Published in final edited form as:

Biopolymers. 2013 August ; 99(8): . doi:10.1002/bip.22205.

Affinity and Kinetic Modulation of Polyamide–DNA Interactions by N-Modification of the Heterocycles

Joseph P. Ramos¹, Balaji Babu², Sameer Chavda², Yang Liu¹, Adam Plaunt², Amanda Ferguson², Mia Savagian², Megan Lee², Samuel Tzou², Shicai Lin³, Konstantinos Kiakos³, Shuo Wang¹, Moses Lee², John A. Hartley³, and W. David Wilson¹

¹Department of Chemistry, Georgia State University, Atlanta, GA 30303

²Department of Chemistry and the Division of Natural and Applied Sciences, Hope College, MI 49423

³Cancer Research, UK Drug–DNA Interactions Research Group, UCL Cancer Institute, Paul O' Gorman Building, 72 Huntley Street, London WC1E 6BT, UK

Abstract

Synthetic N-methyl imidazole and N-pyrrole containing polyamides (PAs) that can form “stacked” dimers can be programmed to target and bind to specific DNA sequences and control gene expression. To accomplish this goal, the development of PAs with lower molecular mass which allows for the molecules to rapidly penetrate cells and localize in the nucleus, along with increased water solubility, while maintaining DNA binding sequence specificity and high binding affinity is key. To meet these challenges, six novel f-ImPy*Im PA derivatives that contain different orthogonally positioned moieties were designed to target 5'-ACGCGT-3'. The synthesis and biophysical characterization of six f-ImPy*Im were determined by CD, ΔT_M , DNase I footprinting, SPR, and ITC studies, and were compared with those of their parent compound, f-ImPyIm. The results gave evidence for the minor groove binding and selectivity of PAs **1** and **6** for the cognate sequence 5'-ACGCGT-3', and with strong affinity, $K_{eq} = 2.8 \times 10^8 \text{ M}^{-1}$ and $K_{eq} = 6.2 \times 10^7 \text{ M}^{-1}$, respectively. The six novel PAs presented in this study demonstrated increased water solubility, while maintaining low molecular mass, sequence specificity, and binding affinity, addressing key issues in therapeutic development.

Keywords

polyamides; diamino; dicationic; gene control; ACGCGT; *MCB*; MluI; cell-cycle box; pyrrole; imidazole; DNA; minor groove binder; sequence selectivity; binding affinity; kinetics

INTRODUCTION

Synthetic pyrrole (Py) and imidazole (Im) containing analogs (Figure 1) of the naturally occurring polyamide (PA) compounds distamycin and netropsin have been shown to selectively bind as “stacked” dimers to the minor groove of DNA with high affinity and selectivity.^{1–8} The addition of a formamido (f) group to the N-terminus of PAs increases the binding affinity and causes the stacking to be in a “staggered” position as opposed to “overlapped” stacking.^{1,9,10} Because of their high affinity and sequence selectivity, PAs

have demonstrated potential as therapeutics. Based on their ability to target, alter, and control gene expression, especially those genes associated with such diseases as cancer, PAs are of interest for the design and development of new types of therapeutics.^{1-4,11-16}

The PA, f-ImPyIm (Figure 1), binds to the cognate DNA sequence 5'-ACGCGT-3' as a stacked, staggered homodimer with a high binding affinity, $K_{eq} > 10^8 M^{-1}$, for a small molecule.^{9,10,17,18} Closely related trimer PAs bind to their cognate sites from 10 to 100 fold more weakly.^{1,9} The sequence 5'-ACGCGT-3' is of significance due to its occurrence in the core sequence of the *MluI* cell-cycle box (MCB), a transcriptional element found in the promoter region of the human *Dbf4* (huDbf4 or ASK, activator S-kinase) gene. High levels of Cdc7 (cyclin dependent) kinase have been implicated in the development of various cancers and *Dbf4* is a regulatory subunit of this kinase.¹⁹⁻²¹

Our approach to developing modified derivatives of PAs, such as f-ImPyIm, as therapeutic candidates is to develop analogues that will facilitate an increase in binding affinity, sequence selectivity, solubility, and biological activity, including cellular uptake and nuclear localization. An attractive approach for the design of potentially improved analogues of f-ImPyIm is the addition of different functional groups at the N1-position of the center pyrrole coupled with a detailed study of their DNA binding affinity and sequence selectivity.^{22,23} A dicationic PA derivative of f-ImPyIm with an alkyl-amine substituent in place of the usual N1 methyl group, **1**, showed improved solution and DNA binding properties.^{22,23} It is not clear what part of the improved DNA binding is due to the cations and how much is due to possible interactions of the alkyl substituents with each other and with DNA. To evaluate the separate features of the substituent, f-ImPy*Im analogs with modified substituents were synthesized (shown in Figure 1), where Py* is the modified heterocycle. Both neutral and modified cationic substituents were evaluated for DNA binding by using surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), DNase I footprinting, thermal denaturation, and circular dichroism (CD). All synthesized PAs investigated in this study retain the same f-ImPy*Im core and, therefore, have the same 5'-ACGCGT-3' cognate sequence.

