Bioorganic & Medicinal Chemistry 21 (2013) 3907–3918

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Design, synthesis, and DNA binding characteristics of a group of orthogonally positioned diamino, *N*-formamido, pyrrole- and imidazole-containing polyamides

Sameer Chavda [a], Balaji Babu [a], Pravin Patil [a], Adam Plaunt [a], Amanda Ferguson [a], Megan Lee [a], Samuel Tzou [a], Robert Sjoholm [a], Toni Rice [a], Hilary Mackay [a], Joseph Ramos [b], Shuo Wang [b], Shicai Lin [b], Konstantinos Kiakos [c], W. David Wilson [b], John A. Hartley [c], Moses Lee [a,*]

[a] Division of Natural and Applied Sciences and Department of Chemistry, Hope College, Holland, MI 49423, USA
[b] Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA
[c] Cancer Research UK Drug-DNA Interactions Research Group, UCL Cancer Institute, Paul O'Gorman Building, 72 Huntley Street, London WC1E 6BT, United Kingdom

ARTICLE INFO

Article history:
Received 1 February 2013
Revised 26 March 2013
Accepted 2 April 2013
Available online 10 April 2013

Keywords:
Polyamides
Diamino
Minor groove
Sequence specificity
Core rules

ABSTRACT

Orthogonally positioned diamino/dicationic polyamides (PAs) have good water solubility and enhanced binding affinity, whilst retaining DNA minor groove and sequence specificity compared to their monoamino/monocationic counterparts. The synthesis and DNA binding properties of the following diamino PAs: f-I*P*I (**3a**), f-I*P*P (**4**), f-*P*IP (**5**), and f-P*P*P (**6**) are described. *P* denotes the site where a 1-propylamino group is attached to the N1-position of the heterocycle. Binding of the diamino PAs to DNA was assessed by DNase I footprinting, thermal denaturation, circular dichroism titration, biosensor surface plasmon resonance (SPR), and isothermal titration calorimetry (ITC) studies. According to SPR studies, f-I*P*I (**3a**) bound more strongly ($K_{eq} = 2.4 \times 10^8\ M^{-1}$) and with comparable sequence selectivity to its cognate sequence 5′-ACGCGT-3′ when compared to its monoamino analog f-IPI (**1**). The binding of f-I*P*I (**3a**) to 5′-ACGCGT-3′ via the stacked dimer motif was balanced between enthalpy and entropy, and that was quite different from the enthalpy-driven binding of its monoamino parent f-IPI (**1**). f-I*P*P (**4**) also bound more strongly to its cognate sequence 5′-ATGCAT-3′ ($K_{eq} = 7.4 \times 10^6\ M^{-1}$) via the side-by-side stacked motif than its monoamino analog f-IPP (**2a**). Although f-P*P*P (**6**) bound via a 1:1 motif, it bound strongly to its cognate sequence 5′-AAATTT-3′ ($K_{eq} = 4.8 \times 10^7\ M^{-1}$), 15-times higher than the binding of its monoamino analog f-PPP (**2c**), albeit f-PPP bound via the stacked motif. Finally, f-*P*IP **(5)** bound to its target sequence 5′-ATCGAT-3′ as a stacked dimer and it has the lowest affinity among the diamino PAs tested ($K_{eq} < 1 \times 10^5\ M^{-1}$). This was about two times lower in affinity than the binding of its monoamino analog f-PIP (**2b**). The results further demonstrated that the 'core rules' of DNA recognition by monoamino PAs also apply to their diamino analogs. Specifically, PAs that contain a stacked IP core structure bind most strongly (highest binding constants) to their cognate GC doublet, followed by the binding of PAs with a stacked PP structure to two degenerate AT base pairs, and finally the binding of PAs with a PI core to their cognate CG doublet.

1. Introduction

The DNA minor groove is an attractive target for the design and development of molecules able to specifically recognize predetermined DNA sequences.[1a,b] In the discovery of gene control agents, DNA minor groove binders have represented an attractive source of novel antitumor molecules. The increasing scientific attention directed towards this group of compounds stems from their ability to interact in a sequence-selective fashion across various DNA binding sites thus implying the possibility of targeting specific DNA sequences in the human genome. Such compounds are often based on natural products, which have been extensively investigated for their ability to selectively interact with the DNA minor groove.[1]

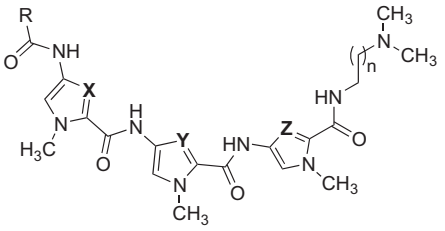
The antibiotics netropsin and distamycin A, along with their synthetic *N*-methylpyrrole (P) and *N*-methylimidazole (I) based analogs have been shown to bind within the DNA minor groove with a high degree of sequence specificity with affinities ranging from 10^5 to $10^8\ M^{-1}$.[2] The molecular recognition and binding affinities of such molecules follow a well-documented set of pairing rules confirmed by experimental and theoretical studies.[1,3,4] Polyamides (PAs) normally bind in a 2:1 fashion aligning themselves in

* Corresponding author. Tel.: +1 616 395 7190; fax: +1 616 395 7923.
E-mail address: lee@hope.edu (M. Lee).

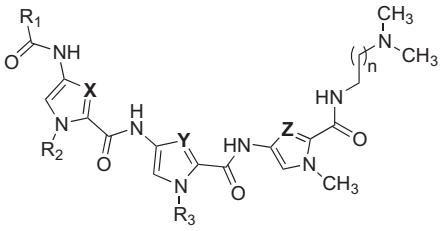
0968-0896/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2013.04.001

an anti-parallel fashion within the minor groove (tail to head and head to tail) to form stacked dimers which possess the ability to recognize specific DNA base pairs. A 'P' overlapped with another 'P' (a P/P pairing) recognizes A/T or T/A base pairs whilst an 'I' overlapped with P forming an I/P pairing recognizes a G/C base pair; a P overlapped with I (P/I pairing) recognizes a C/G base pair; and an I overlapped with I (I/I pairing) recognizes T/G or G/T mismatched pairs and tolerates G/C or C/G base pairs.[5a,b] These relationships have been elegantly accounted for by various experimental findings, which include NMR[5] and X-ray diffraction[6] studies.

Over time, further experimental investigations have led to modifications within the heterocyclic core structure and at the terminal ends of PAs. Such modifications have led to biophysical investigations on modified N-terminus PAs.[1,2,4b,7] In comparison to non-modified PAs, it has been ascertained within the authors' laboratories that those PAs with an N-terminus formamido substituent (f) exhibit increased sequence specific binding within the DNA minor groove along with a staggered binding mode where the stacked heterocyclic units are off-set by one unit.[8] A prime example of this would be the PAs f-IPI (**1**), f-IPP (**2a**), f-PIP (**2b**) and f-PPP (**2c**) depicted in Figure 1. PAs, such as f-IPI (**1**) and f-IPP (**2a**) bind more strongly to their cognate DNA sequences (5′-ACGCGT-3′ and 5′-ATGCAT-3′, respectively) than their non-formamido trimer congeners.[2,9] Furthermore, molecular docking studies on the binding of f-IPI (**1**) to 5′-ACGCGT-3′ carried out within the authors' laboratories showed that the formamido substituent rotated about the carboxamide C–N bind forming specific hydrogen bonds with electron-rich sites within the floor of the minor groove.[4c] Other important modifications to PAs which have shown to be of significance in increasing binding affinity and sequence selectivity include (i) increasing of the size and content of cationic groups present at C-terminus[9] and (ii) the order of the core heterocycles. To this end, research from the author's laboratory has ascertained that the binding affinities rank in the following order when considering the arrangement of the stacked heterocycles in triamide PAs, that is, the 'core rules': IP > PP ≫ PI ⩾ II.[4b] In accordance to these rankings, **1** was found to bind to its cognate sequence 5′-ACGCGT-3′ (in an anti-parallel, staggered stacked mode) with a higher affinity (10-fold) and selectivity than that of distamycin for its cognate sequence 5′-AAATTT-3′.[4b,7b]



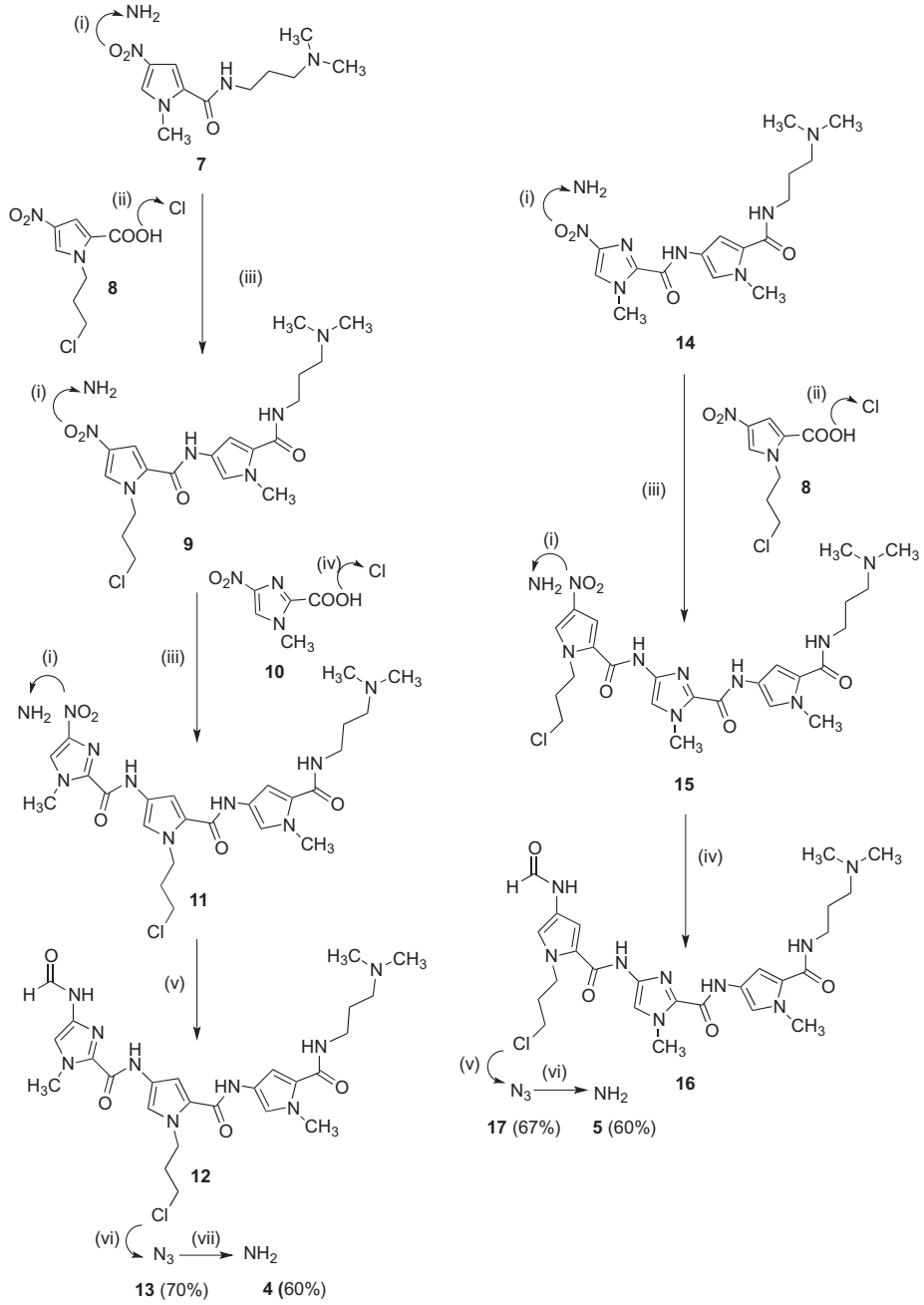
R=H, X=N, Y=CH, Z=N, n=1: f-IPI;1
R=H, X=N, Y=CH, Z=CH, n=2: f-IPP; 2a
R=H, X=CH, Y=N, Z=CH, n=2: f-PIP; 2b
R=H, X=CH, Y=CH, Z=CH, n=2: f-PPP; 2c
R=Ph, X=CH, Y=CH, Z=CH, n=2:PhIPI; 2d



R_1=H,R_2=Me R_3=-$(CH_2)_3NH_2$, X=N, Y=CH, Z=N, n=1: f-I*P*I;3a
R_1=H, R_2=-$(CH_2)_3NH_2$, R_3=Me, X=N, Y=CH, Z=N, n=1; f-*I*PI:3b
R_1=Ph, R_2=Me, R_3=-$(CH_2)_3NH_2$, X=N, Y=CH, Z=N, n=1; Ph-I*P*I:3c
R_1=H,R_2=Me, R_3= -$(CH_2)_3NH_2$, X=N, Y=CH, Z=CH, n=2: f-I*P*P; 4
R_1= H, R_2= -$(CH_2)_3NH_2$, R_3=Me, X=CH, Y=N, Z=CH, n=2: f-*P*IP; 5
R_1=H, R_2=Me,R_3 -$(CH_2)_3NH_2$, X=CH, Y=CH, Z=CH, n=2: f-P*P*P; 6

Figure 1. (A) Non-orthogonal trimers **1** and **2a–d** along with orthogonal diamino PAs and non-formamido tetramers **3–6**.

