. However, the selective impacts of TMG overshadowed the β-insert effects in the cationic

group, making it a determinant factor upon binding to SC1 sequence. For GAGA sequence, β modification did not significantly increase the affinity among formamido group PAs, yet a single β replacement at position 3 (KA2034) still showed stronger binding than the parent PA, KA2035. Especially, both KA2115 and KA2040 showed weaker binding to GAGA than to the other two sequences, while the other molecules in the group have comparable affinities to all three sequences. The *N*-TMG FH1024, on the other hand, is around 12 times more favorable to GAGA sequence than λB sequence. The difference in binding strength of those molecules caused by alteration in flanking sequences, together with the selectivity of the TMG group, indicates that the DNA context can affect some PA-DNA binding, thus increasing the selectivity of those PAs. This provides a way to target one specific transcription factor among a transcription factor family, for which all proteins have consensus binding site with different flanking sequences for the different specific targets. A larger DNA pool is required for further tests of this selectivity.

Conclusions

Taken together, a binding pattern has been developed that links the modification with binding affinity compared to the parent PA, KA2035: 1) For single-β-containing PA, β/Py increases binding affinity when targeting T·A; Py/β decreases binding affinity when targeting T·A. Im/β decreases binding affinity when targeting G·C. 2) For double-β-containing PA, Imβ/βPy increases binding affinity when targeting GT·CA. β/β can increase or decrease binding affinity when targeting T·A. 3) When compared to the TMG-PA KJK6162, any β modification can increase the binding affinity (in the above, a slash "/" separates pairs of PA building blocks from opposite strands of the hairpin).

Overall, when Py between two Im is replaced by β, the binding affinity tends to go up, while if β appears on a lower strand consisting of four Py, the binding affinity tends to go down. Fortunately, a combination of the two positions can bring the affinity back to even higher levels. Sugiyama *et al* has shown that Im heterocycles are more planar than Py rings, which makes an Im-derived molecule more rigid than a corresponding Py-polyamide, e.g. when binding to DNA.[30] This could partially account for the slow association rate of the unmodified parent PA, KA2035. However, when we substitute the Py in the middle (number 3) with β, the whole molecule, especially the upper strand, becomes more flexible and more able to adjust to the curvature of DNA's minor groove upon binding. The finding that replacing Py between two imidazoles with β increases affinity is consistent with Turner *et al.* results, where PA10 (Im βImPyPyPy-γ-ImPyPyPyβPyβ-Dp) had ca. five times stronger binding affinity to its cognate DNA than the non β-inserted PA9 (ImPyImPyPyPy-γ-ImPyPyPyPyPyβ-Dp), and PA12(Im βImPy-γ-Im βImPyβ-Dp), with two ImXIm (X = β, Py) groups involved, is 100 times stronger than its unsubstituted analog PA11 (ImPyImPy-γ-ImPyImPyβ-Dp) in binding to its cognate DNA.[17] Similarly, Dickinson *et al.* also showed an increase in affinity when substituting Py in PA1 (ImPyImPy-γ-PyPyPyPy-β-Dp) with β in PA2 (Im βImPy-D-PyPyPyPy-β-Dp, where D denotes diaminobutyric acid).[31] Other studies also showed increased binding affinities of ImβIm to 5'-GCGC-3' sequence than ImPyIm.[32,33]

As for the bottom strand, the curvature of four Py fits nicely with that of DNA with optimal hydrogen bonding. Replacing Py with β on the other hand would reduce the van der Waals interaction between PA and DNA without allowing the PA to form better hydrogen bonds. This is also shown in Bashkin's docking model.[19] As a result, the overall binding is weaker than unmodified PA. Interestingly, when two of the above modifications are combined, the positive binding effects of the upper strand can reverse the negative binding effects of the lower strand, especially when Py/Py is replaced by β/β.

The *N*-terminal cationic group can have a positive effect on accelerating the association process of PA and DNA. However, because of the bulky nature of TMG group, steric clash and sometimes repulsive electrostatic interactions are generated upon binding to DNA, resulting in impaired DNA binding affinity of PAs. Thus, the TMG group has to be carefully combined with β-alanine inserts to achieve optimal PA-DNA binding strength.