MATERIALS AND METHODS

Methods

Surface Plasmon Resonance—SPR measurements were performed with a four-channel Biacore T200 optical biosensor system (Biacore, GE Healthcare). The 5'-Biotin-labeled DNA hairpin duplex samples were immobilized onto a streptavidin-coated sensor chip (Biacore SA) as previously described.^{10,23,24} All SPR experiments were performed at 25°C and used 0.01M cacodylic acid (CCA), 0.1M NaCl, 0.001M ethylenediaminetetraacetic (EDTA), 0.05% (by vol) P20 surfactant, and pH 6.25 as the running buffer. The amount of DNA immobilized was ~400 response units (RU). This was achieved by continuously injecting ~20 μ L of an ~50 nM DNA solution at a rate of 2 μ L min⁻¹ onto the sensor chip surface until a relative response of 400 U was reached. Binding data were obtained by injecting known concentrations over each flow cell of the chip and were analyzed with a two site binding model as previously described^{24,25}: $r = (K_1 \times C_{free} + 2 \times K_1 \times K_2 \times C_{free}^2) / (1 + K_1 \times C_{free} + K_1 \times K_2 \times C_{free}^2)$ 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.²⁵

Isothermal Titration Calorimetry (ITC)—ITC analysis was performed using a VP-ITC microcalorimeter (MicroCal, MA). The DNA and compound samples were prepared in 0.01M CCA, 0.1M NaCl, 0.001M EDTA and pH 6.25 buffer solution. The instrument was equilibrated at a temperature of 25°C (298 K). The ITC experiments were conducted using the “excess DNA” method (“model-free” method), where an excess of DNA is used and does not allow for the reaction to reach an equilibrium end point, as previously reported.²⁵ Briefly, after an initial delay of 300 s, compounds (80 μ M) were titrated via 30 injections (10 μ L for 20 s, repeated every 300 s), into 30 μ M DNA. All ITC experiments were repeated in triplicate and reproducibility noted. The data were analyzed using the method previously reported.²⁶ Origin 7.0 was used and the area under each curve integrated as a function of time. A linear fit of the integrated heats was then employed using KaleidaGraph 4.0 and this was subtracted from the reaction integrations to normalize for nonspecific heat components. ΔG was calculated from $\Delta G = -RT \ln K_{eq}$, where, R is 1.987 cal mol⁻¹ K⁻¹ and T is measured in K .

DNase I Footprinting—DNase I footprinting experiments were conducted using a 130-bp 5'-[³²P]-radiolabeled DNA fragment to determine sequence specificity for PAs **1-6**. The engineered DNA fragment contained each of the following sequences, the cognate sequence 5'-ACGCGT-3', as well as two other DNA sequences with the same CG composition in differing arrangements, 5'-AGCGCT-3' and 5'-ACCGGT-3'. The protocol was reported previously.²⁷

Thermal Denaturation—Thermal denaturation 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.²⁷ ΔT_M values were determined from the difference in melting temperatures of the DNA-PA complex and duplex DNA alone. For each experiment 1.0 μ M of DNA and 3.0 μ M of PA were used.

Circular Dichroism (CD)—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 pathlength cuvette and a band pass of 2.4 nm.²⁷ 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 μ M was used as were the ratios of PAs (1, 2, 3, 4, 5, 6, 8, and 10 molar equivalents) titrated into the solution.

Synthesis of Modified ImPy*Im Derivatives

f-IP(C₃NH₂)I (1)—The synthesis of diamino f-IPI (**1**) was reported earlier.²⁸ A more efficient synthetic approach is reported herein. A solution of f-IP(C₃Cl)I (**2**) tail (100 mg, 0.183 mmol) in dry MeOH (5 mL) in a sealed tube was purged with dry NH₃ (g) for 6 h. The tube was sealed tightly and heated at 60°C for 24 h. The solvent was removed by evaporation and the residue purified by flash column chromatography eluting in MeOH:CHCl₃ (0:100–100:0) to give diamino f-IPI (**1**) as a yellow solid (30 mg, 31%). The ¹H NMR spectrum is identical to the product obtained from hydrogenation of the azide precursor.²⁸

f-IP(C₃Cl)I (2)—Nitro-IP(C₃Cl)I (**11**) (200 mg, 0.366 mmol) was reduced over hydrogen and 5% Pd/C (100 mg) in cold MeOH (100 mL) for 18 h at room temperature and atmospheric pressure. The reaction mixture was filtered over celite and the catalyst washed thoroughly with MeOH. The solvent was removed by evaporation and the residue was coevaporated with dry DCM (3 × 2 mL). The resulting dark orange solid was dried under high vacuum and protected from light. Formic acid (1 mL, kept at 0°C prior to use) was

added to acetic anhydride (2 mL) at 0°C and the solution heated at 50°C for 30 min. The solution was then added slowly dropwise (~10 min) to an ice-cold solution of the above amine dissolved in dry DCM (20 mL). The solution was stirred overnight. The excess anhydride was quenched with methanol (30 mL), and the solvent was removed by evaporation. A basic aqueous work-up was performed and the residue purified by flash column chromatography eluting in MeOH:CHCl₃ (0:100–100:0) to give chloro f-IPI (**2**) as a yellow solid (113 mg, 56%), mp 122°C, *R*_f = 0.43 (50:50 MeOH:CHCl₃); IR (NaCl) ν 3390, 3184, 2925, 2359, 1651, 1537, 1384, 1070; ¹H NMR (400 MHz, CDCl₃) δ 8.91 (s, 1H), 8.53 (s, 1H), 8.47 (s, 1H), 8.39 (s, 1H), 7.67 (t br, 1H), 7.53 (s, 1H), 7.42 (s, 1H), 7.29 (s, 1H), 6.94 (s, 1H), 4.61 (t, *J* = 6.0 Hz, 2H), 4.00 (s, 3H), 3.95 (s, 3H), 3.54 (t br, *J* = 6.0 Hz, 4H), 2.66 (t, *J* = 6.0 Hz, 2H), 2.37 (s, 6H), 2.33 (t, *J* = 6.0 Hz, 2H); LRMS (ES⁺) *m/z* (rel intensity) 548 (M+H⁺, 100%); HRMS (M+H⁺, C₂₃H₃₁ClN₁₀O₄⁺) calcd: 547.2294; found: 547.2286.