PAs such as **1**, **2d** and **3a–c** (Fig. 1), which contain the f-IPI core structure, target the 5′-ACGCGT-3′ sequence, which occurs within the central sequence of the Mlu1 cell cycle box (MCB), a transcriptional element found in the promoter of the human DBf4 gene. This gene codes for a regulatory subunit of Cdc7 (cyclin independent 7) kinase. High levels of this kinase have been implicated in the development of various cancers.[10] In an effort to improve the sequence specificity, solubility in biological media, cell penetration and nuclear localization properties of **1**, we have recently reported the synthesis and biophysical characteristics of the diamino formamido PA f-I*P*I (**3a**) (*P* denotes the site where the orthogonal alkylamino group is attached), f-*I*PI (**3b**) and the non-formamido variant PhI*P*I (**3c**). We also incorporated an additional amino group (cationic at physiological pH) into **2a–c** after careful consideration that multiple cationic groups may gravitate to AT rich sequences arising from attraction to the negative molecular electrostatic potential within the minor groove of A/T rich sequences.[11] In addition, we chose to position the amino group at an optimum distance from the other amino/cationic sites on **3a–c** (at the end of a pyrrole-N1-alkyl group) so as to minimize the possibility of the cationic charges repelling one another which could potentially lead to disfavored binding orientations thus compromising binding affinity. To this end, Satz and Bruice have elegantly reported that AT-sequence targeting PAs containing a pyrrole-N1-alkyl spermine/spermidine group target A/T rich sequences with higher binding affinities compared to their congeners lacking an alkylamino groups.[12] Furthermore, it was also shown experimentally that the alkylammonium side chain is capable of forming electrostatic bonds with the negatively charged phosphodiester backbone of DNA. In line with these observations, our investigations on the binding and selectivity of diamino/dicationic **3a** to its cognate sequence 5′-ACGCGT-3′ ascertained, through biosensor surface plasmon resonance (SPR) studies, a binding constant ~4 times higher than that for the monocationic **1** ($2.4 \times 10^8\ M^{-1}$ for **3a** vs $5.4 \times 10^7\ M^{-1}$ for **1**).[11] In support of this result, the binding affinity of the non-formamido diamino/dicationic congener **3c** showed a ~threefold enhancement over its monocationic analog **2d** ($1.5 \times 10^3\ M^{-1}$ for **3c** vs $4.8 \times 10^6\ M^{-1}$ for **2d**) upon binding to 5′-ACGCGT-3′.[13] In contrast, the recently reported orthogonal diamino congener f-*I*PI **3b** (containing the aminopropyl chain attached to the N-terminus imidazole) showed a slightly weaker affinity than **1** ($2.3 \times 10^7\ M^{-1}$ for **3b**) leading to the suggestion that the position of the aminopropyl group can affect binding affinity.[14] Based on an interest to expand our research on diamino PAs, and based upon the rationale used for the design of **3a**, herein we report the design, synthesis and biophysical evaluation of a group of formamido PAs containing a pyrrole-N1 substituted with a propylamine chain (Scheme 1), namely f-I*P*P **4**, f-*P*IP **5** and f-P*P*P **6** (see Fig. 1) The binding affinities, as well as the minor groove and sequence specific properties of **3–6** were studied against their cognate and non-cognate sequences using thermal melts (ΔT_M), circular dichroism (CD), surface plasmon resonance, isothermal titration calorimetry (ITC) and DNase I footprinting, and the results were compared to their respective *N*-formamido trimers **2a–c**.



Scheme 1. Synthesis of f-I*P*P (**4**) and f-*P*IP (**5**). Synthesis of **4**: Reagents and conditions: (i) 5% Pd-C, cold MeOH; (ii) $SOCl_2$, reflux, 15 min; (iii) dry Et_3N, dry DCM; (iv) acetic formic anhydride; (v) NaN_3, dry DMF, 80 °C, 12 h; (vi) 10% Pd/C, cold MeOH. Synthesis of **5**: Reagents and conditions: (i) 5% Pd-C, cold MeOH; (ii) $SOCl_2$, reflux, 15 min; (iii) dry Et_3N, dry DCM; (iv) acetic formic anhydride; (v) NaN_3, dry DMF, 80 °C, 12 h; (vi) Ph_3P in THF/H_2O 10:1, 3 days rt.

2. Results and discussion

2.1. Chemistry

The synthesis of diamino f-I*P*I (**3a**) was communicated earlier.[15,16] The approach employed in the synthesis of **3a** and related PAs[17] was applied to the preparation of diamino PAs **4** and **5** (Scheme 1). For the synthesis of **4**, reduction of the nitro group of monoamide **7** by catalytic hydrogenation over 5% Pd-C, and coupling of the amino intermediate with the acid chloride of *N*-(3-chloropropyl)-4-nitropyrrole-2-carboxylic acid (**8**), produced from reaction with thionyl chloride, gave the diheterocylic intermediate O_2N-*P*P-*N*,*N*-dimethyl-2-aminoethylamine **9** in 77% yield. Catalytic hydrogenation of PA **9** gave an intermediate amine and that was immediately coupled with 4-nitro-1-*N*-methylimidazole -2-carbonyl chloride, formed by refluxing the carboxylic acid **10** in oxalyl chloride and dry THF for 45 min. Triamide **11** was obtained in 48% yield, and it was transformed to the *N*-formamido compound by first reducing the nitro group to an amino using catalytic hydrogenation. Subsequent reaction of the amine with acetic formic anhydride gave the desired PA **12** in 44% yield. S_N2 displacement of the chloride with sodium azide gave intermediate **13** in 70% yield. Catalytic hydrogenation of azide **13** gave the target diamino PA f-I*P*P (**4**) in 60% yield. The synthesis of diamino f-*P*IP (**5**) is also depicted in Scheme 1 and it follows the same strategy used for the synthesis of diamino PA **4**. Starting with NO_2-IP-*N*,*N*-

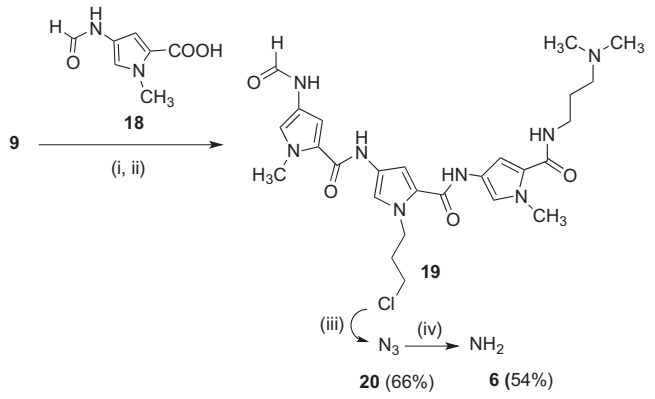
dimethyl-2-aminopropylamine **14**,[4b,7b,8] reduction of the nitro group, followed by coupling of the amine with the acid chloride derived from carboxylic acid **8** gave triamide **15** in 66% yield. Subsequent conversion of the nitro-moiety to the *N*-formamido compound **16** was achieved in 66% yield. S_N2 azidation gave PA **17** in 67%, and Staudinger reduction gave the desired diamino PA **5** in 60% yield.

Synthesis of f-P*P*P (**6**) (Scheme 2) began with reduction of the nitro group of NO_2-*P*P **9** by catalytic hydrogenation. The resulting amine was coupled to 4-*N*-formamido-1-methylpyrrole-2-carboxylic acid **18**[18] in the presence of PyBOP to afford the desired chloride **19** in 38%. Azidation of PA **19** gave the desired product **20** in 66% yield. Subsequent Staudinger reduction of the azido group in compound **20** furnished the target compound f-P*P*P (**6**) in 53%. The structures of all the PAs and their precursors were confirmed by IR, NMR and mass spectrometry (LRMS and HRMS). It is worthy to note that the target diamino compounds **3a**, **4**–**6** have appreciable solubility in water. As hydrochloride salts, these compounds readily dissolved in water to prepare solutions at 1 mM concentration.

2.2. DNA footprinting

The sequence specificity of diamino PAs **3a**, **4**–**6** was investigated by DNA footprinting studies using 5′-[32P]-radiolabeled engineered DNA fragments containing the DNA sequences 5′-ATGCAT-3′, 5-TACGAT-3′, and 5′-AAATTT-3′ (Fig. 2).[7] These are cognate sequences for f-IPP (**2a**) and f-I*P*P (**4**),[4b,8] f-PIP (**2b**) and f-*P*IP (**5**),[4b] and f-PPP (**2c**) and f-P*P*P (**6**),[4b] respectively. As previously reported DNase I footprinting of f-I*P*I (**3a**) showed that it bound strongly and selectively to its cognate sequence, 5′-ACGCGT-3′.[15]

The gel in Figure 2A shows that at 1 μM f-I*P*P (**4**) a footprint begins to appear for the cognate sequence 5′-ATGCAT-3′. However, it was also found to bind 5′-ACGCGT-3′ and 5′-ACTAGT′3′ indicating that this molecule has limited DNA sequence selectivity. This result is in contrast to that reported for f-I*P*I (**3a**), which demonstrated a preference for the cognate sequence 5′-ACGCGT-3′ at 0.05 μM.[15] Figure 2B shows the footprinting autoradiograph for diamino PA f-*P*IP (**5**) which binds to 5′-TACGAT-3′, immediately adjacent to the 5′-flank of ICB2 (5′-***TACGAT***TGG-3′, ICB2 underlined) of the topoisomerase IIα promoter.[4b,19] At 10 μM f-*P*IP (**5**), a clear footprint is evident at the target sequence. At 30 μM a secondary footprint is evident at the 3′-flank of ICB1 (5′-CTGCTTC-3′), thus indicating a three-fold selectivity for the cognate sequence. For comparison, f-PIP (**2b**) also produced a footprint at 30 μM indicating a threefold enhancement in binding affinity by the second amino group while maintaining sequence selectivity. Interestingly, an H-pin of f-PIP (f-P_LIP in which the stacked PAs were connected by an ethylene glycol linker at P_L) was found to give a similar sequence selectivity and affinity to f-PIP (**2b**).[20] However, unlike f-PIP, which could enter cells and traffic into the nucleus and modulate the expression of the topoisomerase IIα gene in confluent cells in culture[19] the H-pin analog had no effects on cells.[20] We rationalized that the H-pin was too large in molar mass and was unable to enter the cell and localize in the nucleus. This was part of the rationale for the development of diamino PA analogs, which are small, more soluble in water, and should more readily concentrate in the nucleus.



