Polyamide Curvature and DNA Sequence Selective Recognition: Use of 4-Aminobenzamide to Adjust Curvature

Jamie Lajiness[a], Alan Sielaff[a], Hilary Mackay[a], Toni Brown[a], Jerome Kluza[b], Binh Nguyen[c], W. David Wilson[c], Moses Lee[a,#,*] and John A. Hartley[b]

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

Abstract: Imidazole and pyrrole-containing polyamides belong to an important class of compounds that can be designed to target specific DNA sequences, and they are potentially useful in applications of controlling gene expression. The extent of polyamide curvature is an important consideration when studying the ability of such compounds to bind in the minor groove of DNA. The current study investigates the importance of curvature using polyamides of the form f-Im-Phenyl-Im, in which the imidazole heterocycles are placed in *ortho*-, *meta*-, and *para*-configurations of the phenyl moiety. The synthesis and biophysical evaluation of each compound binding to its cognate DNA sequence (5'-A**CGCG**T-3') and a negative control sequence (5'-A**AATT**T-3') is reported, along with their comparison to the parent binder, f-Im-Py-Im (**3**). A**CGCG**T is a medicinally significant sequence present in the M*lu*I cell-cycle box (MCB) transcriptional element found in the promoter of a gene associated with cell division. The results demonstrated that the *para*-derivative has the greatest affinity for its cognate sequence, as indicated *via* thermal denaturation, CD, ITC, SPR analyses, and DNase I footprinting. ITC studies showed that binding of the *para*-isomer (**2c**) to A**CGCG**T was significantly more exothermic than binding to A**AATT**T. In contrast, no heat change was observed for binding of the *meta*- (**2b**) and *ortho*- (**2a**) isomers to both DNAs, due to low binding affinities. This is consistent with results from SPR studies, which indicate that the *para*-derivative binds in a 2:1 fashion to A**CGCG**T and binds weakly to A**CCGG**T ($K = 1.8 \times 10^6$ and 4.0×10^4 M^{-1}, respectively). Interestingly, it binds in a 1:1 fashion to A**AATT**T ($K = 5.4 \times 10^5$ M^{-1}). The *meta*-compound does not bind to any sequence. The *para*-derivative also was the only compound to show an induced peak *via* CD at 330 nm, indicative of minor groove binding, and produced a ΔT_m value of 5.8 ºC. Molecular modeling experiments have been performed to determine the shape differences between the three compounds, and the results indicate that the *para*-derivative **2c** has a closest curvature to previously synthesized polyamides. DNase I footprinting studies confirmed earlier observations that only the *para*-derivative **2c** produced a footprint with A**CGCG**T (1 μM) and no significant footprint was observed at any sites examined for *meta*-**2b** and *ortho*-**2a** analogs up to 40 μM. The results of these studies suggest that the shape of the *ortho*- and *meta*- derivatives is too curved to match the curvature of the DNA minor groove to facilitate binding. The *para*-derivative gives the highest binding affinity in the series and the results illustrate that 4-aminobenzamide is a reasonable substitute for 4-aminopyrrole-2-carboxylate.

Key Words: Curvature, polyamide, DNA, gene regulation, minor groove.

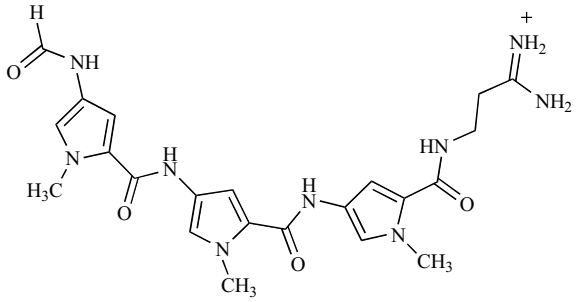
[#]**Author Profile:** Moses Lee received his B.Sc. and Ph.D. degrees from the University of Guelph, Canada. After post-doctoral studies at the University of Alberta, he worked as a research scientist at Synphar Laboratories in Edmonton before joining Furman University in South Carolina. In 2005, he moved to Hope College to become Dean of the Natural and Applied Sciences Division and Professor of Chemistry.

INTRODUCTION

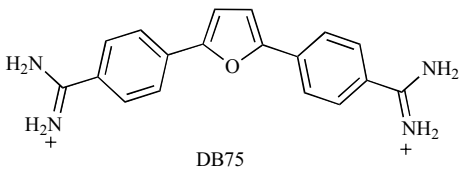
The ability of polyamides such as distamycin A (**1**) to bind to the minor groove of DNA has been extensively studied. This naturally occurring compound, composed of three *N*-methylpyrrole-2-carboxamido moieties, a formamido *N*-terminus and an amidine group at the C-terminus (Fig. (**1**)) binds in the minor groove in a stacked, anti-parallel 2:1 (ligand:DNA) fashion, and targets A•T rich DNA [1]. Modifications to distamycin A have led to the substitution of the *N*-methylpyrrole (Py) units with other heterocycles, most notably *N*-methylimidazoles (Im). These imidazole substituted polyamides bind in the same fashion as distamycin A (**1**). Investigations into the substitution have led to the development of pairing rules, in which a stacked Py/Py targets an A•T or T•A base pairing, a Im/Py targets G•C, Py/Im targets C•G, and Im/Im targets a T•G mismatch [2].

*Address correspondence to this author at the Division of Natural and Applied Sciences and Department of Chemistry, Hope College, Holland, MI, 49423, USA; Tel: +1 616 395 7190; Fax: +1 616 395 7923;
E-mail: lee@hope.edu

1573-4064/09 $55.00+.00



Distamycin, **1**



DB75

Fig. (1). Structures of distamycin A (**1**) and furamidine or DB273.

Imidazole and pyrrole-containing polyamides have significant appeal to medicinal chemists because they can be designed to target any DNA sequence for the ultimate goal of developing novel gene control agents [3]. A number of polyamides have been reported to have the capacity of controlling the expression of specific genes in cells. For example, the authors have recently reported polyamides that target the inverted CCAAT box (ICB, 5'-ATTGG-3') of the topoisomerase IIα gene [4]. A hairpin polyamide, called JH-37, was found using a ChIP assay to bind its target ICB site in the nucleus and inhibited the binding of the transcriptional factor NF-Y thereby turning "on" the expression of the topoisomerase IIα gene as indicated by western blot and RT-PCR analyses [4a].

Less extensively studied has been the relationship between polyamide 3D structure and the conformation of the DNA binding site. The minor groove of DNA has a specific curvature, with which the polyamide must match, or adapt to fit, in order to facilitate binding. Polyamides attempting to bind to DNA with incorrect structure to fit this curvature have reduced or no binding at all [5]. Wilson and co-workers have investigated the importance of curvature with the position of positively charged groups on diphenylfuramidine derivatives, for example, furamidine or DB75 (Fig. (**1**)). The amidine moieties were placed in three relative positions; *para-para*, *meta-para* and *meta-meta.* The experiments were repeated with imidazolines as well as dimethylimidazolines. Results from the experiments concluded that the *para*- substituted analogs bound strongly to the minor groove of DNA, whereas *meta*-substituted derivatives bound relatively weakly [5]. The *para*-compounds utilized a bridging H_2O molecule and this was found to be the reason for the superior binding [6]. The Cory group has also studied the effect of curvature with pentamidines, amidinobenzimidazoles and amidoindoles and found that the planarity of the molecules was important with respect to DNA binding [7].