f-IP(C₃N₃)I (3**)**—To a solution of the f-IP(C₃Cl)I (**2**) (185 mg, 0.342 mmol) in dry DMF (2 mL), NaN₃ (103 mg, 1.58 mmol) was added. The solution was heated at 80°C overnight. The DMF was removed under vacuum (0.1 mm Hg, 60°C). A basic aqueous work-up was performed and the residue was purified by flash column chromatography eluting in MeOH:CHCl₃ (0:100–100:0) to give azido f-IPI (**3**) as a yellow solid (180 mg, 0.33 mmol, 96%), mp 112°C, *R*_f = 0.33 (30:70 MeOH:CHCl₃); IR (NaCl) ν 3329, 2921, 2359, 2099, 1659, 1531, 1470, 1402, 1082; NMR (400 MHz, CDCl₃) δ 8.85 (s, 1H), 8.37 (s, 1H), 8.00 (s br, 1H), 7.63 (s br, 1H), 7.53 (s br, 1H), 7.45 (s, 1H), 7.39 (s, 1H), 7.31 (s, 1H), 7.29 (s, 1H), 6.85 (s, 1H), 4.45 (t, *J* = 4.2 Hz, 2H), 4.08 (s, 3H), 4.03 (s, 3H), 3.49 (quart, *J* = 5.0 Hz, 2H), 3.32 (quart, *J* = 5.0 Hz, 2H), 2.52 (t, *J* = 5.0 Hz, 2H), 2.32 (s, 6H), 2.10 (quint, *J* = 4.2 Hz, 2H); LRMS (ES⁺) *m/z* (rel intensity) 554 (M+H⁺, 100%); HRMS (M+H⁺, C₂₃H₃₂N₁₃O₄⁺) calcd: 554.2700; Found: 554.2692.

f-IP(C₃alkyne)I (4**)**—NO₂P(C₃-alkyne)I (**14**) (100 mg, 0.241 mmol) was reduced by SnCl₂ • 2H₂O (354 mg, 1.57 mmol) and 12M HCl (1.1 mL) while refluxing in ethanol (10 mL) for 2 h. The reaction mixture was poured into water and the pH was brought to 10 (2M NaOH). The solution was extracted with chloroform (3 × 20 mL), the organic layers dried over anhydrous Na₂SO₄, and the solvent removed by evaporation. The corresponding amine was allowed to dry under high vacuum for 30 min. Separately, DiPEA (0.21 mL, 0.97 mmol) was added to a solution of 1-methyl-4-*N*-for-mamidoimidazole-2-carboxylic acid (45 mg, 0.26 mmol) and PyBOP (138 mg, 0.265 mmol) in dry DMF (2 mL) under argon atmosphere and the solution was stirred for 45 min. A solution of the dried amine dissolved in dry DMF (2 mL) was added to the activated acid solution over a period of 10 min. The reaction mixture was stirred overnight at room temperature and under an argon atmosphere. 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:CHCl₃ (0:100–100:0) to give alkynyl f-IPI (**4**) as a light pink solid (38 mg, 24%), mp 197–200°C, *R*_f = 0.31 (5:15:1 MeOH:CHCl₃:NH₄OH); IR (KBr) ν 3373, 3285, 2950, 2113, 1669, 1663, 1539, 1472, 1405, 1243, 1123, 1018, 904, 752; ¹H NMR (400 MHz, CDCl₃) δ 8.81 (s br, 1H), 8.37 (s, 1H), 8.18 (s br, 1H), 7.58 (s, 1H), 7.46 (s, 1H), 7.37 (s, 1H), 7.26 (s, 1H), 4.59 (t, *J* = 5.6 Hz, 2H), 4.08 (s, 3H), 4.02 (s, 3H), 3.49 (quart, *J* = 6.0 Hz, 2H), 2.56 (t, *J* = 6.0 Hz, 2H), 2.35 (s, 6H), 2.28 (t, *J* = 5.6 Hz, 2H), 2.01 (quint, *J* = 5.6 Hz, 2H); LRMS (ES⁺) *m/z* (rel intensity) 571 (M+H⁺, 75%); HRMS (M+H⁺, C₂₅H₃₁ClN₁₀O₄⁺) calcd: 571.2300; Found: 571.2297.

f-IP(C₃NHAc)I (5**)**—Acetyl chloride (24 μ L, 0.034 mmol) was added dropwise to a room temperature solution of f-IP(C₃NH₂)I (**1**) (18 mg, 0.03 mmol) and dry Et₃N (4 μ L, 0.03 mmol) in dry DCM (10 mL) and the solution was stirred overnight. The solution was quenched with water and pH was brought to 7 (1M NaOH). The solution was extracted with