Scheme 2. Synthesis of f-P*P*P (**6**): Reagents and conditions: (i) 5% Pd-C, cold MeOH; (ii) acid **18**, PyBOP; (iii) NaN_3, dry DMF, 80 °C, 12 h; (iv) Ph_3P in THF/H_2O 10:1, 3 days rt.

Figure 2C shows the DNA footprinting gel for f-P*P*P (**6**). The radiolabelled oligonucleotide was cleaved with micrococcus nuclease and generated a footprint at 5 μM. At 30 μM, no other site of cleavage inhibition was observed; thus, demonstrating the specificity of the tripyrrole heterocyclic core for 5′-AAATTT-3′. For comparison, the results of f-PPP (**2c**) (data not shown)[21a,b] showed that a footprint at 5′-AAATTT-3′ emerged at 20 μM, which was four-times lower in affinity than f-P*P*P (**6**). Overall, the footprinting studies revealed the following results: First, the diamino PAs bind strongly to their cognate sequences. f-I*P*I (**3a**), f-*P*IP (**5**), and f-P*P*P (**6**) showed a 3- to 4-fold enhancement in binding affinity over their monoamino counterparts, **1**, **2b**, and **2c**, respectively. Second, except for f-I*P*P (**4**), the sequence selectivity of f-I*P*I (**3a**), f-*P*IP (**5**), and f-P*P*P (**6**) was retained.

2.3. Thermal denaturation

The binding of *N*-formamido diamino PAs **3a** and **4**–**6** was tested using the DNA sequences given in Table 1 in thermal denaturation experiments. An excess of each PA to DNA (3–1 μM, respectively) was used for these studies and in all cases, and the concentration of DNA was identical. The following DNA sequences were examined: 5′-ACGCGT-3′, the cognate sequence for f-I*P*I (**3a**) found in the core sequence of the *Mlu1* cell-cycle box, 5′-ATGCAT-3′, the cognate sequence for f-I*P*P **4**, 5′-ATCGAT-3′, the cognate sequence for f-*P*IP (**5**), degenerate with the 5′-TACGAT-3′ of the 5′-flank to the ICB2 site of the human topoisomerase IIα promoter,[19] and 5′-AAATTT-3′, the cognate site for distamycin A and that anticipated for f-P*P*P (**6**). These thermal denaturation experiments screened the preferred binding sequences of **3a** and **4**–**6** by measuring their ability to stabilize double stranded (ds) DNA from thermal denaturation upon binding. ΔT_M values were ascertained from the difference in melting temperature of the DNA–PA complexes and duplex DNA (Table 1). For comparative purposes, the thermal denaturation data for the non-diamino variants of **3a** (also including the non-formamido variants) and **4–6** reported previously are also shown in Table 1. It is also worth noting that the reported ΔT_M value for distamycin A against its cognate sequence 5′-AAATTT-3′ is 14 °C.[7b] The data presented in Table 1 show that diamino PAs f-I*P*I (**3a**), f-I*P*P (**4**) and f-P*P*P (**6**) bind to their cognate sequences and stabilize the DNA much more strongly (ΔT_M of 24, 26 and 25 °C, respectively) relative to their non-diamino counterparts **1**, **2a** and **2c** (ΔT_M <12 °C). This is an indication that the additional diamino group can contribute favorable stabilization properties to these PAs when bound to their cognate sequences.

2.4. Circular dichroism (CD) titrations

Results from CD studies provided evidence for the minor groove binding motif of f-I*P*I (**3a**) (Fig. 3A), f-I*P*P (**4**) (Fig. 3B), f-*P*IP (**5**) (Fig. 3C), and f-P*P*P (**6**) (Fig. 3D) to their respective cognate DNA sequences. The appearance of strong and positive induced CD bands at 330 nm is consistent with minor groove binding for

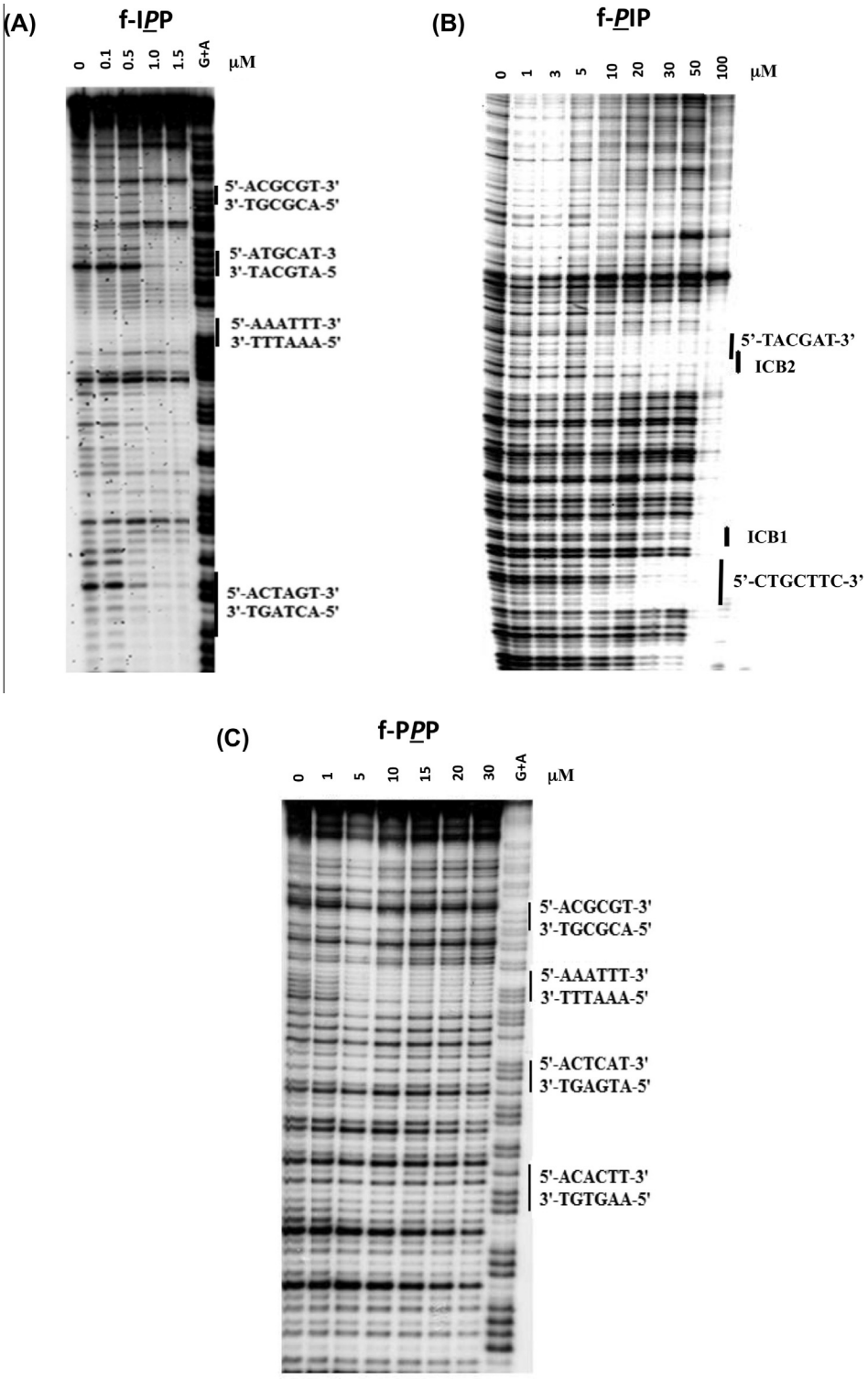


Figure 2. DNA footprinting studies of diamino compounds: (A) f-I*P*P (**4**), (B) f-*P*IP (**5**), and (C) f-P*P*P (**6**). (A) and (B) use cleavage by DNase I and (C) by micrococcal nuclease. All reactions contain 1000 cps DNA fragment, 10 mM Tris, pH 7, 1 mM EDTA, 50 mM KCl, 1 mM $MgCl_2$, 0.5 mM DTT, 20 mM Hepes. Figure 2B shows only the resolved portion of the gel in order to highlight the footprints. There were no artifacts in the full gel (see Supplementary data).

PAs.[6–10,12–17,19] The presence of an isodichroic point at ~310 nm provides evidence that the diamino PAs bind to the DNA sequence by a single mechanism, presumably by interacting in the minor groove as a stacked dimer.[6b,13–17,19] This binding mode is similar to what has been reported for binding of the monoamino counterparts f-IPI (**1**), f-IPP (**2a**), f-PIP (**2b**)[19] and f-PPP (**3**) to their cognate sequences.[4b,7b]

2.5. Biosensor-surface plasmon resonance (SPR)

The binding constants for the diamino PAs f-I*P*I (**3a**), f-I*P*P (**4**), f-*P*IP (**5**), and f-P*P*P (**6**) for their respective cognate and non-cognate sequences were determined by SPR studies. The sensorgrams for binding of these diamino PAs to their respective sequences are given in Figure 4, and the binding constants derived from these

Table 1
Thermal denaturation and SPR binding constants for PAs **1**–**6**

PA	5′-ACGCGT-3′		5′-ATGCAT-3′		5′-AAATTT-3′		5′-ATCGAT-3′	
	ΔT_M[a] (°C)	K_{eq}[b] (M^{-1})	ΔT_M (°C)	K_{eq} (M^{-1})	ΔT_M (°C)	K_{eq} (M^{-1})	ΔT_M (°C)	K_{eq} (M^{-1})
f-IPI, **1**[7b,15]	9	5.4×10^{7}[c]	9	1.2×10^{7}[d]	1	5.3×10^{4}	1	1.0×10^{5}[d]
f-I<u>P</u>I, **3a**[15]	>24	2.4×10^{8}	24	3.5×10^{7}	10	-	11	1.7×10^{6}
f-<u>I</u>PI , **3b**[14]	>20	2.3×10^{7}	—	—	11	2.2×10^{6}	—	—
Ph-I<u>P</u>I, **3c**[13]	>25	1.5×10^{7}	—	2.1×10^{6}	8	NR	—	—
Ph-IPI, **2d**[13]	9	4.8×10^{6}	—	2.4×10^{5}	0	NR	—	—
f-IPP, **2a**[4b,8]	2	8.8×10^{4}[d]	11	1.2×10^{7}[d]	—	—	3	9.4×10^{4}[d]
f-I<u>P</u>P, **4**	0	2.3×10^{6}	26	1.3×10^{8}	0	—	7	2.7×10^{6}
f-PPP, **2c**[8]	—	—	—	—	9	3.2×10^{6}[d]	1	—
f-P<u>P</u>P, **6**	0	—	—	—	25	4.8×10^{7}	10	—
f-PIP, **2b**[4b,17]	3	—	—	—	—	—	0	2.0×10^{5}
f-<u>P</u>IP, **5**	0	—	—	—	10	—	0	<1.0×10^{5}[e]

Cognate sequences are typed in bold face. '—' Not determined. NR = no response. References are superscripted. All DNAs used in the studies were hairpin oligonucleotides. For DNA melt experiments, 1.0 μM of DNA, 3.0 μM of PA, and a $PO_4$0 buffer at pH 6.2 were used. The concentration of Na^+ was 12.5 mM. The SPR experiments were performed in cacodylic acid buffer (10 mM cacodylic acid, 100 mM NaCl, 1 mM EDTA, pH 6.25) with 0.005 % v/v Surfactant P20, and the data have a ±20% experimental error in K_{eq}. The fit values for any experiment have a lower error but the experimental value is what observed for overall reproducibility.