Materials and Methods

The compound synthesis strategies and details, biophysical experimental procedures and additional results are provided in the ESI.†

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank NIH GM111749 to W.D.W. and the UMSL College of Arts and Sciences and Office of Research Administration to J.K.B. for financial support. The UMSL Agilent 600 MHz NMR spectrometer was obtained using funds from the NSF (0959360). J.K.B. thanks Prof. B. J. Bythell (UMSL) and the Danforth Plant Sciences Center (NSF DBI 0922879) for HRMS.

References

1. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. Proc. Natl. Acad. Sci. U. S. A. 1985; 82:1376. [PubMed: 2983343]
2. Wemmer DE. Annu. Rev. Biophys. Biomol. 2000; 29:439.
3. Dervan PB. Bioorg. Med. Chem. 2001; 9:2215. [PubMed: 11553460]
4. Neidle S. Nat. Prod. Rep. 2001; 18:291. [PubMed: 11476483]
5. Moretti R, Donato LJ, Brezinski ML, Stafford RL, Hoff H, Thorson JS, Dervan PB, Ansari AZ. ACS Chem. Biol. 2008; 3:220. [PubMed: 18422304]
6. Edwards TG, Koeller KJ, Slomczynska U, Fok K, Helmus M, Bashkin JK, Fisher C. Antiviral Res. 2011; 91:177. [PubMed: 21669229]
7. Koeller KJ, Harris GD, Aston K, He G, Castaneda CH, Thornton MA, Edwards TG, Wang S, Nanjunda R, Wilson WD, Fisher C, Bashkin JK. Med. Chem. 2014; 4:338.
8. Pelton JG, Wemmer DE. Proc. Natl. Aca. Sci. U. S. A. 1989; 86:5723.
9. Coll M, Frederick CA, Wang AH, Rich A. Proc. Natl. Aca. Sci. U. S. A. 1987; 84:8385.
10. Mrksich M, Parks ME, Dervan PB. J. Am. Chem. Soc. 1994; 116:7983.
11. Woods CR, Ishii T, Wu B, Bair KW, Boger DL. J. Am. Chem. Soc. 2002; 124:2148. [PubMed: 11878968]
12. Parks ME, Baird EE, Dervan PB. J. Am. Chem. Soc. 1996; 118:6147.
13. Kelly JJ, Baird EE, Dervan PB. Proc. Natl. Aca. Sci. U. S. A. 1996; 93:6981.