The aim of the current report was to study the effect of curvature on DNA binding of three novel polyamides (Fig. (**2**), compounds **2a**-**c**). These compounds are analogs of the polyamide f-Im-Py-Im (Fig. (**2**, compound **3**)), shown by the authors to be a strong binder to its cognate sequence 5'-A**CGCG**T-3' [8]. The 5'-A**CGCG**T-3' sequence is very important in cancer research because it is present in the core sequence of the M*lu*I cell-cycle box (MCB) transcriptional element found in the promoter of the human Dbf4 (huDbf4 or ASK , activator of S-phase kinase) gene. Dbf4 is the regulatory subunit of Cdc7 (cyclin dependent 7) kinase, and high levels of this kinase have been implicated for development of various cancers [9]. Cdc 7 kinase promotes cancer cells to enter into and traverse through S phase. In an attempt to improve the binding affinity of polyamides to DNA, specifically to A**CGCG**T, the focus of this paper examines the effects of polyamide curvature on binding affinity. By replacing the Py moiety of compound **2** with a phenyl (Ph) and substituting the imidazole groups at the *ortho* (**2a**)- *meta* (**2b**)- and *para* (**2c**)- positions, three polyamides of varying curvature were produced of the form f-Im-Ph-Im (Fig. (**2**)). This study illustrates the synthesis of the three Ph-derivatives and the binding characteristics of these compounds to the cognate A**CGCG**T (Fig. (**3**)), and non-cognate A**AATT**T (Fig. (**3**)) DNA sequences; determined by thermal denaturation (ΔT_m), circular dichroism (CD), and isothermal titration microcalorimetry (ITC) studies. Surface Plasmon Resonance (SPR), molecular modeling and DNase I footprinting studies were also performed.

RESULTS & DISCUSSION

Synthesis

Compounds **2a**-**c** were synthesized according to Scheme **1**. The nitro group of imidazolecarboxamide (**4**) was reduced using standard Pd-C catalyzed hydrogenation conditions as reported previously [10] and the resulting amine was reacted with the relevant commercially available nitrobenzoyl chloride using standard Schotten-Bauman coupling conditions [11]. The final imidazole heterocyle was added in the same manner as described in previously reports [9], and the formylation step was performed using the published procedure [10]. All compounds were obtained in good yield and pure according to TLC and ^{1}H-NMR analyses. Their structures were characterized by standard techniques: ^{1}H NMR, IR, LRMS and/or accurate mass measurements.

Thermal Denaturation (T_m) Studies

Thermal denaturation studies were conducted in order to gauge the ability of the polyamide to stabilize DNA upon heating. The results, summarized in Table **1**, indicate that the *para*- derivative (**2c**) has the strongest affinity to the cognate sequence (A**CGCG**T) with a ΔT_m of 5.8 °C. Both *ortho*- (**2a**) and *meta*- (**2b**) derivatives produced ΔT_m < 1°C, suggesting minimal to no binding. In comparison to f-Im-Py-Im

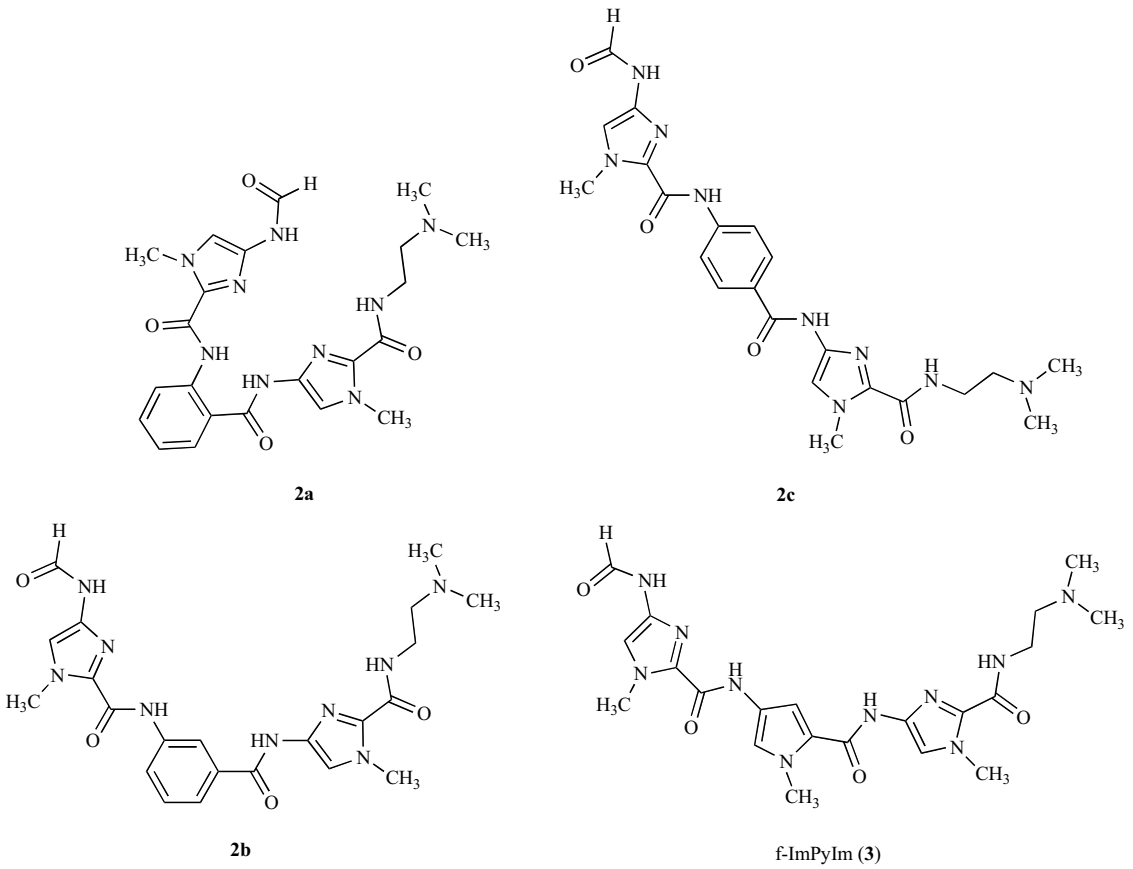


Fig. (2). Structures of the phenyl-derivatives: *ortho* (**2a**), *meta* (**2b**) and *para*-substituted (**2c**); and f-ImPyIm (**3**).

(**3**), the *para*-derivative (**2c**) binds with less affinity (*c.f.* ΔT_m = 5.8 *vs.* 7.8 °C) and also reduced selectivity, as it also binds to the A/T rich negative control (ΔT_m = 7.5 *vs.* 0.9 °C).

Circular Dichroism (CD) Studies

CD studies were performed to confirm that the polyamides were binding in the minor groove. CD reflects the ability of chiral molecules to differentially absorb circularly polarized light. B-form DNA produces a characteristic peak at 275 nm, due to its inherent chirality. When the non-chiral polyamide binds to DNA, the chiral complex gives an induced band at 330 nm [12], a characteristic typically observed for the binding of polyamides in the minor groove. Results from these studies, (Fig. (**4**)) suggest the preferential binding of the *para*-derivative (**2c**) to its cognate sequence A**CGCG**T over A**AATT**T as indicated by the more defined and larger induced CD band at ~330 nm (36 and 12 mdeg, respectively) (Fig. (**4A**, **B**)). The results also suggest preferential binding of the *para*-derivative (**2c**) over and above the *meta*- (**2b**) (Fig. (**4C**, **D**)) and *ortho*- (**2a**) derivatives (Fig. (**4E**, **F**)), as both compounds illustrate a minimal and/or no induced band at ~330 nm. Tha *para*-compound (**2c**) also produces a 2:1 binding stoichiometry. With regard to f-ImPyIm (**3**), similar results are observed for both compound **2c** and f-Im-Py-Im (**3**), with respect to binding stoichiometry and CD response. Both compounds produce a CD ellipticity of ~35 mdeg for the cognate sequence (A**CGCG**T) and a response of ~ 13 mdeg for the non-cognate (A**AATT**T) DNA sequence [10].