DCM (3 × 10 mL), the organic layers dried over anhydrous Na₂SO₄ and the solvent was removed by evaporation to yield a brown residue purified by flash column chromatography eluting in MeOH:CHCl₃ (0:100–100:0) to give acetamido f-IPI (**5**) as a light yellow solid (15 mg, 79%), *R*_f = 0.27 (5:15:1 MeOH:CHCl₃:NH₄OH); IR (ATR) ν 3482, 3257, 3079, 2952, 1654, 1630, 1551, 1534, 1455, 1438, 1306, 1257, 1211, 1185, 1167, 1151, 1126, 1090, 1053, 1039, 1013, 897, 864, 797, 775; NMR (400 MHz, CDCl₃) δ 8.90 (s, 1H), 8.41 (s, 1H), 8.38 (s, 1H), 8.33 (s, 1H), 7.72 (t br, *J* = 4.0 Hz, 1H), 7.45 (s, 2H), 7.36 (s, 1H); 7.35 (d, *J* = 1.5 Hz, 1H), 6.86 (d, *J* = 1.5 Hz, 1H); 6.76 (t br, *J* = 4.0 Hz, 1H), 4.41 (t, *J* = 6.0 Hz, 2H); 4.07 (s, 3H); 4.03 (s, 3H); 3.52 (quart, *J* = 6.0 Hz, 2H); 3.23 (quart, *J* = 6.0 Hz, 2H); 2.60 (t, *J* = 6.0 Hz, 2H); 2.36 (s, 6H), 2.01 (s, 3H), 1.99 (quint, *J* = 6.0 Hz, 2H); LRMS (ES⁺) *m/z* (rel intensity) 570 (M+H⁺, 100%); HRMS (M+H⁺, C₂₅H₃₅N₁₁O₅⁺) calcd: 570.2909; Found: 570.2901.

f-IP(C₃NHGly)I (6**)**—A solution of f-IP(C₃ *N*-Cbz glycine)I (**12**) (10 mg, 0.0139) was hydrogenated over 10% Pd/C (10 mg) in cold MeOH (10 mL) for 2 days at room temperature and atmospheric pressure. The solution was filtered through celite and the solvent was removed by evaporation to yield glycine f-IPI (**6**) as a white solid (3 mg, 37%), mp 180°C, *R*_f = 0.27 (5:15:1 MeOH:CHCl₃:NH₄OH); IR (ATR) ν 2961, 1649, 1533, 1465, 1437, 1401, 1260, 1985, 1019, 894, 865, 767, 759; NMR (400 MHz, CD₃OD) δ 8.25 (s, 1H); 7.44 (s, 3H); 7.02 (s, 1H), 4.43 (t, *J* = 5.0 Hz, 2H), 4.05 (s, 3H), 4.02 (s, 3H), 3.55 (t, *J* = 5.0 Hz, 4H), 3.26 (s, 2H), 2.80 (t, *J* = 5.0 Hz, 2H), 2.51 (s, 6H), 2.03 (quint, *J* = 5.0 Hz, 2H); LRMS (ES⁺) *m/z* (rel intensity) 585 (M+H⁺, 25%), 293 (100%); HRMS (M+H⁺, C₂₅H₃₆N₁₂O₅⁺) calcd: 585.3008; Found: 585.3010.

RESULTS

Synthesis

As depicted in Scheme 1, synthesis of the target PAs **1–6** were accomplished using solution phase aromatic amine-acid chloride or carboxylic acid coupling reactions. Chloro f-IPI **2** was synthesized by catalytic reduction of the nitro group in amide **7** to an amine, which was immediately coupled to the freshly prepared acid chloride **8**²⁹ giving diamide **9** in 21% yield (Supporting Information). Subsequent reduction of the nitro group of diamide **9** followed by coupling of the resulting amine with the acid chloride of imidazole **10** gave the triamide **11** in 25%. Catalytic hydrogenation of **11** followed by reaction of the amine with acetic formic anhydride afforded the desired chloro f-IPI **2** in 56% yield. The chlorine atom was displaced by an azide to afford PA **3** in 96% yield. Even though the synthesis of diamino f-IPI **1** was recently reported by us, and that involved catalytic hydrogenation of the corresponding azide f-IPI PA **3**,²⁸ herein we report that the chloro moiety of PA **2** could also be directly displaced with ammonia to give the diamino **1** via the S_N2 reaction. The isolated yield of **1** was 31%. Reaction of diamine **1** with acetyl chloride yielded acetamide **5** in 79%, and reaction of **1** with *N*-cbz-glycine in presence of EDCI gave the desired PA **12** in 42% yield. The protecting cbz group was removed by catalytic hydrogenolysis to afford glycine f-IPI **6** in 37% yield.

The preparation of alkyne f-IPI **4** is depicted in Scheme 2 and it required a 4-nitropyrrole-*N*1-pentyne synthon. Coupling of the amine obtained from the reduction of amide **7** with the acid chloride of nitro-pyrrolealkyne acid **13**²⁹ afforded diamide **14** in 75% (Supporting Information). Subsequent reduction of the nitro group in PA **14** afforded an amine intermediate which was immediately coupled with *N*-foramido-1-methylimidazole-2-carboxylic acid **15**³⁰ using PyBOP and DiPEA. The desired polyamide **4** was isolated in 24% yield.

Surface Plasmon Resonance—Surface plasmon resonance (SPR) biosensor experiments were performed to obtain quantitative binding affinities, stoichiometries, and selectivities of PAs **1-6**. Sensorgrams for the binding of PA **6**, which like **1** is a dication but **6** has a significantly different structure, to the cognate sequence (5'-ACGCGT-3') are shown as an example in Figure 2. The sensorgrams and fitting of the data show a strong, highly-positive, cooperative process for binding PA **6** as well as the other PAs to the sequence 5'-ACGCGT-3' with a concentration dependent observed on-rate as expected for a PA-DNA complexation process and a first-order, extremely slow, off-rate. Steady-state RU values were obtained by equilibrium fitting and the observed $RU_{\max}$ values clearly indicate a 2:1 complex formation. The binding constants for PA **6** are $K_1 = 2.4 \times 10^5$, $K_2 = 1.1 \times 10^9$, and $K_{eq} = 1.6 \times 10^7$, where $K_{eq} = \sqrt{K_1 \times K_2}$, and were determined using a 2:1 binding model. The binding constants for compounds **1-3**, **5**, **6** are listed in Table I and were calculated using the same 2:1 binding model. PA **1** exhibited the strongest overall binding to the cognate DNA sequence, 5'-ACGCGT-3'. Binding affinities for PA **4** were not determined by SPR because the compound aggregated and exhibited surface problems with the biosensor flow cells, something occasionally observed with modified PAs. The binding affinity and kinetic behavior for PAs **1-3**, **5**, and **6** are consistent for strong DNA binders. The rate constant for dissociation, k_d , for PA **6** is extremely low as shown in Figure 2. Because of the extremely slow off-rate for PA **6**, an accurate value for k_d could not be determined by a kinetic fit of the data without an unrealistically long dissociation time being used. This is also true for PAs which have extremely fast on (k_a) and off rates that are close to the instrument injection mixing time. However, it is possible to estimate an apparent half-life for the dissociation of PA-DNA complexes from the sensorgrams. These are very useful for comparison and are included in Table III. The apparent half-lives for the dissociation of dicationic PAs, **1** and **6**, from the PA-DNA complex are >8.3 and 4.2 h, respectively. The higher binding affinities for PAs **1** and **6** are the direct result of the extremely slow rate of dissociation of the PAs from the DNA complex.