14. Kielkopf CL, Baird EE, Dervan PB, Rees DC. Nat. Struct. Biol. 1998; 5:104. [PubMed: 9461074]
15. Zhang W, Bando T, Sugiyama H. J. Am. Chem. Soc. 2006; 128:8766. [PubMed: 16819870]
16. Trauger JW, Baird EE, Mrksich M, Dervan PB. J. Am. Chem. Soc. 1996; 118:6160.
17. Turner JM, Swalley SE, Baird EE, Dervan PB. J. Am. Chem. Soc. 1998; 120:6219.
18. Wang CC, Ellervik U, Dervan PB. Bioorg. Med. Chem. 2001; 9:653. [PubMed: 11310600]
19. Bashkin JK, Aston K, Ramos JP, Koeller KJ, Nanjunda R, He G, Dupureur CM, Wilson WD. Biochimie. 2013; 95:271. [PubMed: 23023196]
20. Wang S, Aston K, Koeller KJ, Harris GD Jr, Rath NP, Bashkin JK, Wilson WD. Org. Biomol. Chem. 2014; 12:7523. [PubMed: 25141096]
21. Wang S, Nanjunda R, Aston K, Bashkin JK, Wilson WD. Biochemistry. 2012; 51:9796. [PubMed: 23167504]
22. Hargrove AE, Raskatov JA, Meier JL, Montgomery DC, Dervan PB. J. Med. Chem. 2012; 55:5425. [PubMed: 22607187]
23. Liu B, Kodadek T. J. Med. Chem. 2009; 52:4604. [PubMed: 19610652]
24. He GF, Bashkin JK. Future Med. Chem. 2015; 7:1953. [PubMed: 26496156]
25. Bashkin, JK., Edwards, TG., Fisher, C., Harris, GD., Koeller, KJ. US Patent. 9,133,228. 2015.
26. Penaranda C, Kuswanto W, Hofmann J, Kenefeck R, Narendran P, Walker LS, Bluestone JA, Abbas AK, Dooms H. Proc. Natl. Aca. Sci. U. S. A. 2012; 109:12668.
27. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, Akashi K, Fiering S, Tenen DG. Nat. Genet. 2004; 36:624. [PubMed: 15146183]
28. Lyng R, Roger A, Norden B. Biopolymers. 1992; 32:1201. [PubMed: 1420988]
29. Munde M, Wang S, Kumar A, Stephens CE, Farahat AA, Boykin DW, Wilson WD, Poon GM. Nucleic Acids Res. 2014; 42:1379. [PubMed: 24157839]
30. Han YW, Matsumoto T, Yokota H, Kashiwazaki G, Morinaga H, Hashiya K, Bando T, Harada Y, Sugiyama H. Nucleic Acids Res. 2012; 40:11510. [PubMed: 23042247]
31. Dickinson LA, Trauger JW, Baird EE, Ghazal P, Dervan PB, Gottesfeld JM. Biochemistry. 1999; 38:10801. [PubMed: 10451376]
32. Han YW, Kashiwazaki G, Morinaga H, Matsumoto T, Hashiya K, Bando T, Harada Y, Sugiyama H. Bioorg. Med. Chem. 2013; 21:5436. [PubMed: 23810670]
33. Meier JL, Yu AS, Korf I, Segal DJ, Dervan PB. J. Am. Chem. Soc. 2012; 134:17814. [PubMed: 23013524]