[a] $\Delta T_M = T_M$ (bound DNA) $- T_M$ (free DNA).

[b] $K_{eq} = (K_1 \times K_2)^{1/2}$, determined at 25 °C.

[c] Reported in Ref. 15. It was reported earlier at 1.9×10^{8} M^{-1} (Ref. 4b,7b).

[d] Determined using 11 bp hairpin oligonucleotides.

[e] K <1.0×10^{5} M^{-1}, too weak to be accurately evaluated.

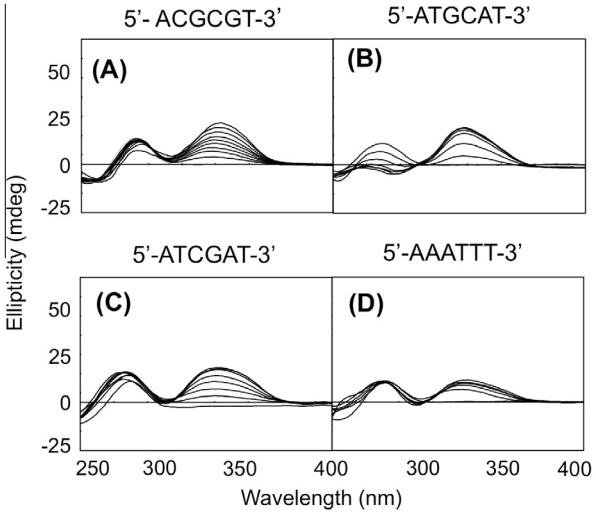


Figure 3. CD titration studies of diamino compounds: (A) f-I<u>P</u>I (**3a**), (B) f-I<u>P</u>P (**4**), (C) f-<u>P</u>IP (**5**), and (D) f-P<u>P</u>P (**6**). CD experiments were carried out using 160 μL of a 9 μM DNA solution, which was titrated with 1 mol equiv of each PA, past the point of saturation. The experiments were performed in a 10 mM phosphate buffer (1 mM EDTA, 50 mM [Na^+], pH 6.2). Data was obtained using an OLIS DSM20 spectropolarimeter.

studies are given in Table 1. The results are compared to their monoamino counterparts [f-IPI (**1**), f-IPP (**2a**), f-PIP (**2b**) and f-PPP (**3**)] as well as the PhIPI (**2d**).[13] In addition, the DNA binding affinity of a non-formamido diamino tetraamide PhI<u>P</u>I (**3c**)[13] as well as the imidazole-derivatized diamino f-<u>I</u>PI (**3b**)[14] are also included in the table for comparison purposes.

The results are interesting: First, the f-IPI diamino analogs **3a**–**c** bound more strongly to their cognate sequence 5′-ACGCGT-3′ than their non-cognate counterparts. As reported earlier f-I<u>P</u>I (**3a**) bound more strongly than its parent f-IPI (**1**).[15] A model for the binding of f-I<u>P</u>I bound to 5′-ACGCGT-3′ is given in Figure 5A. Similarly, PhI<u>P</u>I (**3c**) bound more strongly and as selectively to ACGCGT than its monoamino counterpart PhIPI (**2d**).[13] PA **1**, however, binds twofold more strongly than f-<u>I</u>PI (**3b**), indicating that the position of the second amino group can have some effect on binding affinity, which could be a result of the repulsive positive charges between the C-terminus quaternary ammonium group and the propyl ammonium species in the opposing stacked PA.[14] Second, the results for f-I<u>P</u>P (**4**) and f-P<u>P</u>P (**6**) provided further evidence that addition of a second aminoalkyl group could enhance the binding affinity over their monoamino counterparts. Specifically, f-I<u>P</u>P (**4**) and f-P<u>P</u>P (**6**) bound to their respective cognate sequences 18 and 15-fold more strongly than their monoamino counterparts f-IPP (**2a**) and f-PPP (**2c**). However, even though f-I<u>P</u>P (**4**) bound through the stacked side-by-side motif, f-P<u>P</u>P (**6**) bound as a monomer, that is, 1:1 ligand to DNA, according to the SPR results. These findings are not surprising because the minor groove of GC-containing sequences is typically wider than AT-rich sequences, and it can accommodate the binding of stacked PAs in a positively cooperative manner.[4b,8] Also, earlier studies on the binding of distamycin to the minor groove of four base-pair AT-rich sequences by NMR and X-ray crystallography studies revealed a consistent 1:1 binding motif.[21] Furthermore, distamycin, and f-PPP (**2c**)[4b,8] were later shown to bind in the longer AT minor groove at 5′-AAATTT-3′ as a stacked dimer, but with negative cooperativity.[22] This indicated that binding of the first molecule hindered the binding of the second molecule presumably due to the narrow minor groove.[23] In the case of f-P<u>P</u>P (**6**) and according to results from SPR studies, the second alkylamino group on <u>P</u> hindered the ability of the molecules to bind in a stacked motif. These results suggest that care must be taken in designing *N*-alkylamino PAs with regard to the composition of pyrrole and imidazole units in the context of sequence recognition as well as the relative position of the *N*-propylamino moiety. Models for the binding of f-I<u>P</u>P (**4**) and f-P<u>P</u>P (**6**) to their cognate sequences, 5′-ATGCAT-3′ and 3′-AAATTT-3′, are depicted in Figure 5B and D, respectively.

Third, the binding of f-<u>P</u>IP (**5**), in which the alkylamino group is located at the N-terminus end of the molecule, to its cognate sequence is significantly lower than for its monoamino counterpart f-PIP (**2b**). This result is consistent with that observed for f-<u>I</u>PI (**3b**) and f-IPI (**1**). It tells us that addition of an alkylamino group to the N-terminus region of the molecule is less preferred than derivatizing the middle section of the molecule. In the case of f-<u>P</u>IP the SPR sensorgrams illustrate the very weak binding to its

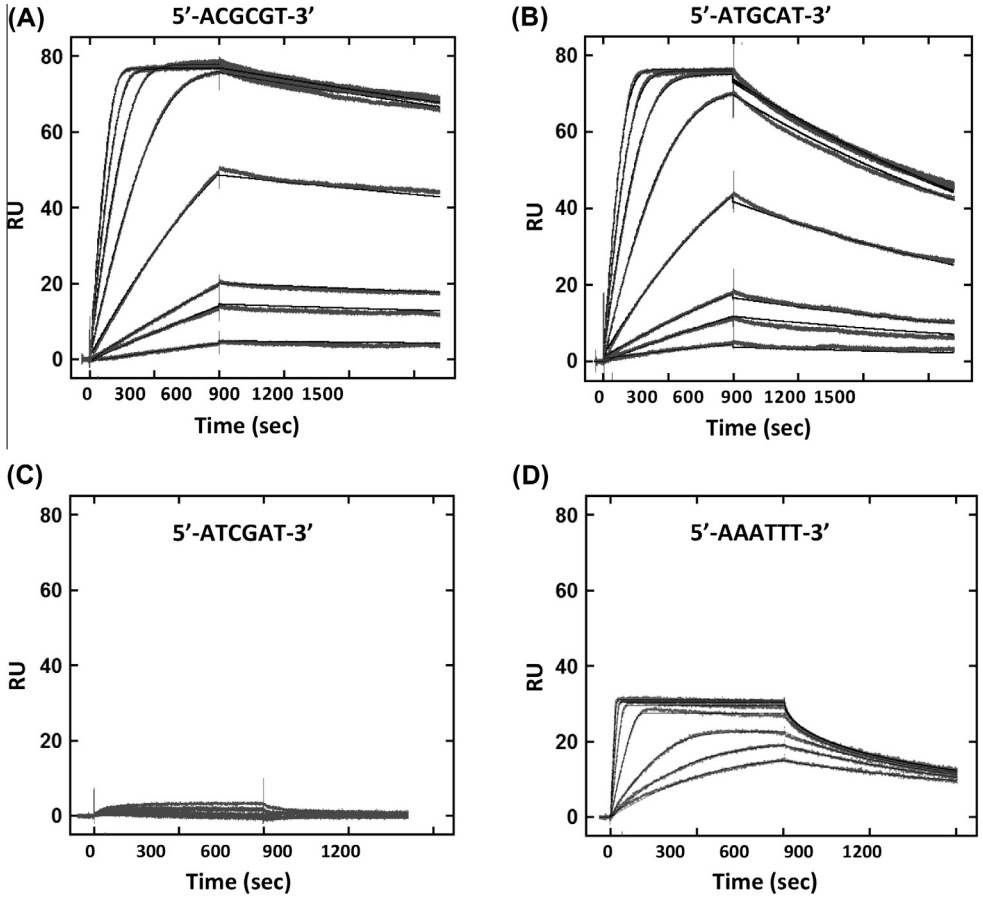


Figure 4. Biosensor SPR sensorgrams for the binding of the diamino PAs to their respective cognate sequences: (A) f-I<u>P</u>I (**3a**), (B) f-I<u>P</u>P (**4**), (C) f-<u>P</u>IP (**5**), and (D) f-P<u>P</u>P (**6**). The concentrations in both experiments from the bottom to the top sensorgrams are 4, 8, 10, 20, 40, 60, 80 and 100 nM.

cognate sequence 5′-ATCGAT-3′ (Fig. 5C). One explanation for this observation is the possibility of unfavorable electrostatic clash between the N-terminus alkylamino charged group and the C-terminus dimethylaminoalkyl group in the stacked binding motif. It is evident that small changes in the structure of the PAs can have dramatic effects on DNA binding properties, including affinity, sequence selectivity, as well as binding motif. In the case of f-<u>I</u>PI (**3b**) and f-<u>P</u>IP (**5**) attachment of the propylamino moiety at the N-terminus heterocycle negatively impacted binding affinity when compared to diamino analogs in which the alkylamino unit was attached to the central heterocycle, such as, f-I<u>P</u>I and f-I<u>P</u>P.

Lastly, the trend of binding constants of the diamino PAs is consistent with those of the monoamino counterparts.[4b,7b] Specifically, PAs bearing an IP core of stacked heterocycles bind to a GC doublet with the most favorable affinity (ca. 10^8 M^{-1}). Next are PAs that contain a PP core that recognizes a degenerate (A/T)(A/T) doublet with affinity ca. 10^7 M^{-1}. As expected PAs bearing a stacked PI core that bind a CG doublet exhibit the weakest binding affinity in the range of 10^5 M^{-1}. Due to the mass transfer in the association phase and the very slow dissociation of some PAs, such as **3a**, it is impossible to present a meaningful comparison of the rate constants (k_{on} and k_{off}). Therefore, only the equilibrium constants are shown in Table 1.