Isothermal Titration Calorimetry—To better understand the influence of substituents at the central pyrrole of f-ImPyIm, the interactions of PAs **1-6** with the DNA sequence 5'-ACGCGT-3', ITC experiments were conducted and complete thermodynamic profiles are given in Table I. An ITC titration thermogram and the corresponding integrated heats of the reaction for PA **6** with the sequence 5'-ACGCGT-3' are shown as an example in Figure 3. The excess DNA method for ITC is employed to determine ΔH and not K_{eq} because some of the binding affinities are generally too high to be determined by ITC. Because binding affinities are determined via SPR, there is not a need to use ITC to determine ΔG . The excess DNA method for ITC allows for a very accurate determination of ΔH since a consistent measurement of Q (heat) up to a molar range of ~0.6:1 (moles of ligand: moles DNA) allows a linear fit to determine ΔH of the reaction.²⁶ The enthalpy of the reaction for PA **6** is $-4.5 \text{ kcal mol}^{-1}$ from the linear fit of the integrated heats shown in Figure 3. The Gibbs-Free energy (ΔG) for the interactions of PAs **1-6** with the cognate DNA sequence 5'-ACGCGT-3' are collected with ΔH values in Table I. ΔG of the reaction for PA **6** with 5'-ACGCGT-3' is quite favorable and was calculated to be $-10.6 \text{ kcal mol}^{-1}$. $T\Delta S$ for the interaction of PA **6** is calculated to be $6.1 \text{ kcal mol}^{-1}$ using the ΔG value determined by SPR and the ΔH value determined by ITC. The binding of PA **6** to its cognate sequence is thus driven by a combination of both favorable enthalpy and entropy contributions, though a slightly larger entropy contribution is noted. PAs **1** and **5** have very similar enthalpy and entropy contributions (Table I).

Thermal Denaturation—Thermal denaturation studies were used as a qualitative evaluation of the affinity and sequence specificity of PAs **1-6** using the DNA sequences 5'-ACGCGT-3' and 5'-A₃T₃-3'. Melting temperatures (ΔT_m) values were determined from the

difference in melting temperatures of the DNA–PA complex and duplex DNA alone. ΔT_m for PAs **1–6** for both sequences are shown in Table II. PAs **1** and **6**, with cationic substituents, showed very high ΔT_m values while ΔT_m values for PAs **2–4** with 5'-ACGCGT-3' were less remarkable. All six PAs showed little or no significant change in the DNA T_m value with the sequence 5'A₃T₃-3', as shown in Table II.

DNase I Footprinting—DNase I footprinting studies provide high-resolution sequence selectivity of the PAs synthesized for this study. The autoradiogram given in Figure 4 shows that a footprint for PA **2** at the 5'-ACGCGT-3' sequence is clearly evident at a concentration of 1 μ M. No footprints other than for the cognate sequence appear for PA **2**. The audiogram for PA **1** shows a footprint emerging at 0.01 μ M and a clear footprint by 0.05 μ M for the sequence 5'-ACGCGT-3' (Table II). This supports the SPR results showing that PA **1**, the diamino derivative, exhibits the highest binding affinity to the cognate sequence 5'-ACGCGT-3'. Table II lists the concentrations at which a clear footprint was evident for the sequence 5'-ACGCGT-3' for all six PAs.

Circular Dichroism—Structural investigation of the binding properties of PAs **1–6** were conducted with circular dichroism (CD) methods and the results for PA **6** are shown in Figure 5. The titration of PA **6** into the cognate oligonucleotide 5'-ACGCGT-3' produced a strong, positive DNA-induced band at ~330 nm which is indicative of the compound binding to the minor groove of DNA.³¹ A weaker induced CD band was observed for the noncognate sequence 5'-A₃T₃-3'. Compounds **1–5** also showed a DNA-induced band at 330 nm, thus indicating all these PAs also bind to the minor groove of DNA. Further evidence that the PAs are binding to the DNA via a single mechanism is the presence of an isodichroic point for the overlaid spectra at ~310 nm (shown for PA **6** in Figure 5), most likely as a stacked homodimer. This holds true for compounds **1–5** also.