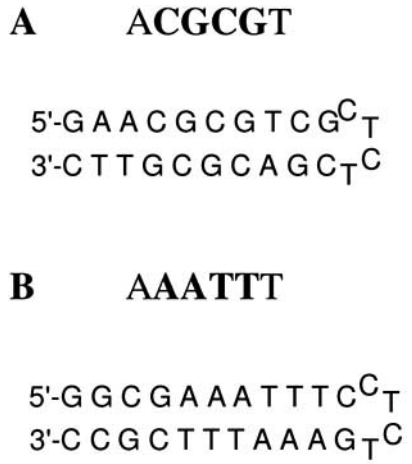
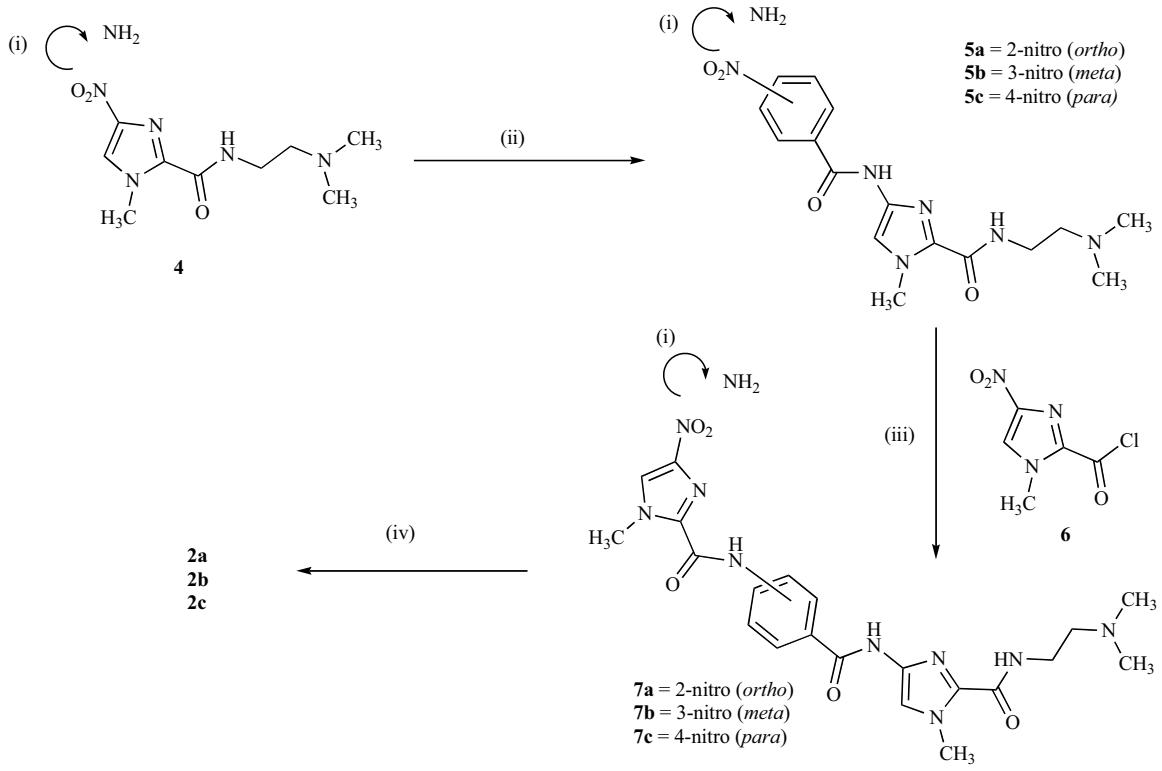


Fig. (3). DNA sequences used in the study: cognate sequence A**CGCG**T (**A**) and non-cognate sequence A**AATT**T (**B**).

Isothermal Titration Microcalorimetry (ITC) and Surface Plasmon Resonance (SPR) Studies

ITC studies enable the enthalpy of the binding reaction (ΔH) to be determined (Table **2**), and example thermograms are given in Fig. (**5**). Heat change was only observed with the *para*-derivative (**2c**) suggesting a significant preference



Scheme 1. (i) 5% Pd/C, H_2, cold MeOH, RT, ~18 h; (ii) 2-, 3-, or 4-nitrobenzoyl chloride, dry TEA, dry DCM, 0-RT ºC, ~ 18h; (iii) **6**, dry TEA, dry DCM, 0-RT ºC, ~18 h; (iv) acetic formic anhydride, dry DCM, 0-RT ºC, ~ 18h.

for the cognate sequence (A**CGCG**T, Fig. (**5A**)) over A**AATT**T (Fig. (**5B**)).

Surface Plasmon Resonance (SPR) Studies

SPR provides a direct method for determining the binding affinity and the kinetics of binding of small molecules with macromolecules, such as DNA [13]. Consistent with data from ITC experiments, the results show that the *para*-derivative (compound **2c**) does bind to its cognate sequence (A**CGCG**T (Fig. (**6A**, **Part A**))) with a binding affinity of 1.8 x 10^6 M^{-1} (Table **2**). Some binding is observed with A**CCGG**T (K = 4 x 10^4 M^{-1}). With another non-cognate DNA: A**AATT**T, 1:1 binding was observed with a binding affinity of 5.4 x 10^5 M^{-1} ((See Fig. (**6B**, **Part A**)). Fig. (**6 Part B**) shows the minimal binding observed for the *meta*-compound (**2b**). f-ImPyIm (**3**) does have higher binding affinity (*c.f.* 10^8 *vs*. 10^6 M^{-1}) and this compound has less affinity for A**AATT**T [8]. Interestingly, the dissociation rate con-

Table 1. Results from Thermal Denaturation Studies

	(ºC)					
	ACGCGT			AAATTT		
Compound	**T_m DNA**	**T_m DNA + PA**	**ΔT_m**	**T_m of DNA**	**T_m DNA + PA**	**ΔT_m**
Ortho- (**2a**)	71.8	71.6	**<1.0**	56.6	56.4	**<1.0**
Meta- (**2b**)	71.3	71.2	**<1.0**	58.3	59.2	**<1.0**
Para- (**2c**)	71.3	77.1	**5.8**	57.4	64.9	**7.5**
f-ImPyIm (**3**) [8]	71.2	79.0	**7.8**	54.9	55.8	**0.9**