2.6. Isothermal titration calorimetry (ITC)

ITC experiments were performed on the binding of **4**, **5** and **6** and compared to published enthalpy results (Table 2). The binding of **3a** to the cognate DNA sequence was strongly exothermic.[15] Using the binding constant (K_{eq}) for **3a** and 5′-ACGCGT-3′ determined by SPR studies, the free energy of interaction (ΔG) was calculated to be -11.4 kcal mol^{-1}. With a measured enthalpy of -5.9 kcal mol^{-1}, the entropy of reaction or $T\Delta S$ was determined to be 5.5 kcal mol^{-1}. For comparison, the reported thermodynamic data for binding of the parent f-IPI (**1**) to 5′-ACGCGT-3′ were the following: K_{eq} was 5.4×10^7 M^{-1}; ΔG was -10.5 kcal mol^{-1}; ΔH was -7.6 kcal mol^{-1}; and $T\Delta S$ was 2.9 kcal mol^{-1}.[15] These results are striking: despite having only a slight increase in the binding affinity of f-I<u>P</u>I (**3a**) over f-IPI (**1**) for 5′-ACGCGT-3′, their thermodynamic profiles were very different. Binding of f-IPI (**1**) to its cognate sequence was dominated through enthalpy. However, binding of its diamino counterpart (**3a**) to the same DNA sequence was more balanced between enthalpy and entropy. The binding enthalpy of **3a** is less negative and less favorable than that for **1**, while **3a** has a more favorable entropy which compensates its enthalpy loss and provides an overall increased affinity relative to **1**. To completely understand the basis for this effect we will require additional mono- and dicationic PAs. A likely cause of the decrease in enthalpy, however, is that the central side chains on dimer complexes, such as **3a**, bring some charge repulsion and steric clash on binding to reduce the favorable enthalpic interaction terms. Such a change would yield additional flexibility of the complex and a resulting compensating increase in entropy. On contrast, the terminal side chains on **3b** dimer complex have less charge repulsion and steric clash, and can further stabilize the dimer formation through interacting with the DNA backbone. This tight

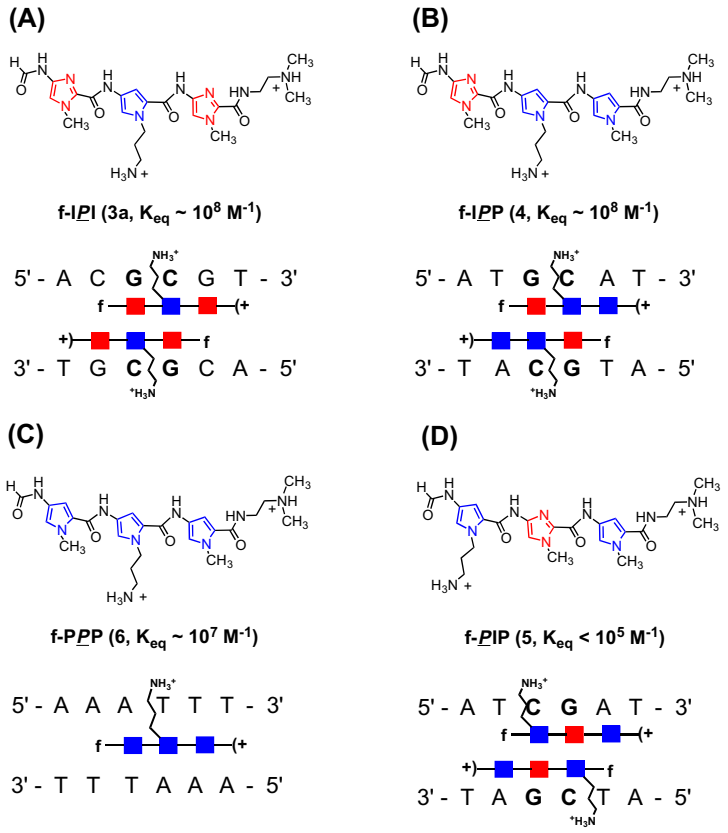


Figure 5. Models for the binding of (A) f-IPI (**3a**), (B) f-IPP (**4**), (C) f-PPP (**6**), and (D) f-PIP (**5**) with their respective cognate DNA sequence according to results from SPR studies.

Table 2
Thermodynamic profile for PAs **1**–**6** with their cognate DNAs

	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	TΔS (kcal mol^{-1})
5′-ACGCGT-3′			
f-IPI, **1**[15]	−10.5	−7.6	2.9
f-IPI, **3a**[15]	−11.4	−5.9	5.5
f-IPI , **3b**[14]	−10.0	−8.4	1.6
Ph-IPI, **3c**[13]	−9.8	−2.9	6.9
Ph-IPI, **2d**[13]	−9.1	−2.4	6.7
5′-ATGCAT-3′			
f-IPP, **2a**[8]	−9.4	–	–
f-IPP, **4**	−11.1	−6.1	5.0
5′-AAATTT-3′			
f-PPP, **2c**[8]	−8.9	−4.5	4.4
f-PPP, **6**	−10.5	−3.2	7.3
5′-ATCGAT-3′			
f-PIP, **2b**[17]	−7.2	–	–
f-PIP, **5**	<−6.8[a]	–	–

"–" Not determined.
[a] K <1.0 × 10^5 M^{-1}, too weak to be accurately evaluated.

binding of **3b** results in less favorable entropy compared to **1**. PA **4** is similar in thermodynamic profile (see Fig. 6A) to **3a** and binding is balanced between enthalpy and entropy. The binding of PA **5** is too weak to allow accurate enthalpy determination. The enthalpies of PAs f-PPP (**2c**) and its diamino counterpart f-PPP (**6**) (see Fig. 6B) binding to 5′-AAATTT-3′ as monomer complexes are much lower than those for PAs **1** and **3a** (Table 2). This is in agreement with low enthalpy values for other minor groove binders in A-tract sequences which resulted from the displacement of strongly bound water.[24] Binding of f-PPP (**2c**) to its cognate sequence was balanced between enthalpy and entropy, while the binding of dicationichas a less favorable enthalpy and more favorable entropy giving a higher affinity compared to **2c**. The thermodynamic profiles thus change in a very similar manner as PAs **3a** and **1**, but due to the monomeric binding mode the influence of the side chain is smaller than that for the dimer complex (compare **3a**–**1**). The dicationic derivatives are also more effective in displacing cations from the DNA than monocations and under low salt conditions this results in a higher binding entropy for the dications.[25] The results show that small modifications can directly impact how PAs bind their cognate sequences and the thermodynamics of interaction.

Overall, the results described in this study provide evidence that the N-position of pyrroles and imidazoles are attractive sites for further derivatization as a venue to improve the physiochemical and DNA binding properties of PAs. The molecular design described in this study continues to enable our team to create novel diamino PAs that have increased water solubility and binding affinity, whilst keeping the molecules small, that is, low molar mass. One key finding in this study is that incorporation of the second amino group did not visibly compromised DNA sequence selectivity. The second key finding is that N1 modification at the central heterocycle improves binding affinity relative to the parent monocation (compare **3a**–**1**, and **2a**–**4**), while modification at the N-terminal heterocycle weakens binding with respect to the parent PA (compare **3b**–**1**, and **5**–**2b**). Studies to show that the water-soluble diamino PAs are readily taken up by cells and concentrate in the nucleus are in progress.

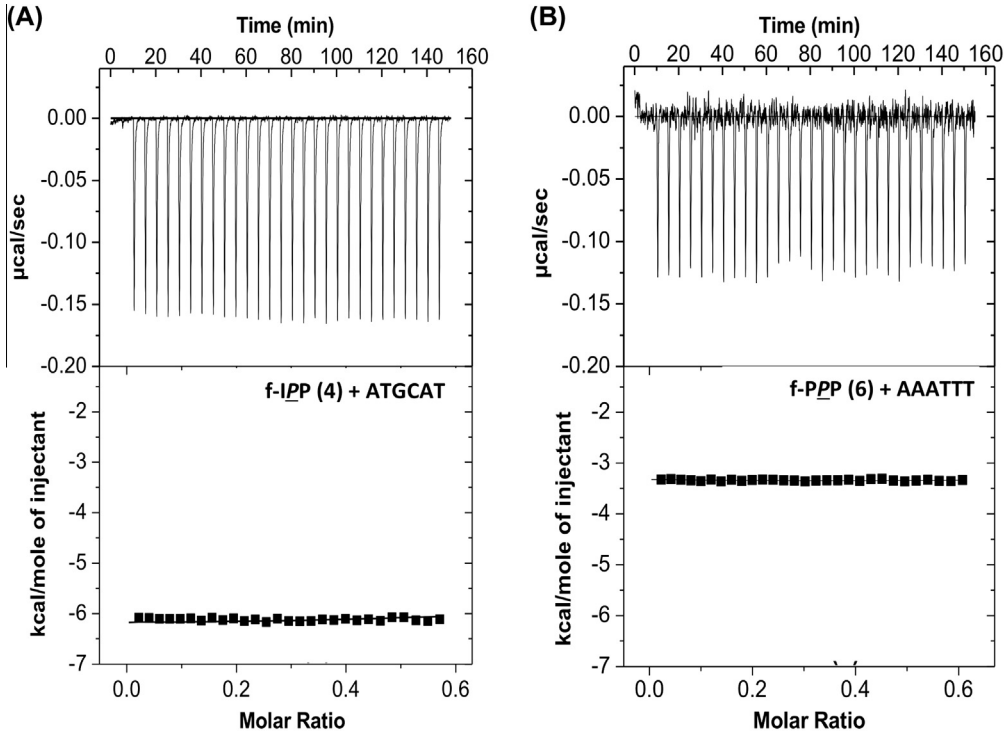


Figure 6. ITC thermograms (top) and integrated data (bottom) of (A) f-IPP (**4**) and (B) f-PPP (**6**) binding to their cognate DNA sequences.

3. Experimental

3.1. DNA footprinting

DNA footprinting experiments were conducted using 5′-[32P]-radiolabeled DNA fragments to determine the DNA sequence specificity of PAs **4–6**. For **4**, a 130-bp DNA fragment, engineered to contain the cognate sequence 5′-ATGCAT-3′ and the control, non-cognate sequences 5′-ACGCGT-3′, 5′-AAATTT-3′ and 5′-ACTAGT-3′ was utilized. For **5**, a 479-bp radiolabeled probe corresponding to the human topoIIα promoted and containing the ICBs 1 and 2 (at the resolvable portion of the gel) was utilized. Footprinting experiments using DNaseI were performed as previously described.[7] For **6**, a 130-bp DNA fragment, designed to contain the consensus sequence 5′-AAATTT-3′ and the control sequences 5′-ACGCGT-3′, 5′-ACTCAT-3′ and 5′-ACACTT-3′ was subjected to micrococcal nuclease digestions (due to the inefficient cleavage of DNaseI at AT-rich sequences) as described previously.[17]

3.2. Thermal denaturation

The synthetic DNA hairpins used in these studies were purchased from Operon (Huntsville, AL) with no HPLC purification: 5′-GA**ACGCGT**CG-CTCT-CGACGCGTTC-3′ (**5′-ACGCGT-3′**); 5′-GA**ATGCAT**CG-CTCT-CGATGCATTC-3′ (**5′-ATGCAT-3′**); 5′-CG**AAATTT**CC-CTCT-GGAAATTTCG-3′ (**5′-AAATTT-3′**); 5′-GA**ATCGAT**TG-CTCT-CAATCGATTC-3′ (**5′-ATCGAT-3′**). Thermal denaturation (ΔT_M) studies were performed in duplicate using a Cary Bio 100 spectrophotometer UV/vis instrument (Palo Alto, CA) and cells with a 10 mm pathlength using a previously reported method.[7,17] ΔT_M values were determined from the difference in melting temperatures of the DNA–PA complex and duplex DNA alone. 1.0 μM of DNA and 3.0 μM of PA were used in a $PO_4$0 buffer at pH 6.2 and containing a concentration of Na^+ of 12.5 mM.

3.3. Circular dichroism

CD spectroscopy was used to probe the binding of PAs **1–6** in the minor groove of double-stranded DNA using an OLIS (Bogart, GA) DSM20 spectropolarimeter with a 10 mm path length cuvette and a band pass of 2.4 nm.[7,17] These experiments were conducted in duplicate by titrating each PA with DNA solutions comprised of the DNA sequences tested in the thermal denaturation studies. In all cases, a fixed DNA concentration of 9 μM was used as were the ratios of PAs (1, 2, 3, 4, 5, 6, 8, and 10 molar equiv) titrated into the solution.