DISCUSSION

Since discovery and isolation of netropsin by Finlay et al. over 60 years ago,³² the study of netropsin, distamycin and their complexation with AT-rich sequences of the minor groove of DNA has become well recognized and accepted.^{33–39} The discovery that distamycin could form a “stacked,” anti-parallel dimer in the minor groove of DNA of varying AT sequences opened the door and created a rationale for synthetic PA drug design based on DNA sequence.^{1–6} The systematic definition of Im and Py containing synthetic PAs binding rules for sequence specific base pair recognition of the minor groove for Watson-Crick duplex DNA by Dervan and coworkers has paved the way for the development of an unlimited library of synthetic PAs which can be programmed to bind to any particular DNA sequence.^{2,3} PAs have shown excellent in vitro activity and the development and progress of PAs as therapeutics, although hindered by solubility and cellular uptake, is progressing rapidly and positive biological results have been reported.^{2–4,13–16,40}

Our approach to deal with these problems is to synthesize heterocyclic PAs with improved PA-PA and PA-DNA recognition and affinity. A promising development area is modification of the N1-position of the pyrrole and imidazole ring systems. While a number of changes have been reported at that position and H-pin PAs are generally linked through the same N1-position of the center heterocycle, more detailed studies with systematic modifications are needed. To accomplish this goal we added alkyl cationic groups to the N1-position and have now extended the systematic analysis with the analogs shown in Figure 1.

The strategy is to use PAs, such as f-ImPyIm, as a test system and through the addition of different alkyl based side chains at the N1-position of the central pyrrole heterocycle or at other heterocycles to improve the properties of the PAs for targeting DNA. The importance

of using f-ImPyIm derivatives to target the sequence 5'-ACGCGT-3' is the presence of that particular sequence in the promoter region of the core sequence of the *MluI* cell-cycle box which is a transcriptional element of the human *Dbf4* gene. Again, the *Dbf4* gene is a regulatory subunit of the Cdc7 kinase (a cyclin dependent kinase) which in high levels has been implicated in various cancers.

The f-ImPyIm PA has a very high binding affinity and the analogs cause moderate to significant decreases in DNA binding. The K for PA **1** is within experimental error of that of the parent compound (Table I) and it has a very slow dissociation rate. PAs **1** and **6** demonstrated the highest binding affinity of the new f-ImPyIm analogue compounds to the cognate sequence 5'-ACGCGT-3' with good sequence selectivity. Both of these PAs have a second positive charge that contributes to the higher binding affinities compared to the other four PA derivatives. This is presumably due to an additional electrostatic interaction between the positively charged NH_3^+ present on the distal end of the alkyl side chain on both PAs **1** and **6** and the negatively charged phosphate backbone of DNA. The length of the alkyl side chain allows it to extend out of the minor groove of the DNA and interact electrostatically with the phosphate back bone. The larger enthalpic contribution to the overall Gibbs-free energy of the reaction for PA **1** binding to the sequence 5'-ACGCGT-3' would support this explanation for the increased binding affinity. The cationic modification of PA **6** is more bulky than the N-Me substituent of PA **1** and this clearly causes some steric hindrance and reduced binding affinity. In agreement with this observation, the uncharged compounds bind significantly more weakly than the parent f-ImPyIm as well as the dicationic compounds. This again probably represents some decrease in DNA interactions due to unfavorable steric effects without the favorable interactions of a second cation. The results of this study demonstrate that by making systematic modifications to PAs, such as f-ImPyIm, the binding affinity can be altered through the modulation of k_d and/or k_a .

The dicationic PAs **1** and **6** demonstrated comparable binding affinities and slower off rates than the parent compound f-ImPyIm while having increased water solubility. Increased water solubility and reduced aggregation address one of three major advancements that must be made in order for PAs to become approved therapeutics. The other two major areas of concern, sequence specificity and binding affinity, remained within experimental error of the parent compound, f-ImPyIm. Our initial thoughts on the noncationic substituent PAs of Figure 1 were that we might see some improvements in binding due to their offset stacking and possible hydrophobic/van der Waals interactions of the substituents. The unfavorable steric clash is the opposite effect and presumably canceled some of the positive effects of the cationic substituents. Clearly, however, modifications of the $-\text{N}-\text{CH}_3$ groups of both Py and Im heterocycles are promising for improving their solution properties and DNA interactions. Additional studies on such functionalized PAs should yield new ideas for these very useful, sequence-specific compounds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Contract grant sponsor: NIH