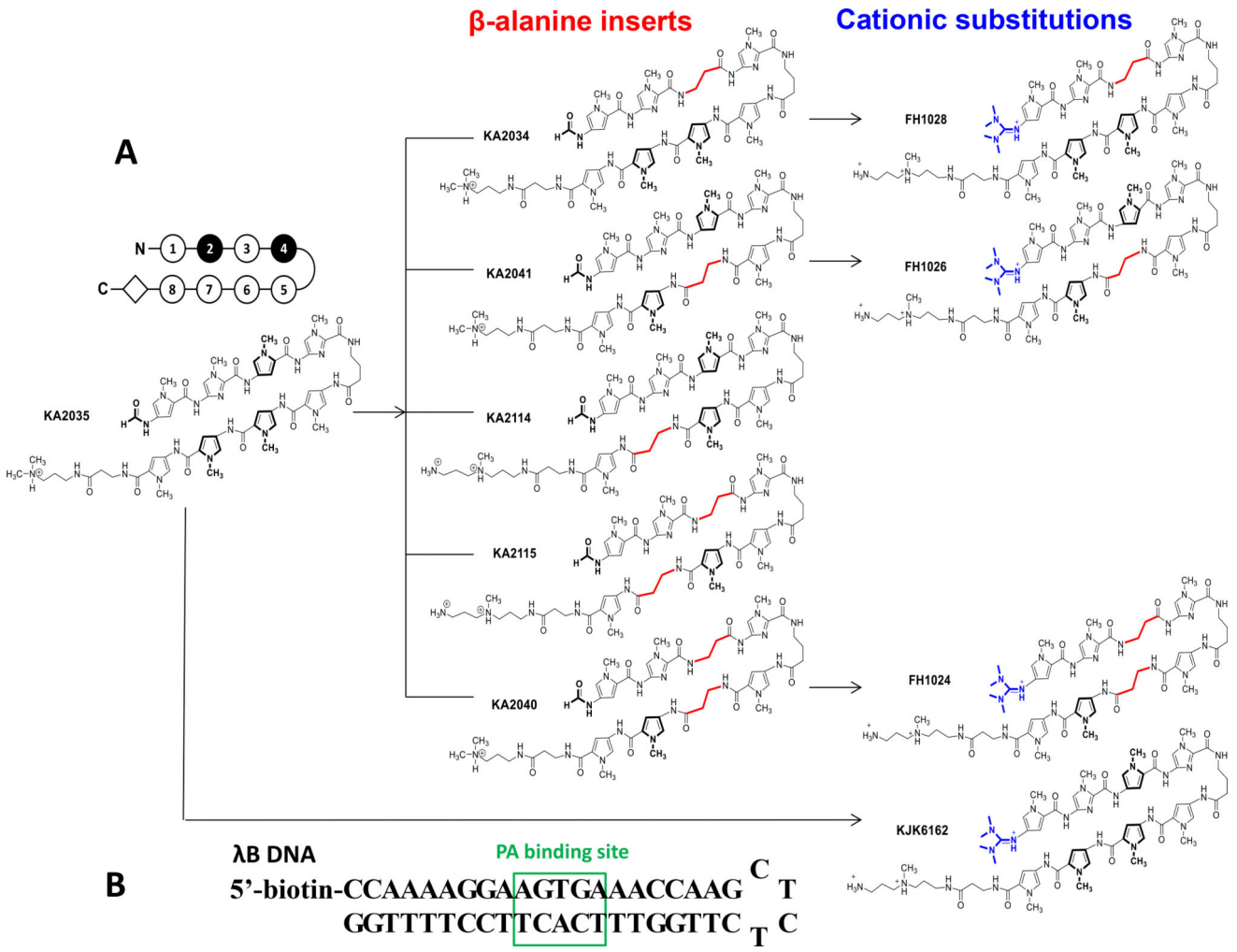


Fig. 1.

A)Systematically designed polyamides that are categorized into two modified groups: β-alanine inserts (red) and cationic substitutions (blue). The illustration above KA2035 is the simplified representative of KA2035. The open and closed circles stand for pyrrole and imidazole, respectively. The diamond represents β-alanine. The numbering starting from *N*-terminus to *C*-terminus applies to all PAs. B) The biotinylated cognate binding sequence of the PAs: λB DNA. The predetermined PA binding site is highlighted in the green frame.

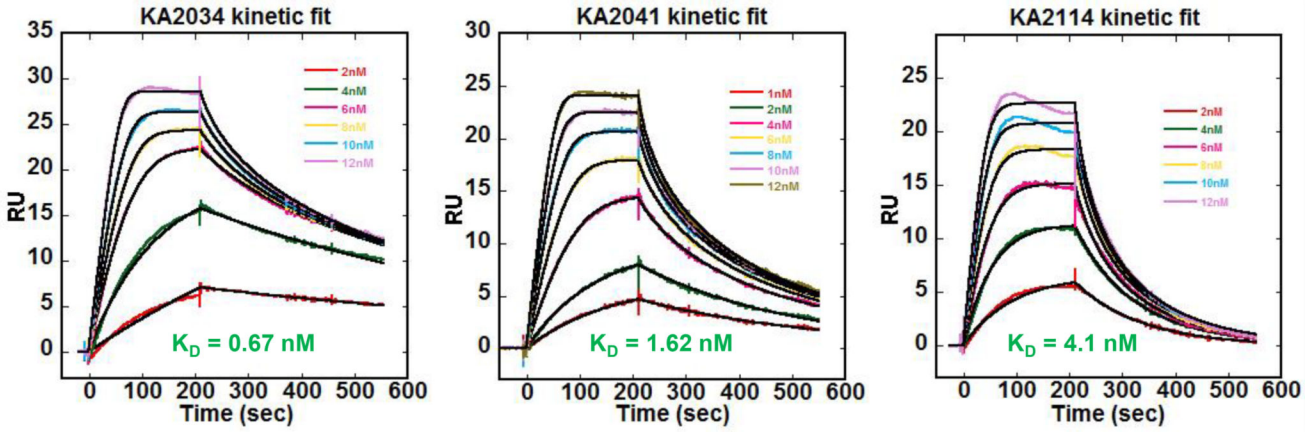


Fig. 2.

Representative SPR sensorgrams of PAs binding to λB DNA. From left to right, there are strong (K_D = 0.67 nM), intermediate (K_D = 1.62 nM), and weak (K_D = 4.1 nM), binders. The colored lines are experimental sensorgrams. The black overlays are 1:1 global kinetic fits.