3.4. Surface plasmon resonance

SPR measurements were performed with a BIAcore T100 optical biosensor system (Biacore, GE Healthcare Inc.). The same DNA sequences used in the thermal denaturation experiments with 5′-biotion label were immobilized onto streptavidin—coated sensor chips (BIAcore SA) as previously described.[25] The SPR experiments were performed in filtered, degassed cacodylic acid buffer (10 mM cacodylic acid, 100 mM NaCl, 1 mM EDTA, pH 6.25) with 0.005 % v/v Surfactant P20. Solutions of known ligand concentration were injected through the flow cells. Compound solution flow was then replaced by buffer flow resulting in dissociation of the complex. The reference response from the blank cell was subtracted from the response in each cell containing DNA to give a signal (RU, response units) that is directly proportional to the amount of bound compound. Kinetic analysis was performed with multiple injections of different compound concentrations over the immobilized DNA surface at a flow rate of 10 μL/min and 25 °C. The number of binding sites and the binding constant were obtained from kinetic fitting sensorgrams using 2:1 binding with mass transfer kinetic model as previously described[7,17,26]: $r = (K_1 \times C_{free} + 2 \times K_1 \times K_2 \times C_{free}^2)/(1 + K_1 \times C_{free} + K_1 \times K_2 \times C_{free}^2)$ where r represents the moles of bound compound per mole of DNA hairpin duplex, K_1 and K_2 are macroscopic binding constants, and C_{free} is the free compound concentration in

equilibrium with the complex. The stoichiometry of the reaction, *n*, can be determined directly from an SPR experiment and from the RU of DNA immobilized, the maximum RU for the bound PA at steady-state, and the molecular-weight of DNA and the PA.

3.5. Isothermal titration calorimetry

ITC analysis was performed at 25 °C using a VP-ITC microcalorimeter (MicroCal, Northampton, MA) as previously reported.[4b,7] The concentrations of DNA sequences are 30 μM and those of PAs are 80 μM. Origin 7.0 was used to analyze the data, where the integration of each titration was plotted against ligand concentration. A linear fit was then employed and this was subtracted from the reaction integrations to normalize for non-specific heat components. Upon determining the ΔH, and using the ΔG calculated from K_{eq} obtained from SPR studies using the equation $\Delta G = \Delta H - T\Delta S$, the entropy term $T\Delta S$ was calculated. $\Delta G = -RT\ln K_{eq}$, $R = 1.987$ cal mol^{-1} K^{-1}.

3.6. Synthesis of diamino analogs

(f-I*P*I **3a**[15,16], f-I*P*P **4**, f-*P*IP **5**, and f-P*P*P **6**). Solvents and organic reagents were purchased from Aldrich or Fisher, and in most cases were used without further purification. DCM (P_2O_5), and DMF (BaO) were distilled prior to use. Melting points (mp) were performed using a Mel-temp instrument and are uncorrected. ^{1}H NMR spectra were obtained at 400 MHz on either a Varian INOVA or Bruker UltraShield spectrometer and the chemical shifts are reported in ppm. Infrared spectra were recorded on either a Midac FT-IR or Bruker Alpha-P spectrophotometer. The major bands are reported in wave numbers (cm^{-1}). High-resolution mass spectra (HRMS) and Low-resolution mass spectra (LRMS) were provided by the Mass Spectrometry Laboratory, University of South Carolina, Columbia. Reaction progress was assessed by thin-layer chromatography (TLC) using Merck silica gel (60 F_{254}) on aluminum plates unless otherwise stated. Visualization was achieved with UV light at 254 nm and/or 366 nm, I_2 vapor staining and ninhydrin spray.

3.6.1. NO_2-P(C_3Cl)P (9)

NO_2-P **7** (500 mg, 1.96 mmol) was reduced in the presence of hydrogen over 5% Pd/C (250 mg) in cold MeOH (20 mL) for 18 h. The reaction mixture was filtered over celite and the catalyst was washed thoroughly with MeOH. The solvent was removed by evaporation and the residue was co-evaporated with dry DCM (3 × 5 mL). The compound was coupled with a solution of acid chloride in dry DCM (10 mL). The acid chloride of 1-(3-chloropropyl)-4-nitro pyrrole-2-carboxylic acid **8** was prepared by refluxing the acid (546 mg, 2.36 mmol) in $SOCl_2$ (9 mL) for 15 min. The excess $SOCl_2$ was evaporated and the residual reagent removed by co-evaporation with dry DCM (3 × 5 mL). The acid chloride solution was added dropwise to the reduced compound dissolved in dry DCM (20 mL) and dry Et_3N (0.21 mL, 1.5 mmol) at 0 °C and allowed to stir for 18 h. A basic aqueous work-up was performed and the residue purified by flash column chromatography eluting in MeOH/$CHCl_3$ (0:100–100:0) to give **9** as a yellow solid (667 mg, 77%), mp 255–257 °C, R_f 0.18 [MeOH/$CHCl_3$ (20:80)]; IR: (KBr) υ 3480, 2926, 1673, 1666, 1659, 1651, 1644, 1634, 1623, 1615, 1587, 1574, 1568, 1538, 1504, 1441, 1435, 1423, 1403, 1385, 1372 1313; ^{1}H NMR ($CDCl_3$) δ 7.82 (s br, 1H), 7. 73 (s, 1H), 7.70 (s, 1H), 7.17 (ds, 1H), 6.49 (s, 1H), 4.60 (t, *J* = 4.0 Hz, 2H), 3.94 (s, 3H), 3.54 (t, *J* = 4.0 Hz, 2H), 3.48 (quart, *J* = 6.0 Hz, 2H), 2.56 (t, *J* = 6.0 Hz, 2H), 2.38 (s, 6H), 2.36 (quint, *J* = 4.0 Hz, 2H), 1.79 (quint, *J* = 6.0 Hz, 2H); LRMS (ES$^+$) *m*/*z* = 439 ([M+H]$^+$, 78%).

3.6.2. NO_2-IP(C_3Cl)P (11)

NO_2-P(C_3Cl)P **9** (525 mg, 1.20 mmol) was reduced over hydrogen and 5% Pd/C (262 mg) in cold MeOH (20 mL) for 18 h. The reaction mixture was filtered over celite and the catalyst was washed thoroughly with MeOH. The solvent was removed by evaporation and the residual MeOH removed by co-evaporation with dry DCM (3 × 5 mL). The compound was coupled with the acid chloride in dry DCM (10 mL). The acid chloride of 1-methyl-4-nitro-imidazole-2-carboxylic acid **10**, prepared by refluxing the acid (246 mg, 1.44 mmol) in oxalyl chloride (1.8 mL) and dry THF (1.8 mL) for 45 min. The excess oxalyl chloride was evaporated and the residual reagent removed by co-evaporation with dry DCM (3 × 5 mL). The acid chloride solution was added dropwise to the reduced compound dissolved in dry DCM (20 mL) and dry Et_3N (0.21 mL, 1.5 mmol) at 0 °C and allowed to stir for 18 h. A basic aqueous work-up was performed and the residue purified by flash column chromatography eluting in MeOH/$CHCl_3$ (0:100–100:0) to give **11** as a yellow solid (269 mg, 48 %), R_f 0.14 [MeOH/$CHCl_3$ (50:50)]; IR: (KBr) υ 2952, 2924, 2850, 2360, 2341, 1673, 1667, 1656, 1650, 1644, 1582, 1576, 1563, 1543, 1538, 1532, 1463, 1457, 1443, 1397, 1385, 1308, 1256, 1152, 1118, 1086, 1001, 828, 775, 752; ^{1}H NMR ($CDCl_3$) δ 9.05 (s br, 1H), 7.85 (s, 1H), 7.78 (t br, 1H), 7.64 (s br, 1H), 7.32 (d, *J* = 2.0 Hz, 1H), 7.20 (d, *J* = 2.0 Hz, 1H), 6.85 (d, *J* = 2.0 Hz, 1H), 6.52 (d, *J* = 2.0 Hz, 1H), 4.53 (t, *J* = 4.5 Hz, 2H), 4.23 (s, 3H), 3.94 (s, 3H), 3.52 (t, *J* = 4.5 Hz, 2H), 3.49 (quart, *J* = 6.0 Hz, 2H), 2.58 (t, *J* = 6.0 Hz, 2H), 2.41 (6H, s), 2.32 (quint, *J* = 4.5 Hz, 2H), 1.81 (quint, *J* = 6.0 Hz, 2H); LRMS (ES$^+$) *m*/*z* = 562 ([M+H]$^+$, 100%).

3.6.3. f-IP(C_3Cl)P (12)

Formic acetic anhydride was formed by the addition of formic acid (2.5 mL) to a solution of acetic anhydride (5 mL) at 0 °C. The solution was heated at 50 °C for 30 min. NO_2-I(C_3Cl)PP **11**(269 mg, 0.479 mmol) was reduced in the presence of hydrogen over 10% Pd/C (135 mg) in cold MeOH (20 mL) for 18 h. The reaction mixture was filtered over celite and the catalyst washed thoroughly with MeOH. The solvent was removed by evaporation and the residual MeOH removed by co-evaporation with dry DCM (3 × 5 mL). The pre-formed acetic anhydride was added dropwise (over 10 min) to the amine in dry DCM (15 mL) at 0°Cand stirred for 18 h. The solution was quenched with methanol (75 mL), the solvent removed by evaporation and the residue obtained purified by flash column chromatography eluting in MeOH/$CHCl_3$ (0:100–100:0) to give **12** as a pale yellow solid (118 mg, 44%), R_f 0.22 [MeOH/$CHCl_3$ (50:50)]; IR: (KBr) υ 3376, 2950, 2865, 2245, 1643, 1556, 1538, 1480, 1410, 1260, 1205, 1150, 1117, 909, 803, 733; ^{1}H NMR (CD_3OD) δ 8.32 (s, 1H), 7.51 (s, 1H), 7.46 (d, *J* = 1.5 Hz, 1H), 7.22 (d, *J* = 1.5 Hz, 1H), 7.06 (d, *J* = 1.5 Hz, 1H), 6.88 (d, *J* = 1.5 Hz, 1H), 4.58 (t, *J* = 4.6 Hz, 2H), 4.13 (s, 3H), 3.95 (s, 3H), 3.60 (t, *J* = 4.6 Hz, 2H), 3.41 (t, *J* = 6.0 Hz, 2H), 2.52 (t, *J* =6.0 Hz, 2H), 2.37 (6H, s), 2.32 (quint, *J* = 4.6 Hz, 2H), 1.87 (quint, *J* = 6.0 Hz, 2H); LRMS (ES$^+$) *m*/*z* = 560 ([M+H]$^+$, 100%); HRMS [M+H]$^+$ calcd for *m*/*z* $C_{25}H_{35}ClN_9O_4$ 560.2501; found *m*/*z* 560.2509.