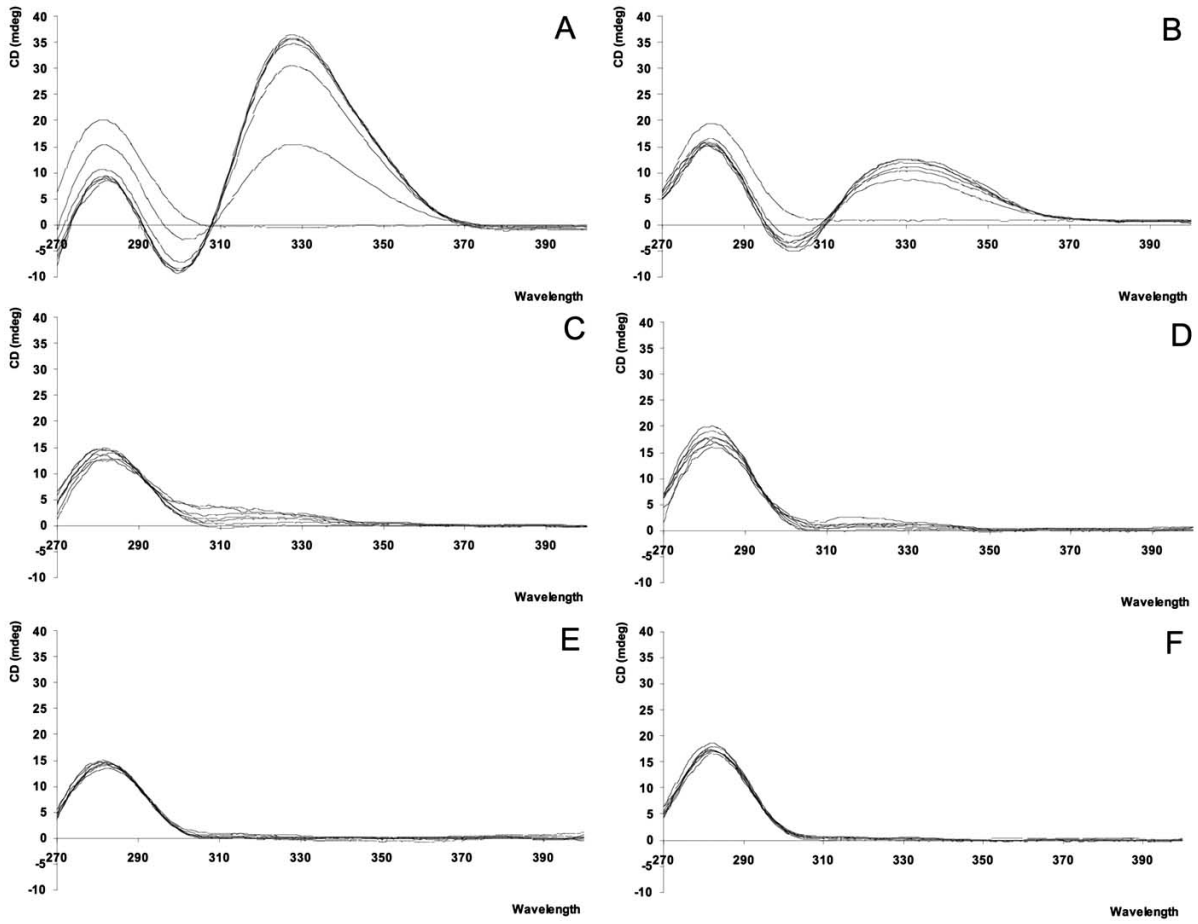


Fig. (4). CD data for the *para*-compound (**2c**) with A**CGCG**T (**A**) and A**AATT**T (**B**) the meta-compound (**2b**) with A**CGCG**T (**C**) and A**AATT**T (**D**) and the *ortho*-compound (**2a**) with A**CGCG**T (**E**) and A**AATT**T (**F**).

stant for compound **3** is significantly slower than that for the *para*-compound **2c** [8]. The reason for this difference in dissociation rates is not clear at this point.

Combining the ITC results with the SPR data enables the thermodynamic parameters (ΔG, ΔH, and TΔS) to be obtained (Table **2**). The ΔG values taken at 25 °C is more negative for

Table 2. Results from ITC and SPR Studies

		Ortho (**2a**)	*Meta* (**2b**)	*Para* (**2c**)	f-ImPyIm (**3**) [8]
ITC kcal mol^{-1}	**ΔH** A**CGCG**T A**AATT**T	No heat No heat	No heat No heat	-4.2	-7.6
kcal mol^{-1}	[a]**TΔS** A**CGCG**T A**AATT**T	na na	na na	4.3	3.1
kcal mol^{-1}	[b]**ΔG_{eq}** A**CGCG**T A**AATT**T	na na	na na	-8.5	-11.2
SPR M^{-1}	**K_{eq}** A**CGCG**T A**AATT**T A**CCGG**T	nd nd nd	nb nb nb	$1.8x10^6$ [c]$5.4x10^5$ $4.0x10^4$	$1.9x10^8$ $5.3x10^4$ $2.2x10^5$

[a]ΔG=ΔH-TΔS; [b]calculated from $\Delta G = -RT\ln K_{eq}$; [c]1:1 binding observed. na = not applicable; nd = not determined; nb = no binding.

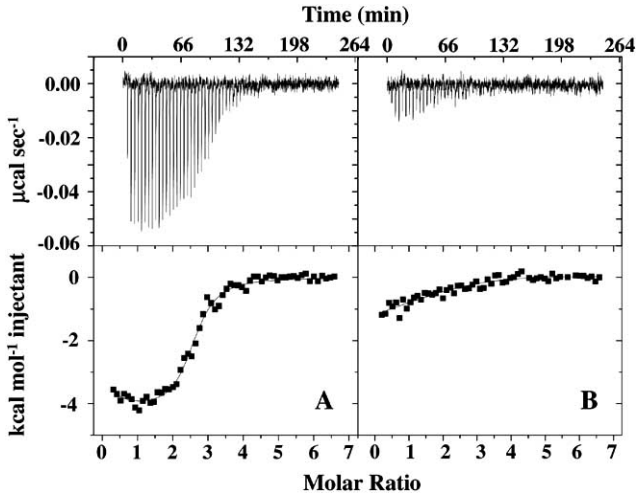


Fig. (5). ITC thermograms for the *para*-compound (**2c**) with A**CGCG**T (**A**) and A**AATTT** (**B**).

f-ImPyIm (**3**) than for the *para*-compound (**2c**) (c.f. -11.2 *vs*. -8.5 kcal mol^{-1}) and the binding of both compounds is strongly driven by enthalpy.

Molecular Modeling Studies

Molecular modeling studies show that the shape of the *para*- (**2c**) and *meta*-compounds (**2b**) is different from f-ImPyIm (**3**) and this could account for the reduction in binding. Fig (**7**) clearly shows that the *para*-derivative (**2c**) is less curved and more closely resemble to that of compound **3**. The *meta*-derivative (**2b**) is more curved than compound **3**; the formamido-*N*-terminus in all cases is aligned very nicely, with the C-terminus *N,N*-dimethylaminoethyl group further accentuating the curvature.