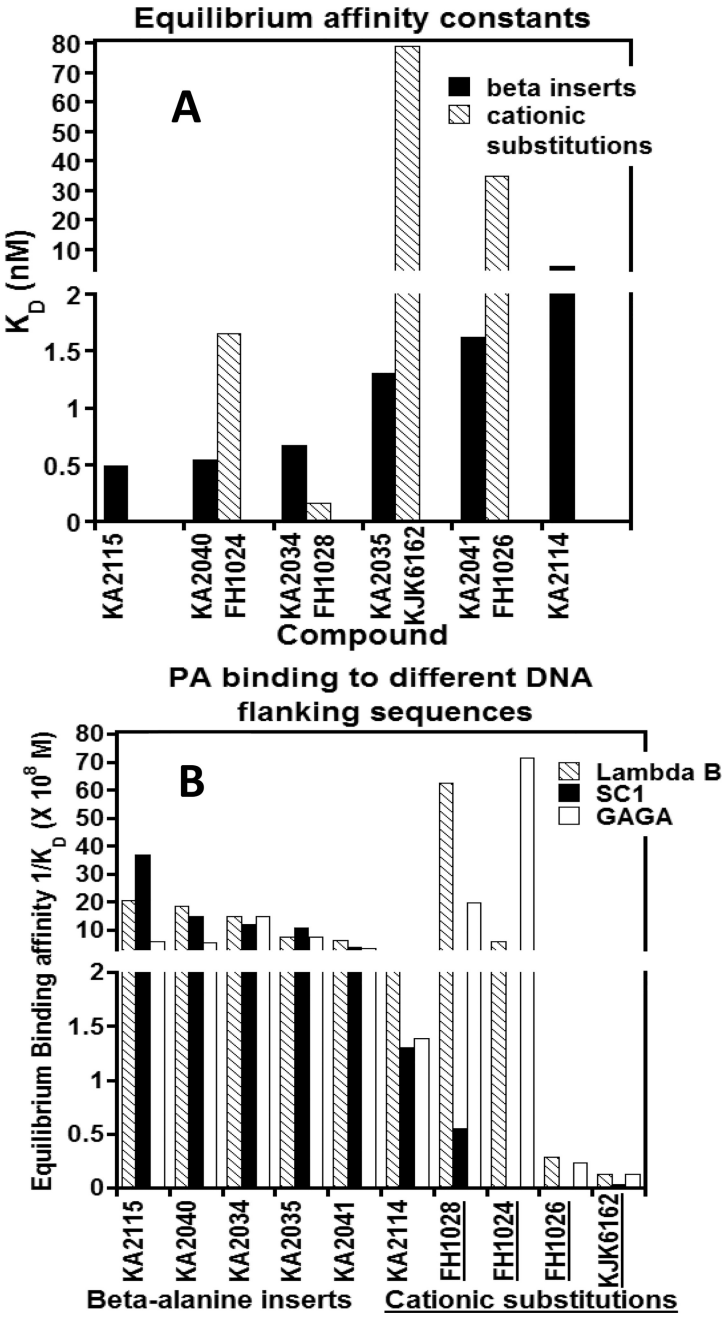


Fig. 3.

A) Direct comparison of equilibrium affinities of all PAs binding to λB sequence. The grouped red and blue columns represent molecules that have the same β modifications but different N-termini. B) Binding affinities of all PAs to λB, SC1 and GAGA mutant sequences. The binding of PAs in β-alanine inserts group to SC1 sequence is comparable to that of λB and GAGA mutant sequences. While extremely weak, some even undetectable binding was observed for PAs in the cationic substitutions group with SC1 sequence. Note that KA2035 is included into β-inserts group for ease of comparison.