Contract grant number: AI-064200

Contract grant sponsor: NSF

Contract grant numbers: CHE 0809162, CHE 0922623

Contract grant sponsor: Cancer Research UK

Contract grant number: C2259/A9994

Contract grant sponsor: Georgia Research Alliance

References

1. Lacy, ER.; Madsen, EM.; Lee, M.; Wilson, WD. DNA and RNA Binders, From Small Molecules to Drugs. Wiley-VCH; Weinheim: 2003. p. 384-413.
2. Dervan PB, Edelson BS. *Curr Opin Struct Biol.* 2003; 13:284–299. [PubMed: 12831879]
3. Dervan PB. *Bioorg Med Chem.* 2001; 9:2215–2235. [PubMed: 11553460]
4. Suckling C. *Future Med Chem.* 2012; 4:971–989. [PubMed: 22650239]
5. Pelton JG, Wemmer DE. *Proc Natl Acad Sci USA.* 1989; 15:5723–5727. [PubMed: 2762292]
6. Pelton JG, Wemmer DE. *J Am Chem Soc.* 1990; 112:1393–1399.
7. Mitra SN, Wahl MC, Sundaralingam M. *Acta Crystallogr D.* 1999; 55:602–609. [PubMed: 10089456]
8. Bando T, Sugiyama H. *Acc Chem Res.* 2006; 39:935–944. [PubMed: 17176032]
9. Lacy ER, Le NM, Price CA, Lee M, Wilson WD. *J Am Chem Soc.* 2002; 124:2153–2163. [PubMed: 11878969]
10. Buchmueller KL, Staples AM, Howard CM, Horick SM, Uthe PB, Le NM, Cox KK, Nguyen B, Pacheco KA, Wilson WD, Lee M. *J Am Chem Soc.* 2005; 127:742–750. [PubMed: 15643900]
11. Satz AL, Bruice TC. *Acc Chem Res.* 2002; 35:86–95. [PubMed: 11851386]
12. Dervan PB, Doss RM, Marques MA. *Curr Med Chem Anticancer Agents.* 2005; 5:373–387. [PubMed: 16101489]
13. Chou CJ, Farkas ME, Tsai SM, Alvarez D, Dervan PB, Gottesfeld JM. *Mol Cancer Ther.* 2008; 7:769–778. [PubMed: 18413791]
14. Shinohara K, Bando T, Sugiyama H. *Anticancer Drugs.* 2010; 21:228–242. [PubMed: 20038830]
15. Hochhauser D, Kotecha M, O'Hare C, Morris PJ, Hartley JM, Taherbhai Z, Harris D, Forni C, Mantovani R, Lee M, Hartley JA. *Mol Cancer Ther.* 2007; 6:346–354. [PubMed: 17237293]
16. Minoshima M, Bando T, Shinohara K, Sugiyama H. *Nucleic Acids Symp Ser.* 2009; 23:69–70.
17. Buchmueller KL, Bailey SL, Matthews DA, Taherbhai ZT, Register JK, Davis ZS, Bruce CD, O'Hare C, Hartley JA, Lee M. *Biochemistry.* 2006; 45:13551–13565. [PubMed: 17087509]
18. Buchmueller KL, Staples AM, Uthe PB, Howard CM, Pacheco KA, Cox KK, Henry JA, Bailey SL, Horick SM, Nguyen B, Wilson WD, Lee M. *Nucleic Acids Res.* 2005; 33:912–921. [PubMed: 15703305]
19. Verma R, Patapoutian A, Gordon CB, Campbell JL. *Proc Natl Acad Sci USA.* 1991; 88:7155–7159. [PubMed: 1871128]
20. Wu X, Lee H. *Oncogene.* 2002; 21:7786–7796. [PubMed: 12420215]
21. Hess GF, Drong RF, Weiland KL, Slightom JL, Scalafani RA, Hollingsworth RE. *Gene.* 1998; 211:133–140. [PubMed: 9573348]
22. Babu B, Liu Y, Plaunt A, Riddering C, Ogilvie R, Westrate L, Davis R, Ferguson A, Mackay H, Rice T, Chavda S, Wilson WD, Lin S, Kiakos K, Hartley JA, Lee M. *Biochem Biophys Res Commun.* 2011; 404:848–852. [PubMed: 21185262]
23. Satam V, Babu B, Chavda S, Savagian M, Sjöholm R, Tzou S, Ramos JP, Liu Y, Kiakos K, Lin S, Wilson DW, Hartley JA, Lee M. *Bioorg Med Chem.* 2012; 22:693–701. [PubMed: 22222156]
24. Nguyen B, Tanious F, Wilson WD. *Methods.* 2006; 42:150–161. [PubMed: 17472897]
25. Nanjunda, R.; Munde, M.; Liu, Y.; Wilson, WD. *Methods for Studying Nucleic Acid/Drug Interactions.* CRC; Boca Raton: 2012. p. 91-122.
26. Ren J, Jenkins TC, Chaires JB. *Biochemistry.* 2000; 39:8439–8447. [PubMed: 10913249]
27. 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 SA, Lin S, Kiakos K, Bruce CD, Wilson WD, Hartley JA, Lee M. *Biochemistry.* 2011; 50:3127–3136. [PubMed: 21388229]

28. Babu B, Liu Y, Plaunt A, Riddering C, Ogilvie R, Westrate L, Davis R, Ferguson A, Mckay H, Rice T, Ramos JP, Chavda S, Wilson WD, Lin S, Kiakos K, Hartley JA, Lee M. *Biochem Biophys Res Commun.* 2011; 404:848–852. [PubMed: 21185262]
29. The synthesis of 1-(3'-chloropropyl)-4-nitropyrrole-2-carboxylic acid **8** and 1-(4'-pentynyl)-4-nitropyrrole-2-carboxylic acid **10** will be reported elsewhere.
30. Mulder K, Serton J, Taherbhai Z, Jones J, Uthe P, Brown T, Lee M. *Synth Commun.* 2008; 38:33–44.
31. Rodger, A.; Norden, B. *Circular Dichroism and Linear Dichroism.* Oxford University; New York: 1997.
32. Finlay AC, Hochstein FA, Sobina BA, Murphy FX. *J Am Chem Soc.* 1951; 73:341–343.
33. Wang S, Munde M, Wang S, Wilson WD. *Biochemistry.* 2011; 50:7674–7683. [PubMed: 21800847]
34. Bailly C, Chaires JB. *Bioconjug Chem.* 1998; 9:513–538. [PubMed: 9736486]
35. Wemmer DE. *Biopolymers.* 2001; 52:197–211. [PubMed: 11295751]
36. Dervan PB, Burli RW. *Curr Opin Chem Biol.* 1999; 3:688–693. [PubMed: 10600731]
37. Wemmer DE, Dervan PB. *Curr Opin Struct Biol.* 1997; 7:355–361. [PubMed: 9204277]
38. White S, Baird EE, Dervan PB. *Chem Biol.* 1997; 4:569–578. [PubMed: 9281524]
39. Chenoweth DM, Dervan PB. *Proc Natl Acad Sci USA.* 2009; 106:13175–13179. [PubMed: 19666554]
40. Edwards TG, Koeller KJ, Slomczynska U, Fok K, Helmus M, Bashkin JK, Fisher C. *Antiviral Res.* 2011; 91:177–186. [PubMed: 21669229]

FIGURE 1.