3.6.4. f-IP(C_3N_3)P (13)

A solution of f-IP(C_3Cl)P **12** (100 mg, 0.179 mmol) and NaN_3 (58 mg, 0.89 mmol) in dry DMF was heated at 80 °C overnight. The DMF was removed by Kugelrohr distillation (0.1 mm Hg, 60 °C) and the residue was purified by flash column chromatography eluting in MeOH/$CHCl_3$ (0:100–100:0) to afford **13** as a pale yellow solid (71 mg, 70%), mp 108 °C, R_f 0.09 [MeOH/$CHCl_3$ (50:50)]; IR: (KBr) υ 3365, 2926, 2862, 2360, 2339, 2097, 1657, 1630, 1573, 1567, 1467, 1456, 1435, 1403, 1387, 1250, 1206, 1118, 1081, 1039, 987, 800, 732;^{1}H NMR (CD_3OD) δ 8.27 (s, 1H); 7.45 (s, 1H), 7.35 (d, *J* = 1.5, 1H), 7.20 (d, *J* = 1.5, 1H), 6.97

(d, J = 1.5, 1H), 6.87 (d, J = 1.5, 1H), 4.43 (t, J = 4.8 Hz, 2H), 4.08 (s, 3H), 3.93 (s, 3H), 3.40 (t, J = 6.0 Hz, 2H), 3.33 (t, J = 4.5 Hz, 2H), 2.82 (t, J = 6.0 Hz, 2H), 2.60 (6H, s); 2.05 (quint, J = 4.5 Hz, 2H), 1.90 (quint, J = 6.0 Hz, 2H); LRMS (ES^+) m/z = 567 ($[M+H]^+$, 100%); HRMS $[M+H]^+$ calcd for m/z $C_{25}H_{35}N_{12}O_4$ 567.2904; found m/z 567.2891.

3.6.5. f-IP(C_3NH_2)P or f-I*P*P (4)

A solution of f-IC_3N_3PP **13** (60 mg, 0.11 mmol) was reduced over hydrogen and 10% Pd/C (30 mg) in cold MeOH (20 mL) for 18 h. The mixture was filtered through celite and the solvent evaporated and allowed to dry under vacuum to yield f-I(C_3NH_2)PP-tail **4** as an off white solid (34 mg, 60%), mp 119 °C, R_f 0.15 [MeOH/$CHCl_3$/NH_4OH (5:15:1)]; IR: (KBr) υ 3320, 2860, 2360, 2342, 1673, 1652, 157, 1538, 1506, 1469, 1463, 1456, 1435, 1261, 1200, 1123, 1033, 899, 802, 738; ^1H NMR (CD_3OD) δ 8.15 (s, 1H), 7.32 (s, 1H), 7.27 (d, J = 1.5 Hz, 1H), 7.09 (d, J = 1.5 Hz, 1H), 6.85 (d, J = 1.5 Hz, 1H), 6.70 (d, J = 1.5 Hz, 1H), 4.35 (t, J = 4.0 Hz, 2H), 3.95 (s, 3H), 3.75 (s, 3H), 3.23 (t, J = 6.0 Hz, 2H), 2.57 (t, J = 6.0 Hz, 2H), 2.32 (t, J = 4.6 Hz, 2H), 2.17 (s, 6H), 1.90 (quint, J = 4.0 Hz, 2H), 1.68 (quint, J = 6.0 Hz, 2H); LRMS (ES^+) m/z = 542 ($[M+H]^+$, 12%).

3.6.6. NO_2-P(C_3Cl)IP (15)

The synthesis of compound **15** was similar to that used for the preparation of diamide **9** except NO_2-IP **14** (200 mg, 0.531 mmol) was used. After purification by silica gel chromatography [eluting in MeOH/$CHCl_3$ (0:100–100:0)] **15** was isolated as an orange solid (197 mg, 66%), mp 180–185 °C, R_f 0.32 [MeOH/$CHCl_3$ (30:70)]; IR: (KBr) υ 2360, 2342, 1669, 1663, 1540, 1558, 1507, 1489, 1323, 725;^1H NMR ($CDCl_3$) δ 9.00 (s br, 1H), 8.61 (s br, 1H), 7.95 (s br, 1H), 7.74 (d, J = 2.0, 1H), 7.64 (s, 1H), 7.44 (s, 1H), 7.25 (d, J = 2.0, 1H), 6.71, (s, 1H), 4.63 (t, J = 4.5 Hz, 2H), 4.07 (s, 3H), 3.90 (s, 3H), 3.55 (m, 4H), 2.71 (t, J = 5.6 Hz, 2H), 2.48 (s, 6H), 2.37 (quint, J = 4.5 Hz, 2H), 1.91 (quint, J = 5.6 Hz, 2H); LRMS (ES^+) m/z = 562 ($[M+H]^+$, 100%).

3.6.7. f-P(C_3Cl)IP (16)

The synthesis of **16** followed a similar procedure to the preparation of **12** except NO_2-P(C_3Cl)IP **15** (236 mg, 0.43 mmol) was used. Upon quenching the reaction with methanol (75 mL) the removal of the solvent yielded **16** as a tan solid (156 mg, 66%), mp 168-170 °C; IR: KBR υ 3256, 3030, 2922, 2850, 1676, 1654, 1648, 1637, 1632, 1577, 1571, 1560, 1540, 1535, 1529, 1522, 1518, 1503, 1491, 1473, 1466, 1458, 1449, 1438, 1431, 1420, 1400, 1400, 1285, 1257, 1206, 1161, 1114, 1098, 1017, 900, 859, 767; ^1H NMR ($CDCl_3$) δ 9.48 (s, 1H), 9.46 (s, 1H), 9.41 (s, 1H), 8.24 (s, 1H), 7.91 (t br, 1H), 7.43 (s, 1H), 7.39 (s, 1H), 7.36 (s, 1H), 6.97 (s, 1H), 6.95 (s, 1H), 4.49 (t, J = 4.5 Hz, 2H), 4.04 (s, 3H), 3.88 (s, 3H), 3.51 (m, 4H), 2.92 (t, J = 6.2 Hz, 2H), 2.70 (6H, s), 2.27 (quint, J = 4.5 Hz, 2H), 2.03 (quint, J = 6.2 Hz, 2H); LRMS (ES^+) m/z = 560 ($[M+H]^+$, 78%).

3.6.8. f-P(C_3N_3)IP (17)

As with the synthesis of **13**, S_N2 reaction of f-P(C_3Cl)IP-tail **16** (250 mg, 0.458 mmol) with NaN_3 (149 mg, 2.29 mmol), and after chromatographic separation [silica gel, gradient, 0:100–100:0 %v/v, $CHCl_3$/MeOH] yielded **17** as a light orange solid (171 mg, 67%), mp 150–152 °C, R_f 0.50 [MeOH/$CHCl_3$ (30:70)]; IR: (KBr) υ 3320, 2931, 2865, 2823, 2782, 2009, 1655, 1539, 1463, 1404, 1260, 1211, 1115, 1201, 901, 799, 754; ^1H NMR ($CDCl_3$) δ 9.06 (s, 1H), 8.45 (s, 1H), 8,80 (s, 1H), 8.32 (s br, 1H), 8.27 (s, 1H), 7.83 (s, 1H), 7.44 (s, 1H), 7.30 (s, 1H), 7.18 (s, 1H), 6.68 (s, 1H), 6.64 (s, 1H), 4.37 (t, J = 4.6 Hz, 2H), 3.97 (s, 3H), 3.84 (3H, s), 3.45 (quart br, 2H), 3.27 (t, J = 4.6 Hz, 2H), 2.42 (t br, 2H), 2.22 (s, 6H); 2.02 (quint, J = 4.6 Hz, 2H); 1.72 (quint br, 2H); LRMS (ES^+) m/z = 567 ($[M+H]^+$, 100%); HRMS $[M+H]^+$ calcd for m/z $C_{25}H_{35}N_{12}O_4$ 567.2904; found m/z 567.2895.

3.6.9. f-P(C_3NH_2)IP or f-*P*IP (5)

Similar to the synthesis of **4**, catalytic hydrogenation solution of f-P(C_3N_3)IP **17** (41 mg, 0.07 mmol) over 10% Pd/C (21 mg) yielded f-P(C_3NH_2)IP-tail **5** as a yellow solid (22 mg, 60%), mp 143-147 °C, R_f 0.22 [MeOH/$CHCl_3$/NH_4OH (5:15:1)]; IR: (KBR) υ 3320, 2962, 2919, 2849, 1650, 1527, 1462, 1401, 1260, 1206, 1091, 1017, 900, 862, 797; ^1H NMR (CD_3OD) δ 8.l5 (s, 1H), 7.45 (s, 1H), 7.32 (d, J = 1.2 Hz, 1H), 7.24 (d, J = 1.2 Hz, 1H), 6.92 (d, J = 1.2 Hz, 1H), 6.81 (d, J = 1.2 Hz, 1H), 4.37 (t, J = 4.6 Hz, 2H), 4.06 (s, 3H), 3.87 (s, 3H), 3.33 (t, J = 6.0 Hz, 2H), 2.47 (t, J = 4.6 Hz, 2H), 2.36 (t, J = 6.0 Hz, 2H), 2.30 (s, 6H), 1.98 (quint, J = 4.6 Hz, 2H), 1.80 (quint, J = 6.0 Hz, 2H); LRMS (ES^+) m/z = 541 ($[M+H]^+$, 100%); HRMS $[M+H]^+$ calcd for m/z $C_{25}H_{37}N_{10}O_4$ 541.2999; found m/z 541.2990.

3.6.10. f-PP(C_3Cl)P (19)

NO_2-P(C_3Cl)P **9** (250 mg, 0.570 mmol) was reduced over hydrogen and 5% Pd/C (125 mg) in cold MeOH (20 mL) for 18 h. The reaction mixture was filtered over celite and the catalyst was washed thoroughly with MeOH. The solvent was removed by evaporation and the residue was co-evaporated with dry DCM (3 × 2 mL). The amine was allowed to dry under high vacuum for 30 min before use. Dry DiPEA (0.35 mL, 1.87 mmol) was added to a solution of f-PCOOH **18**[18] (88 mg, 0.52 mmol) and PyBOP (330 mg, 0.063 mmol) in dry DMF (3 mL) under argon atmosphere and allowed to stir at room temperature for 45 min. The dried amine was added to the activated acid in dry DMF (2 mL) over a period of 10 min and the solution stirred overnight. The DMF was removed by Kugelrohr distillation and the residue was purified by silica gel column chromatography eluting in MeOH/$CHCl_3$ (0:100–100:0) to afford **19** as a light yellow solid (120 mg, 38%), mp 258-260 °C, R_f 0.50 [MeOH/$CHCl_3$/NH_4OH (5:15:1)]; IR: (KBr) υ 3350, 2969, 2931, 2911, 2848, 1702, 1657, 1578, 1554, 1533, 1457, 1395, 1347, 1254, 1095, 1084, 1022 797, 668; ^1H NMR (CD_3OD) δ 8.12 (s, 1H), 7.28 (s, 1H), 7.18 (s, 1H); 7.15 (s, 1H), 6.96 (s, 1H), 6.87 (s, 1H), 6.79 (s, 1H), 4.47 (2H, t, J 6 Hz), 3.89 (3H, s), 3.86 (3H, s), 3.50 (t, J = 6.2 Hz, 2H), 3.30 (t, J = 4.6 Hz, 2H), 2.39 (t, J = 6.2 Hz, 2H), 2.26 (s, 6H), 2.23 (quint, J = 4.6 Hz, 2H), 1.76 (quint, J = 6.2 Hz, 2H); LRMS (ES^+) m/z = 559 ($[M+H]^+$, 100%); HRMS $[M+H]^+$ calcd for m/z $C_{26}H_{36}ClN_8O_4$ 559.2548; found m/z 559.2544.