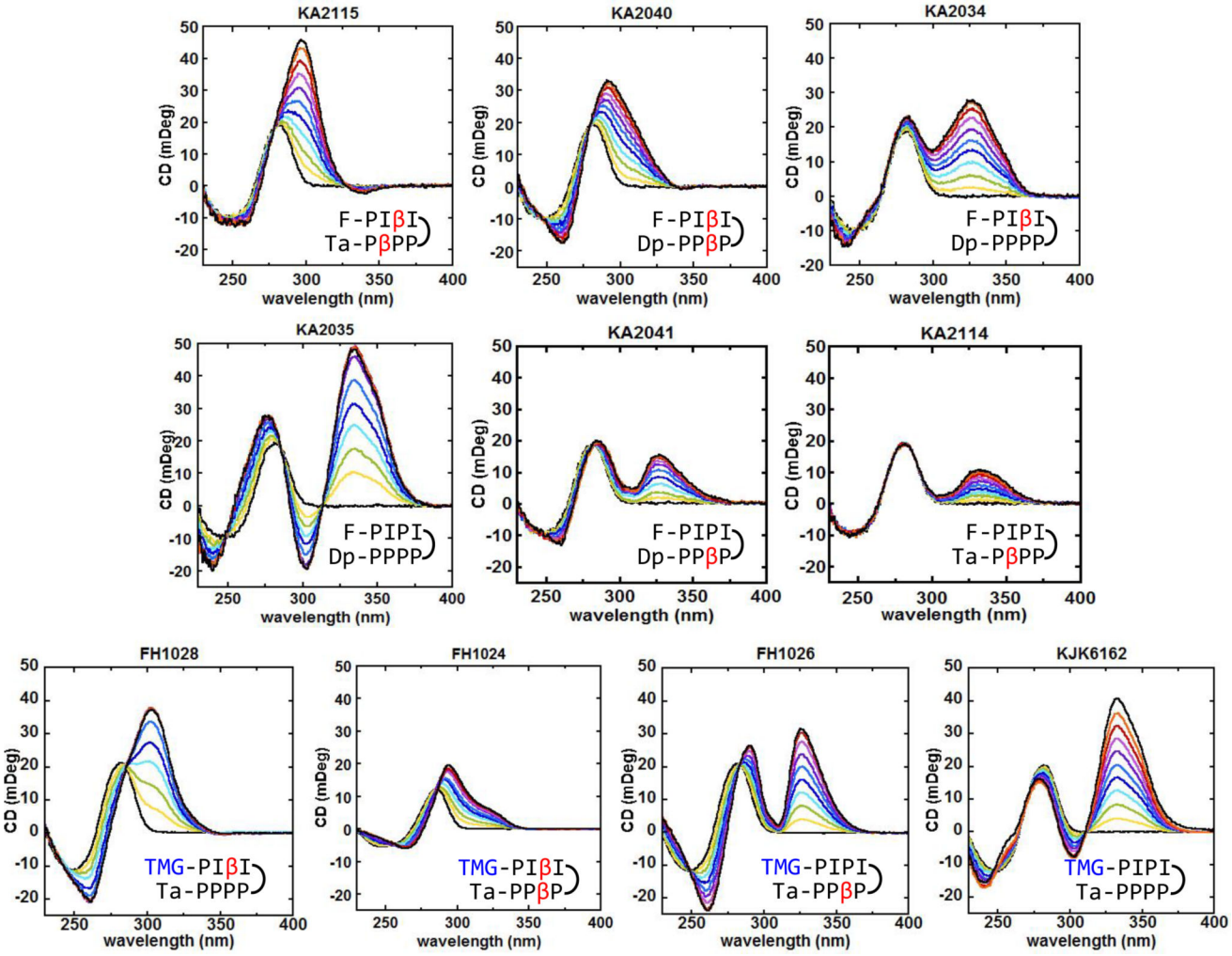


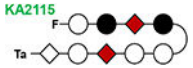
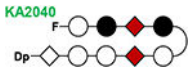
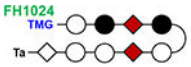
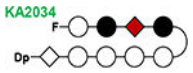
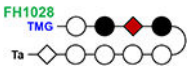
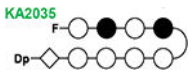
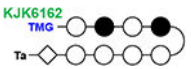
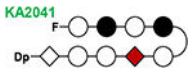
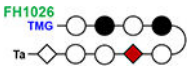
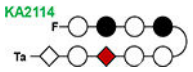
Fig. 4.

CD spectra of PA titration with short λB sequence (5'-GGAAGTGAACCTCTGTTCACTTCC-3').