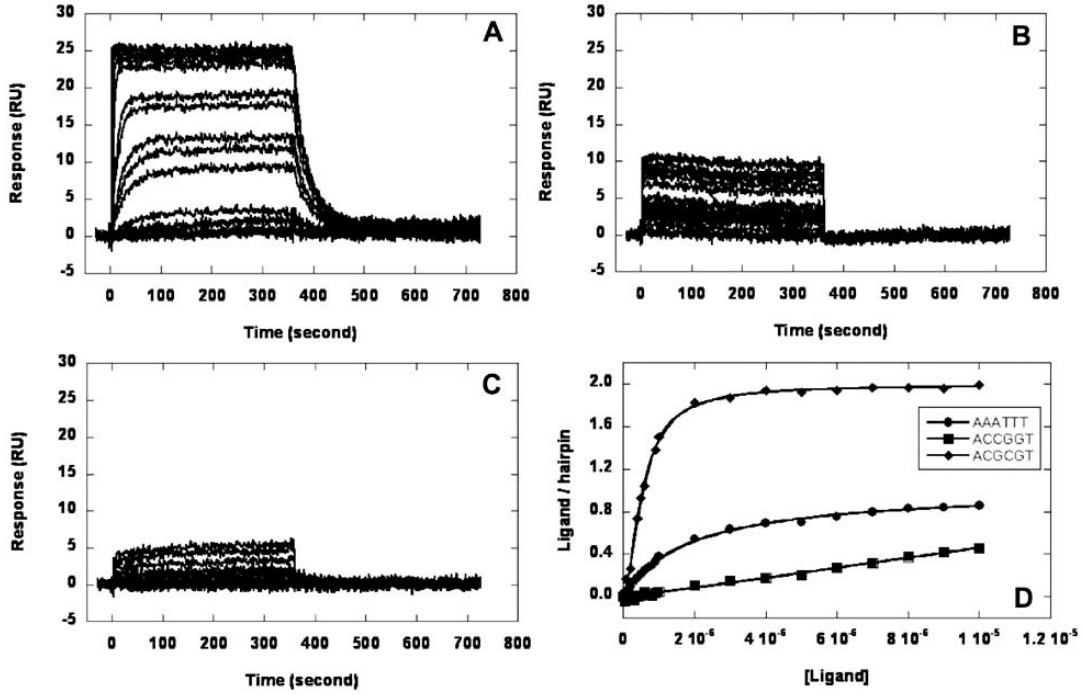
DNase I footprinting Studies

To compare the sequence selectivity and affinity of f-ImPyIm (**3**) with the derivatives *ortho*-, *meta*-, and *para*- f-Im-Ph-Im derivatives (**2a-c**, respectively), DNase I footprinting experiments were carried out using a 131 base pair 5'-^{32}P-radiolabeled DNA fragment that contained the sequences 5'-A**CGCG**T-3', 5'-A**CCGG**T-3', 5'-A**CGTG**T-3' and 5'-A**GCGC**T-3' (Fig. (**8**)) [8]. Consistent with previously reported results protection of DNase I mediated cleavage by f-ImPyIm (**3**) for the cognate sequence A**CGCG**T was observed from 0.05 µM. Binding at 5 µM for the sites containing the sequences A**CCGG**T, A**CGTG**T and A**GCGC**T was also observed. The *para*-isomer **2c** is the only derivative which is able to bind with the cognate A**CGCG**T at 1 µM. No significant footprint was observed at any sites examined for *meta*-(**2b**) and *ortho*-(**2a**) analogs up to 40 µM.

CONCLUSION

This study describes the synthesis of three phenyl substituted polyamide compounds (**2a-c**) with different degrees of curvature and probed their affinity to bind in the minor groove of DNA for comparison with the already synthesized compound f-ImPyIm (**3**). The data suggest that the *para*-derivative (**2c**) best meets the curvature requirements (of the compounds studies) to facilitate polyamide binding in the minor groove. This result is consistent with ΔT_m, CD, ITC, SPR and molecular modeling studies. The CD, ITC and SPR results also indicate selectivity of the *para*-derivative (**2c**) for its cognate sequence (A**CGCG**T) over the non-cognate sequences A**CGCG**T and A**AATT**T. In comparison

Part A



(Fig. (6). Contd….)

Part B

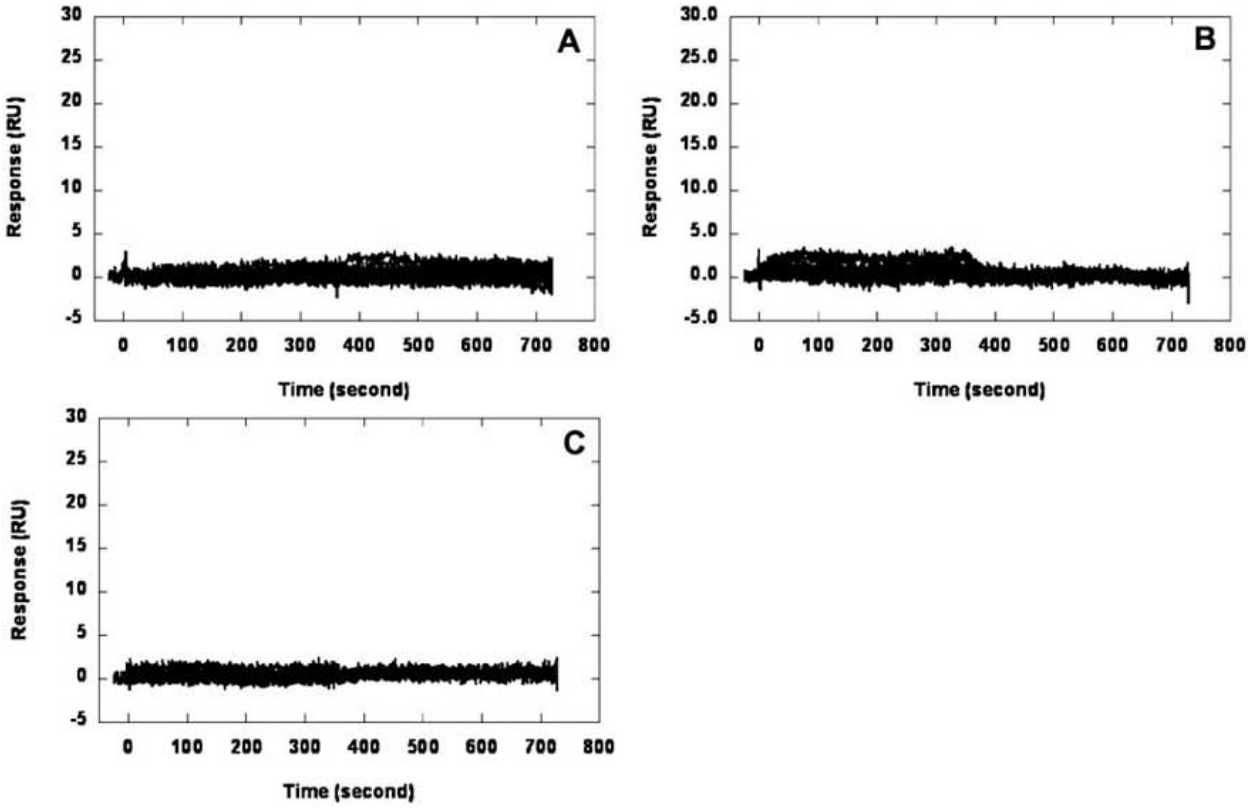


Fig. (6). SPR sensograms of (Part **A**) the *para*-compound (**2c**) with A**CGCG**T (**A**), A**AATT**T (**B**), A**CCGG**T (**C**) and the K plot (**D**) and (Part **B**) the *meta*-compound (**2b**) with A**CGCG**T (**A**), A**AATT**T (**B**), A**CCGG**T (**C**).

to the parent polyamide f-ImPyIm (**3**), the *para*- compound (**2c**) bound with a slightly less affinity and selectivity to the cognate DNA sequence; however, *para*-phenyl moieties are a useful alternative to the -ImPy- moiety. Molecular modeling studies showed that the *para*-compound (**2c**) was less curved and the *meta*-compound more curved than f-ImPyIm (**3**) potentially explaining the reduction in binding affinity. The results are confirmed by data obtained from DNase I experiments.