3.6.11. f-PP(C_3N_3)P (20)

As with the synthesis of 13, of f-PP(C_3Cl)P reaction of **19** (150 mg, 0.268 mmol) with NaN_3 (87 mg, 1.3 mmol) in dry DMF yielded after column chromatography [MeOH/$CHCl_3$ (0:100–100:0)] azide **20** as a light yellow solid (100 mg, 66%), mp 185–190 °C, R_f 0.50 [MeOH/$CHCl_3$/NH_4OH (5:15:1)]; IR: (KBr) υ 3285, 2941, 2866, 2534, 2097, 1637, 1579, 1455, 1441, 1400, 1260, 1133, 1098, 890, 811, 774; ^1H NMR (CD_3OD) δ 8.12 (s, 1H), 7.28 (s, 1H), 7.18 (s, 1H), 7.15 (s, 1H), 6.96 (s, 1H), 6.87 (s, 1H), 6.79 (s, 1H), 4.47 (t, J = 4.6 Hz, 2H), 3.89 (s, 3H), 3.86 (s, 3H), 3.50 (t, J = 6.0 Hz, 2H), 3.33 (t, J = 4.6 Hz, 2H), 2.41 (t, J = 6.0 Hz, 2H), 2.26 (s, 6H), 2.20 (quint, J = 4.6 Hz, 2H), 1.76 (quint, J = 6.0 Hz, 2H); LRMS (ES^+) m/z = 566 ($[M+H]^+$, 100%); HRMS $[M+H]^+$ calcd for m/z $C_{26}H_{36}N_{11}O_4$ 566.2952; found m/z 566.2957.

3.6.12. f-PP(C_3NH_2)P or f-P*P*P (6)

A solution of f-PP(C_3N_3)P **20** (45 mg, 0.080 mmol) and triphenylphosphine (126 mg, 0.477 mmol) in THF:H_2O (2 mL, 10:1) was stirred for 3 days. The solution was evaporated, co-evaporated with dry methanol until dryness. The residue was purified by silica gel column chromatography eluting in MeOH/$CHCl_3$ (0:100–100:0) to give **6** as a yellow solid (23 mg, 53.6%), mp128-130 °C, R_f 0.20

[MeOH/$CHCl_3$/NH_4OH (5:15:1)]; IR: (KBr) υ 3401, 3323, 3318, 2959, 2911, 2871, 2825, 2787, 1664, 1651, 1581, 1558, 1536, 1454, 1436, 1402, 1343, 1257, 1099, 1089, 1030, 819, 778; ^{1}H NMR (CD_3OD) δ 8.13 (s, 1H), 7.29 (d, J = 1.5 Hz, 1H), 7.18 (d, J = 1.5 Hz, 1H), 7.17 (d, J = 1.5 Hz, 1H), 6.92 (d, J = 1.5 Hz, 1H), 6.87 (d, J = 1.5 Hz, 1H), 6.79 (d, J = 1.5 Hz, 1H), 4.44 (t, J = 4.6 Hz, 2H), 3.90 (s, 3H), 3.87 (s, 3H), 3.33 (t, J = 6.0 Hz, 2H), 2.68 (t, J = 4.6 Hz, 2H), 2.42 (t, J = 6.0 Hz, 2H), 2.27 (s, 6H), 1.99 (quint, J = 4.6 Hz, 2H), 1.77 (quint, J = 6.0 Hz, 2H); LRMS (ES^+) m/z = 540 ($[M+H]^+$, 30%), 271 (100%); HRMS $[M+H]^+$ calcd for m/z $C_{26}H_{38}N_9O_4$ 540.3047; found m/z 540.3055.

Acknowledgments

The authors thank the NSF (CHE 0809162, CHE0922623), Cancer Research UK (C2259/A9994 to JAH) and the Georgia Research Alliance for their generous support. The authors also thank Dr. Yang Liu for his helpful assistance in performing some of the experiments.

Supplementary data

Supplementary data (^{1}H NMR spectra of diamino PAs **4**, **5**, and **6**, as well as the full gel for the footprinting of f-*P*IP (**5**) given in Figure 2B) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2013.04.001.

References and notes

1. (a) Cai, X.; Gray, P. J., Jr.; Von Hoff, D. D. *Cancer Treat. Rev.* **2009**, *35*, 437; (b) Neidle, S. *Nat. Chem.* **2012**, *4*, 594; (c) Tse, W. C.; Boger, D. L. *Chem. Biol.* **2004**, *11*, 1607; (d) Bischoff, G.; Hoffmann, S. *Curr. Med. Chem.* **2002**, *9*, 321; (e) Ebbinghaus, S. W. *Chem. Biol.* **2003**, *10*, 895.
2. Buchmueller, K. L.; Uthe, P. B.; Howard, C. M.; Pacheco, K. A.; Cox, K.; Henry, J. A.; Bailey, S. A.; Horick, S. M.; Nguyen, B.; Wilson, W. D.; Lee, M. *Nucleic Acids Res.* **2005**, *33*, 912.
3. (a) White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 2153; (b) Chenoweth, D. M.; Meier, J. L.; Dervan, P. B. *Angew. Chem., Int. Ed.* **2013**, *52*, 415.
4. (a) White, S.; Szewezyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468; (b) Buchmueller, K. L.; Staples, A. M.; Horick, S. M.; Uthe, P. B.; Le, N. M.; Cox, K. K.; Nguyen, B.; Pacheco, K. A.; Wilson, W. D.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742; (c) Collar, C. J.; Lee, M.; Wilson, W. D. *J. Chem. Inf. Model.* **2010**, *50*, 1611; (d) Kopka, M. L.; Goodsell, D. S.; Han, G. W.; Chiu, T. K.; Lown, J. W.; Dickerson, R. E. *Structure* **1997**, *5*, 1033.
5. (a) Yang, X. L.; Hubbard, R. B.; Lee, M.; Tao, Z. F.; Sugiyama, H.; Wang, A. H. *Nucleic Acids Res.* **1997**, *27*, 4183; (b) Yang, X. L.; Kaenzig, C.; Lee, M.; Wang, A. H. *Eur. J. Biochem.* **1999**, *263*, 646; (c) Zhang, Q.; Dwyer, T.; Tsui, V.; Case, D. A.; Cho, J.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **2004**, *126*, 7958.
6. (a) Kielkopf, C. L.; Bremer, R. E.; White, S. S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *J. Mol. Biol.* **1998**, *295*, 557; (b) Chenoweth, D. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13175; (c) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
7. (a) Westrate, L.; Mackay, H.; Brown, T.; Nguyen, B.; Kluza, J.; Wilson, W. D.; Lee, M.; Hartley, J. A. *Biochemistry* **2009**, *48*, 5679; (b) Buchmueller, K. L.; Bailey, S. L.; Matthews, D. A.; Taherbhai, Z. T.; Register, J. K.; Davis, Z. S.; Bruce, C. D.; O'Hare, C.; Hartley, J. A.; Lee, M. *Biochemistry* **2006**, *45*, 13551; (c) Brown, T.; Mackay, H.; Turlington, M.; Sutterfield, A.; Smith, T.; Sielaff, A.; Westrate, L.; Bruce, C. D.; Kluza, J.; O'Hare, C.; Nguyen, B.; Wilson, W. D.; Hartley, J. A.; Lee, M. *Bioorg. Med. Chem.* **2008**, *16*, 5266.
8. Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, D. W. *J. Am. Chem. Soc.* **2002**, *124*, 2153.
9. Brown, T.; Taherbhai, Z.; Sexton, J.; Sutterfield, A.; Turlington, M.; Jones, J.; Stallings, L.; Stewart, L.; Buchmueller, K. L.; Mackay, H.; O'Hare, C.; Kluza, J.; Nguyen, B.; Wilson, D.; Lee, M.; Hartley, J. A. *Bioorg. Med. Chem.* **2007**, *15*, 474.
10. (a) Verma, R.; Patapoutian, A.; Gordon, C. B.; Campbell, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7155; (b) Wu, X.; Lee, H. *Oncogene* **2002**, *21*, 7786; (c) Yamada, M.; Sato, N.; Taniyama, C.; Ohtani, K.; Drong, R. F.; Weiland, K. L.; Slightom, J. L.; Sclafani, R. A.; Hollingsworth, R. E. *Gene* **1998**, *28*, 133.
11. (a) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363; (b) Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802; (c) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12063; (d) Lown, J. W.; Krowicki, K.; Bhat, U. G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 7408.
12. Satz, A. L.; Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 86.
13. Satam, V.; Babu, B.; Chavda, S.; Savagian, M.; Sjoholm, R.; Tzou, S.; Liu, Y.; Kiakos, K.; Lin, S.; Wilson, W. D.; Hartley, J. A.; Lee, M. *Bioorg. Med. Chem.* **2012**, *20*, 693.
14. Satam, V.; Babu, B.; Porte, A.; Savagian, M.; Lee, M.; Smeltzer, T.; Liu, Y.; Ramos, J.; Wilson, D. W.; Lin, S.; Kiakos, K.; Hartley, J. A.; Lee, M. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5898.
15. Babu, B.; Liu, Y.; Plaunt, A.; Riddering, C.; Ogilvie, R.; Westrate, L.; Davis, R.; Ferguson, A.; Mackay, H.; Rice, T.; Chavda, S.; Wilson, D. W.; Lin, S.; Kiakos, K.; Hartley, J. A.; Lee, M. *Biochem. Biophys. Res. Commun.* **2011**, *404*, 848.
16. Ramos, J. P.; Babu, B.; Chavda, S.; Liu, Y.; Plaunt, A.; Ferguson, A.; Savagian, M.; Lee, M.; Tzou, S.; Lin, S.; Kiakos, K.; Wang, S.; Lee, M.; Hartley, J. A.; Wilson, W. D. *Biopolymers*, in press.
17. Chavda, S.; Liu, Y.; Babu, B.; Davis, R.; Sielaff, A.; Ruprich, J.; Westrate, L.; Tronrud, C.; Ferguson, A.; Franks, A.; Tzou, S.; Adkins, C.; Rice, T.; Mackay, H.; Kluza, J.; Tahir, S. A.; Lin, S.; Kiakos, K.; Bruce, C. D.; Wilson, W. D.; Hartley, J. A.; Lee, M. *Biochemistry* **2011**, *50*, 3127.
18. Mulder, K.; Sexton, J.; Taherbhai, Z.; Jones, J.; Uthe, P.; Brown, T.; Lee, M. *Synth. Commun.* **2008**, *38*, 33.
19. Le, N. M.; Sielaff, A.; Cooper, A.; Mackay, H.; Brown, T.; Kotecha, M.; O'Hare, C.; Hochhauser, D.; Lee, M.; Hartley, J. A. *Bioorg. Med. Chem. Lett.* **2006**, *26*, 6161. and references therein.
20. Franks, A.; Tronrud, C.; Kiakos, K.; Kluza, J.; Munde, M.; Brown, T.; Mackay, H.; Wilson, W. D.; Hochhauser, D.; Hartley, J. A.; Lee, M. *Bioorg. Med. Chem.* **2010**, *18*, 5553.
21. (a) Dervan, P. B. *Science* **1986**, *232*, 464; (b) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457; (c) Zimmer, C.; Wahnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31.
22. (a) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393; (b) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723.
23. Hamelberg, D.; McFail-Isom, L.; Williams, L. D.; Wilson, W. D. *J. Am. Chem. Soc.* **2000**, *122*, 10513. and references given therein.
24. Mazur, S.; Tanious, F. A.; Ding, D.; Kumar, A.; Boykin, D. W.; Simpson, I. J.; Neidle, S.; Wilson, W. D. *J. Mol. Biol.* **2000**, *300*, 321.
25. Record, M. T., Jr.; Anderson, C. F.; Lohman, T. M. Q. *Rev. Biophys.* **1978**, *11*, 103.
26. (a) Nguyen, B.; Tanious, F.; Wilson, W. D. *Methods* **2007**, *42*, 150; (b) Nanjunda, R.; Munde, M.; Liu, Y.; Wilson, W. D. In *Methods for Studying Nucleic Acid/Drug Interactions*; Wanunu, Y., Ed.; CRC Press: Boca Raton, 2011; pp 91–122. Chapter 4.