Org Biomol Chem. Author manuscript; available in PMC 2018 November 29.

Table 1

Kinetic rate constants and equilibrium binding constants for all PAs

	k_a ($\times 10^7$ M^{-1} s^{-1})	k_d ($\times 10^{-3}$ s^{-1})	K_D (nM)		k_a ($\times 10^7$ M^{-1} s^{-1})	k_d ($\times 10^{-3}$ s^{-1})	K_D (nM)
KA2115 	1.7 ± 0.1	9.9 ± 0.1	0.49 ± 0.01				
KA2040 	1.5 ± 0.1	8.4 ± 0.1	0.54 ± 0.01	FH1024 	3.2 ± 0.1	52.7 ± 0.6	1.65 ± 0.03
KA2034 	1.0 ± 0.1	6.9 ± 0.1	0.67 ± 0.01	FH1028 	5.0 ± 0.1	7.9 ± 0.1	0.16 ± 0.02
KA2035 	0.58 ± 0.03	7.6 ± 0.1	1.30 ± 0.01	KJK6162 	ND	ND	79 ± 4
KA2041 	2.1 ± 0.1	34.1 ± 0.5	1.62 ± 0.04	FH1026 	ND	ND	35 ± 2
KA2114 	3.9 ± 0.1	161 ± 4	4.1 ± 0.1				

Note: PAs in the β-inserts group are ranked according to their binding affinities from strong to weak. FH1026 and KJK6162 have very fast binding kinetics that are beyond the limitation of instrument detection. Thus no binding kinetics are reported here (ND). The equilibrium binding affinities were measured using steady state fitting. Note that KA2035 is included into β-inserts group for ease of comparison. Diamond, open circle and close circle represent β-alanine, pyrrole and imidazole respectively; Red diamond denotes β-alanine that replaced Py; F means formamido group and TMG means tetramethylguanidinyl group. Dp is short for 3-(dimethylamino)propylamine and Ta is short for 3,3'-diamino-*N*-methyldipropylamine.

Table 2

ΔTm values of PAs with cognate and five mutant sequences

PA \ DNA	AGTGA	Mutant 1 AGAGA	Mutant 2 AGTGT	Mutant 3 AGTTA	Mutant 4 ACTGA	Mutant 5 AGTCA
Tm (° C) of free DNA	68.3	66.0	68.4	63.7	68.8	69.4
KA2115	14.3	14.9	13.6	10.5	5.9	4.4
KA2040	10.1	9.0	11.6	4.5	3.9	2.4
KA2034	9.8	9.2	9.4	6.8	6.5	6.0
KA2035	9.7	9.5	9.4	4.9	3.9	2.6
KA2041	9.7	10.5	9.7	7.9	5.0	4.9
KA2114	7.7	7.2	7.7	7.5	3.0	3.6
FH1028	18.6	18.5	19.1	14.8	12.9	12.6
FH1024	9.6	8.5	11.7	7.0	4.4	5.4
FH1026	7.7	6.7	8.0	4.0	2.6	4.2
KJK6162	5.7	6.3	4.6	3.9	3.6	3.6

Slightly stronger | Equal | Slightly weaker | Weaker | Much weaker

Note: The error of the ΔTm values are within 5%, based on experimental reproducibility.

Table 3

Equilibrium binding affinities of all PAs with DNAs that have mutated flanking sequences.

	Polyamide	λB (K_D/nM) AAGTGA A	SC1 (K_D/nM) AAGTGA G	GAGA (K_D/nM) GAGTGA A
	KA2115	0.49	0.27	1.67
	KA2040	0.54	0.68	1.82
	KA2034	0.67	0.83	0.66
	KA2035	1.30	0.91	1.36
	KA2041	1.62	2.70	2.80
	KA2114	4.10	7.70	7.20
	FH1028	0.16	18	0.51
	FH1024	1.65	No	0.14
	FH1026	35	No	43
	KJK6162	79	277	75

Note: No means no binding was detected between PA and the corresponding DNA. Errors for these equilibrium constant values are standard errors for the fitting and are within 2% for kinetic fitting, within 6% for steady state fitting.