The results from this study indicate that the curvature of polyamides plays a role in binding affinity and selective, which will affect biological outcomes. Moreover, the *para*-compound or 4-aminobenzamide moiety presents a potential substitute for 4-aminopyrrole-2-carboxylate and it provides an opportunity for adjusting the curvature of polyamides to optimize binding affinity. Compounds **2c** and **3** are being investigated for their ability to inhibit the expression of the MCB transcriptional element and the results will be reported in sue course.

EXPERIMENTAL

General Method A

Compound **4** [7] (500 mg, 2.07 mmol) was dissolved in cold MeOH and reduced overnight *via* Parr hydrogenation (50 psi) using 5% Pd/C (250 mg) and reduced overnight. Upon completion the catalyst was removed *via* filtration over celite and washed with MeOH. The product was reacted with the required commercially available acid chloride (462 mg, 2.49 mmol) and allowed to react with TEA (0.35 mL, 2.49 mmol) and dry DCM (12 mL) in ice for 3 days. The product

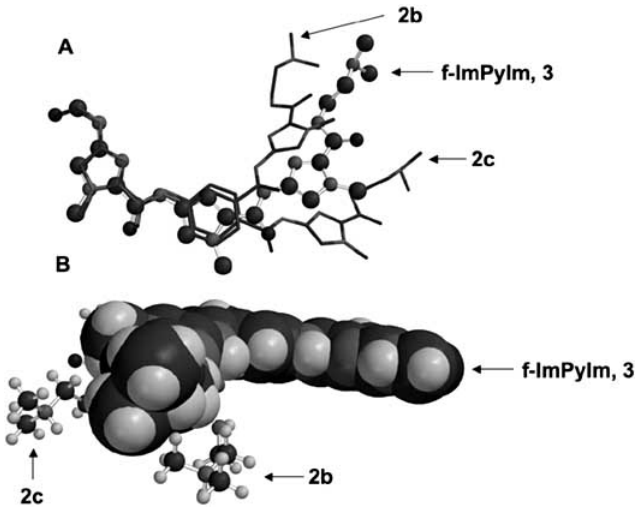


Fig. (7). Molecular modeling of f-ImPyIm (**3**) and the *meta*- and *para*-compounds (**2b** and **2c**, respectively).

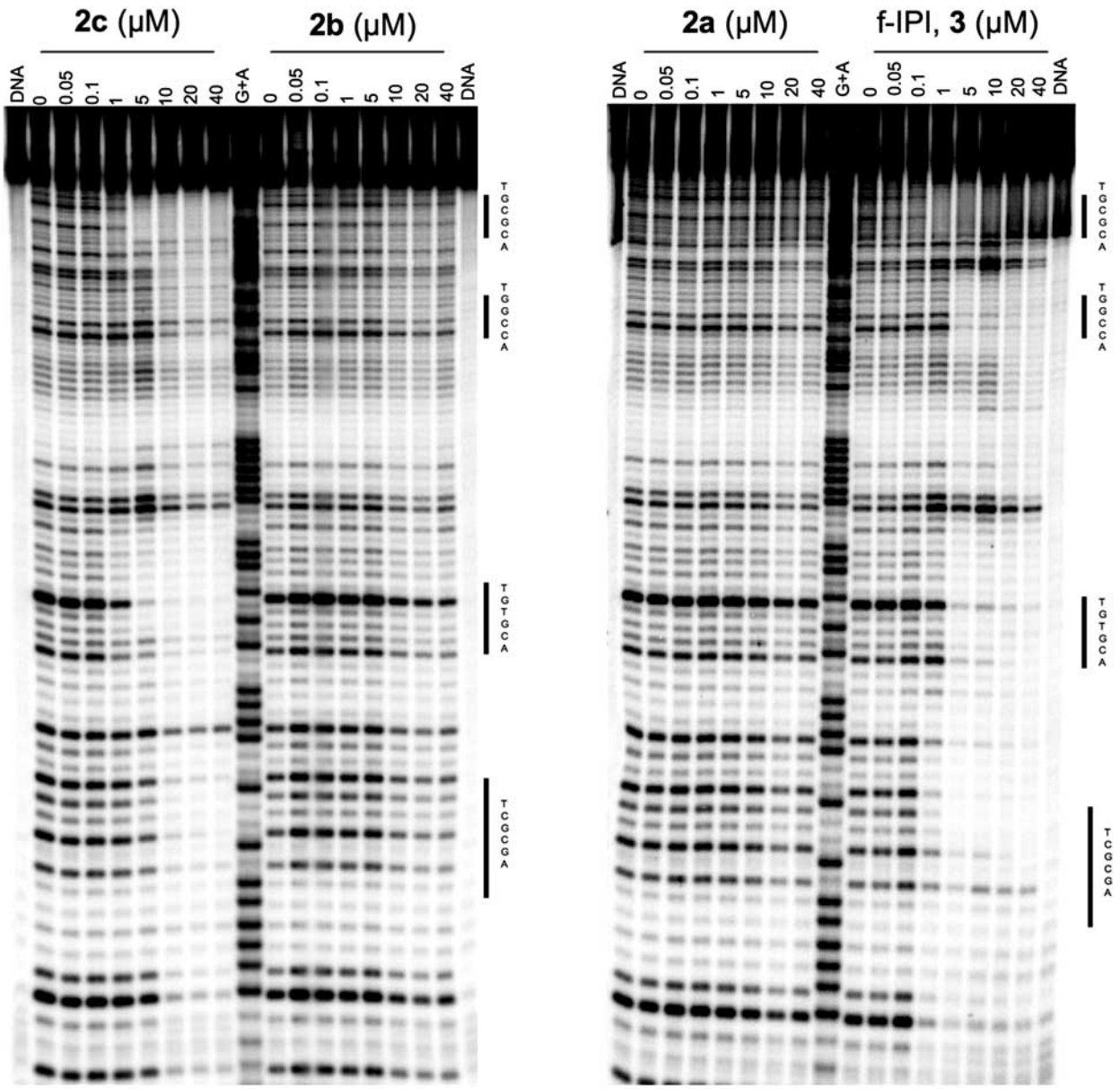


Fig. (8). DNase I footprinting. DNase I footprinting of f-ImPyIm (**3**) and *ortho*-, *meta*-, and *para*-derivatives **2a-c** on the antisense strand of the 5'-^{32}P-radiolabeled 131 bp DNA fragment. DNA denotes undigested DNA and G+A the purine sequencing lane. The sites 5'-A**CGCG**T-3', 5'-A**CCGG**T-3', 5'-A**CGTGT**-3' and 5'-A**GCGC**T-3' are indicated by solid bars.

compounds **5a-c** were extracted with 4 x 20 mL portions of DCM at pH 11 and dried over Na_2SO_4, filtered, and rotary evaporated to dryness. The filtered products were purified over silica gel *via* flash column chromatography using a gradient solvent of $CHCl_3$ and methanol. Starting with $CHCl_3$ the percentage of methanol was increased by 10% every 100 mL of solvent.

Compound 5a

Yellow solid (518 mg, 69%), mp. 58-66 °C: R_f = 0.30 (80:20% v/v $CHCl_3$/MeOH); IR (neat) 2947, 2861, 2824, 2779, 1667, 1568, 1529, 1471, 1349 cm^{-1}; ^{1}H-NMR ($CDCl_3$) 9.25 (br s), 8.11 (dd, *J* = 7.6, 1.6 Hz, 1H), 7.69 (dd, *J* = 7.6, 1.6 Hz, 1H), 7.64-7.58 (m, 2H), 7.48 (s, 1H), 4.02 (s, 3H), 3.44 (q, *J* = 6.0 Hz, 2H), 2.50 (t, J = 6.0 Hz, 2H) 2.17 (s, 6H); LRMS *m/z* 361 ([M+H], 100%) 197 (20%).