Structures of the “parent” compound (f-ImPyIm), f-ImPy((CH₂)₃NH₃)Im (**1**), f-ImPy((CH₂)₃Cl)Im (**2**), f-ImPy((CH₂)₃N₃)Im (**3**), f-ImPy((CH₂)₃alkyne)Im (**4**), f-ImPy((CH₂)₃NHAc)Im (**5**), and f-ImPy((CH₂)₃NHGGly)Im (**6**).

FIGURE 2.

SPR sensorgrams for the binding of f-ImPy((CH₂)₃NHGly)Im (**6**) to the sequence 5'-ACGCGT-3' (A) and a plot of RU_{ss} vs. concentration including a best-fit line (B). The cooperative ($K_2 > K_1$) binding of two f-ImPy((CH₂)₃NHGly)Im molecules per 5'-ACGCGT-3' is observed.

FIGURE 3.

Experimental ITC data (A) for the excess DNA method titration of PA **6** into 5'-ACGCGT-3' and the integrated heats for each injection of **6** used for the titration (solid dots) and a linear fit of that data which is extrapolated to zero and yields of the ΔH of the reaction (B).

FIGURE 4.

DNase I footprinting of compound **2** (f-ImPy((CH₂)₃Cl)Im) on the antisense strand of the 5'-[³²P] 130-bp fragment. All reactions contain ~500 cps DNA fragment, 10 mM Tris, pH 7, 1 mM EDTA, 50 mM KCl, 1 mM MgCl₂, 0.5 mM DTT, 20 mM HEPES. G + A denotes the purine sequencing lane. The 5'-ACGCGT-3' site is indicated by a solid bar.

FIGURE 5.

CD data for compound **6**, (f-ImPy((CH₂)₃NHGly)Im), with the sequence 5'-ACGCGT-3'. CD experiments were carried out using 160 μ L of a 9 μ M DNA solution, which was titrated with 1 mol equivalents of **6**, past the point of saturation. The spectra represent titration from DNA alone to 1 mol equivalent of PA **6** to the DNA hairpin, 2, 3, 4, 5, 6, 8, and 10 molar equivalents.

SCHEME 1.

Toward the synthesis of the target compounds 1-3, 5 and 6. (a) H₂, 5% Pd-C, MeOH, atmospheric pressure; (b) SOCl₂, reflux, 15 min; (c) DCM, TEA, ice-bath to room temperature overnight; (d) oxalyl chloride, THF, reflux; (e) acetic formic anhydride, DCM, ice-bath to room temperature overnight; (f) NH₃, MeOH, sealed-tube, rt, 6 h; (g) H₂, 10% Pd-C, MeOH, atmospheric pressure reported in Ref. ²²; (h) NaN₃, DMF, 80°C, overnight; (i) acetyl chloride, DCM, TEA, ice-bath to room temperature overnight; (j) N-Cbz-glycine, EDCI, DiPEA, DMAP, DMF, rt overnight; (k) H₂, 10% Pd-C, MeOH, atmospheric pressure.

SCHEME 2.

Toward the synthesis of compound 4. (a) H_2 , 5% Pd-C, MeOH, atmospheric pressure; (b) SOCl_2 , reflux, 15 min; (c) DCM, TEA, ice-bath to room temperature overnight; (d) stannous chloride dihydrate, HCl, EtOH, 2 h; (e) PyBOP, DMF, DiPEA, argon, rt, overnight.

Table I

Results for ITC and SPR Experiments on PAs 1–3, 5 and 6 with the Sequence 5'-ACGCGT-3'

PA	ΔH^a (kcal mol ⁻¹)	ΔG^b (kcal mol ⁻¹)	$T\Delta S^c$ (kcal mol ⁻¹)	K_{eq} (M ⁻¹) ^b
1	-5.9	-11.4	5.5	2.8×10^8
2	-4.4	-8.3	3.9	1.2×10^6
3	-4.5	-9.8	5.3	1.5×10^6
5	-5.7	-9.1	3.4	4.9×10^6
6	-4.5	-10.6	6.1	6.2×10^7
f-ImPyIm ¹⁶	-7.6	-10.7	3.1	1.9×10^8

$$K_{eq} = \sqrt{K_1 \times K_2};$$

^aDetermined by ITC.^bDetermined from SPR studies.

^c Calculated from the equation $\Delta S = (\Delta G - \Delta H)/T$, using ΔG calculated from SPR and ΔH calculated from ITC. K values for PA 4 were unable to be determined by SPR because of surface problems. Included at the bottom are the thermodynamic properties and K_{eq} value for the “parent” compound, f-ImPyIm, as where previously determined and from which these modified PAs were derived.

Table II

Results for Thermal Denaturation and DNase I Footprinting Experiments for PAs 1–6 and the Parent Compound f-ImPyIm from Previously Reported Work

PA	ΔT_m (°C)		Footprinting (μM) ^a
	5'-ACGCGT-3'	5'-A ₃ T ₃ -3'	
1	>24	10	0.05
2	1.5	0	1
3	4.8	0	1
4	0	0	10
5	10	1	0.1
6	20	5	0.5
f-ImPyIm ²⁷	7.3	NA	0.05

NA—not available.

^aConcentration when footprint at the cognate site became apparent.

Table III

Estimated Half-Lives (in hours) for the Rate of Dissociation of PAs 1–3, 5, and 6 from the Sequence 5'-ACGCGT-3'

PA	$\sim T_{1/2}$ (app ^a) (hr)
1	>8.3
2	0.22
3	0.57
5	2.6
6	4.2

^a Apparent half-lives based on dissociation rates from SPR experiments.