Compound 5b

Yellow solid (626 mg, 84%), mp. 55-60 °C: R_f = 0.43 (80:20% v/v $CHCl_3$/MeOH); IR (neat) 3089, 2948, 2862, 2823, 2778, 1667, 1565, 1531, 1471, 1349, 1258 cm^{-1}; ^{1}H-NMR ($CDCl_3$) 9.11 (br s, 1H), 8.77 (t, *J* = 2.0 Hz, 1H), 8.41 (dt, *J* = 1.2, 2.4, 8.0 Hz, 1H), 8.29 (d, *J* = 8.0 Hz, 1H), 7.71 (t, *J* = 8.0 Hz, 1H), 7.59 (brt, *J* = 5.2 Hz, 1H), 7.56 (s, 1H),

4.07 (s, 3H), 3.47 (q, *J* = 6.0 Hz, 2H), 2.54 (t, *J* = 6.0 Hz, 2H), 2.27 (s, 6H); LRMS *m/z* 361 ([M+H], 100%).

Compound 5c

Yellow solid (760 mg, 95%), mp. 207 °C: R_f = 0.33 (80:20% v/v $CHCl_3$/MeOH); IR (neat) 2947, 1667, 1602, 1525, 1346, 1259, 1108, 851 cm^{-1}; ^{1}H-NMR ($CDCl_3$) 8.62 (br s, 1H), 8.36 (d, *J* = 8.4 Hz, 2H), 8.08 (d, *J* = 8.4 Hz, 2H), 7.55 (s, 1H), 4.08 (s, 3H), 3.48 (q, *J* = 6.0 Hz, 2H), 2.53 (t, *J* = 6.0 Hz, 2H), 2.29 (s, 6H); LRMS *m/z* 361 ([M+H], 100%).

General Method B

The required intermediate (**5a-c)** were separately dissolved in cold MeOH and reduced overnight by catalytic hydrogenation with 5% Pd/C (50% by wt) at room temperature and atmospheric pressure. Compound **6** was reacted with the previously formed amino-derivatives in TEA (0.22 mL), dry DCM (12 mL) and was stirred in ice for three days. The product compounds **7a**-**b** were extracted with 4 x 20 mL portions of DCM at pH 11 and dried over Na_2SO_4, filtered, and rotary evaporated. The filtered products were purified over silica gel *via* flash column chromatography using a solvent system as described for compounds **5a-c**.

Compound 7a

Pale yellow solid (556 mg, 77%), mp. 155 °C: R_f = 0.50 (70:30% v/v $CHCl_3$/MeOH); IR (neat) 1667, 1510, 1470, 1451, 1435, 1309, 751 cm^{-1}; ^{1}H-NMR (DMSO-d_6) 12.05 (s, 1H), 11.17 (s, 1H), 8.67 (s, 1H), 8.47 (d, *J* = 8.0 Hz, 1H), 7.95 (d, *J* = 8.0 Hz, 1H), 7.76 (br t, *J* = 5.60 Hz, 1H), 7.60 (overlapping s and t, *J* = 8.0 Hz, 2H), 7.26 (t, *J* = 8.0 Hz, 1H), 4.06 (s, 3H), 4.00 (s, 3H), 3.35 (q, *J* = 6.0 Hz, 2H), 2.40 (t, *J* = 6.0 Hz, 2H), 2.17 (s, 6H); LRMS *m/z* 522 ([M+H], 15%) 506 ([M+Na], 20%) 484 ([M+H], 100%).

Compound 7b

Pale yellow solid (489 mg, 56%), mp. 119 °C: R_f = 0.38 (70:30% v/v $CHCl_3$/MeOH); IR (neat) 2956, 2927, 2851, 1661, 1538, 1490, 1470, 1311 cm^{-1}; ^{1}H-NMR ($CDCl_3$) 9.23 (br s, 1H), 8.45 (br s, 1H), 8.29 (d, *J* = 1.6 Hz, 1H), 7.89 (s, 1H), 7.81 (d, *J* = 8.4 Hz, 1H), 7.72 (d, *J* = 6.8 Hz, 1H), 7.55 (s, 1H), 7.52 (q, *J* = 8.4 Hz, 1H), 4.24 (s, 3H), 4.07 (s, 3H), 3.49 (q, *J* = 6.0 Hz, 2H), 2.54 (t, *J* = 6.0 Hz, 2H), 2.31 (s, 6H); LRMS *m/z* 484 ([M+H], 100%).

Compound 7c

Pale yellow solid (217 mg, 33%), mp. 147 °C: R_f = 0.35 (70:30% v/v $CHCl_3$/MeOH); IR (neat) 1658, 1528, 1471, 1379, 1310, 1245, 1186, 1110 cm^{-1}; ^{1}H-NMR ($CDCl_3$) 9.27 (s, 1H), 8.42 (s, 1H), 7.94 (d, *J* = 8.8 Hz, 2H), 7.89 (s, 1H), 7.82 (d, *J* = 8.8 Hz, 2H), 7.54 (s, 1H), 4.24 (s, 3H), 4.07 (s, 3H), 3.49 (q, *J* = 6.0 Hz, 2H), 2.53 (t, *J* = 6.0 Hz, 2H), 2.31 (s, 6H); LRMS *m/z* 484 ([M+H], 100%).

General Method C

The required compound (**7a**-**7c**) was dissolved in cold MeOH and reduced overnight by catalytic hydrogenation using 5% Pd/C (50% by wt) at room temperature and atmospheric pressure. Acetic anhydride (2 mL) was cooled in a 0 °C ice bath, where formic acid was added (1 mL). The solution was refluxed at 50 °C for 30 minutes with stirring, and cooled. The amine compounds were dissolved in dry DCM (8 mL) and cooled in a 0 °C ice bath. The acetic formic anhydride acid solution was then added drop wise to the amine and allowed to stir overnight. A basic extraction was performed, and the products extracted with 4 x 20 mL portions of $CHCl_3$ and dried over Na_2SO_4, filtered, and rotary evaporated. The products were purified on silica gel *via* flash column chromatography using a solvent system as described for compounds **5a-c** giving polyamides **2a**-**2c**.

Compound 2a

Orange-yellow solid (57 mg, 51%), mp. 97 °C: R_f = 0.40 (50:50% v/v $CHCl_3$/MeOH), $\varepsilon_{260}(H_2O)$ = 10729 $M^{-1}cm^{-1}$; IR (neat) 2956, 2923, 2852, 1668, 1652, 1567, 1558, 1515, 1456, 1367, 1259 cm^{-1}; ^{1}H-NMR (DMSO-d_6) δ 11.92 (s, 1H), 11.11 (s, 1H), 10.74 (s, 1H), 8.52 (d, *J* = 8.4 Hz, 1H), 8.21 (s, 1H), 7.91 (d, *J* = 7.6 Hz, 1H), 7.78 (t, *J* = 5.2 Hz, 1H), 7.61 (s, 1H), 7.19 (q, *J* = 7.6 Hz, 1H), 4.00 (s, 3H), 3.99 (s, 3H), 3.34 (q, *J* = 6.0 Hz, 2H), 2.40 (q, *J* = 6.0 Hz, 2H), 2.17 (s, 6H); LRMS *m/z* 504 ([M+Na], 20%) 482 ([M+H], 100%) 341 (50%); HRMS for $C_{22}H_{28}N_9O_4$ calcd. 482.2264 obsd 482.2277.

Compound 2b

Beige solid (102 mg, 73%), mp. 106 °C: R_f = 0.25 (50:50% v/v $CHCl_3$/MeOH), $\varepsilon_{260}(H_2O)$ = 23605 $M^{-1}cm^{-1}$; IR (neat) 3087, 2957, 2866, 2826, 2792, 1668, 1652, 1590, 1532, 1488, 1471, 1442, 755 cm^{-1}; ^{1}H-NMR ($CDCl_3$) δ 9.46 (br s, 1H), 9.39 (s, 1H), 9.10 (s, 1H), 8.38 (s, 1H), 8.03 (s, 1H), 7.95 (d, *J* = 8.0 Hz, 1H), 7.56 (br t, *J* = 5.2 Hz, 1H), 7.60 (d, *J* = 8.0 Hz, 1H), 4.02 (s, 3H), 3.78 (s, 3H), 3.52 (q, *J* = 6.0 Hz, 2H), 2.64 (t, *J* = 6.0 Hz, 2H), 2.32 (s, 6H); LRMS *m/z* 482 ([M+H], 100%); HRMS for $C_{22}H_{28}N_9O_4$ calcd. 482.2264 obsd 482.2263.

Compound 2c

Orange-brown solid (77 mg, 80%), mp. 107 °C: R_f = 0.30 (50:50% v/v $CHCl_3$/MeOH), $\varepsilon_{260}(H_2O)$ = 28835 $M^{-1}cm^{-1}$; IR (neat) 3291, 2925, 2855, 2825, 2779, 1659, 1567, 1510, 1470, 1365, 1320, 1244, 1187, 1110, 1054, 1017, 848, 757 cm^{-1}; ^{1}H-NMR ($CDCl_3$) δ 9.06 (s, 1H), 8.75 (br s, 1H), 8.63 (br s, 1H), 8.40 (s, 1H), 7.86 (d, *J* = 8.6 Hz, 1H), 7.71 (d, *J* = 8.6 Hz, 1H), 7.60 (br t, *J* = 5.0 Hz, 1H), 7.52 (s, 1H), 7.49 (s, 1H), 4.07 (s, 3H), 4.03 (s, 3H), 3.49 (q, *J* = 6.0 Hz, 2H), 2.53 (q, *J* = 6.0 Hz, 2H), 2.28 (s, 6H); LRMS *m/z* 482 ([M+H], 100%) 197 (30%); HRMS for $C_{22}H_{28}N_9O_4$ calcd. 482.2264 obsd 482.2262.

Thermal Denaturation (T_m)

Thermal denaturation data was obtained using a Cary 100 BioMelt (Varian) spectrophotometer with DNA (1 μM) in $PO_4$0 (10 mM phosphate buffer, containing 10 mM Na^+, and 1 mM EDTA, pH 6.2), and compounds **2a-c** (3 μM)**,** using the procedure previously reported [8,10b]. Thermal melts were obtained using both DNA's (A**CGCG**T and A**AATT**T) and each of the three polyamide compounds (**2a-c**).

Circular Dichroism (CD)

Circular Dichroism (CD) titration studies were performed on an OLIS DSM20 spectropolarimeter, using a 1 mm pathlength cuvette using the previously published procedure [10].

Experiments were carried out using $PO_4$5 (10 mM phosphate buffer, containing 50 mM Na^+, and 1 mM EDTA, pH 6.2), and 160 μL of a 9 μM DNA solution (**ACGCGT** and **AAATTT**) was titrated with 1 mol equivalents of the polyamide f-Im-Phenyl-Im **2a-c**, past the point of saturation. Each run was performed over 400-250 nm wavelength range (150 increments) using an integration time of 1 s, and the average of 2 scans were used for analysis.

Isothermal Titration Microcalorimetry (ITC)

ITC analysis was performed using a VP-ITC micromalorimeter (MicroCal) using the previously reported procedure [8,10b]. Compounds **2a-c** were dissolved in $PO_4$5 and the instrument equilibrated at 25 °C. An initial delay of 300 s was used, after which compounds **2a-c** (100 μM) was titrated, *via* 60 injections (3 μL for 7.2 s, repeated every 240 s), into 2 μM DNA ($PO_4$5). Origin 7.0 was used to analyze the data, where the integration of each titration was plotted as a function of time. A linear fit was then employed and this was subtracted from the reaction integrations to normalize for non-specific heat components. The two-sets-of-sites, non-sequential model using the MicroCal version of Origin 7.0 was used to determine the binding constant K_d as previously reported [14].

Surface Plasmon Resonance (SPR)

Biosensor chip surface preparations and biotinated DNA immobilizations were conducted as previously described [8a, 13]. Biotin labeled DNA hairpins (5'-biotin-GA**ACGCGT** CCTCTG**ACGCGT**TC-3', 5'-biotin-GA**ACCGGT**CCTCT G**ACCGGT**TC-3', and 5'-biotin-CG**AAATTT**CCTCTG **AAATTT**CG-3'). The experiments were conducted in a phosphate buffer at 200 mM Na^+ and 0.0005% P20 surfactant. In a typical experiment, 250 μL samples at different concentrations were injected onto the chip surface with a flow rate of 10 μL /min and 1500-sec dissociation period. The surface was regenerated with a glycine pH 2.5 solution and multiple buffer injections. Steady-state analyses were conducted and the response units were converted to mole of ligand per mole of compound as previously described [12]. The data were fitted with equation 1 to obtain macroscopic binding constants.

$$r = (K_1 \times C_{free} + 2 \times K_1 \times K_2 \times C_{free}^2) / (1 + K_1 \times C_{free} + K_1 \times K_2 \times C_{free}^2) \quad (1)$$

Molecular Modeling

The conformation of f-Im-Ph-Im (**3**) and isomers **2a-c** was examined by molecular modeling studies in the gas phase using the suite of programs in MacSpartan, version '04. Upon optimization of the structure using molecular mechanics (MMFF), the structure was energy optimized using Hartree-Fock (3-21G), followed by density function theory (B3LYP and 6-31G*) calculations.

DNase I Footprinting

Preparation of the DNA substrate, radiolabeling and purification. A 131 bp fragment was amplified using PCR. The forward primer 5'-GTCGTTAGGAGAGCTCACTTG-3' (4ng) was radioactively labeled by treatment with γ-[^{32}P] (3 μL) and 1 μL T4 polynucleotide kinase (in-vitrogen) following standard protocols. PCR was performed in thermophilic DNA polybuffer containing dNTPs (50 μL, 125 μM), $MgCl_2$ (1 mM), Flexi Taq polymerase (1 unit) and ^{32}P-labeled forward primer, reverse primer 5'-CTCCAGAAAGCCGGC AACTCAG-3' and the templates ATGCTCCAGAAAGCC GGCACTCAGTCTACAAACGCGTCATCTTGATCACCG GTGTTCACAGAAATTTCTCTAGATCTACACGTAACT CTAGTAGCGCTCTTCAAGCAAGTGGAGCTCTCCTAA CCGACTTT-3' (20 ng) and 5'-AAAGTCGGTTAGGAGA GCTCCACTTGCTTGAAGAGCGCTACTAGAGTTACGT GTAGATCTAGAGAAATTTCTGTGAACACCGGTGATC AAGATGACGCGTTTGTAGACTGAGTGCCGGCTTTCT GGAGCAT-3' (20 ng). Polymerase chain reaction was carried out as follows: an initial denaturation step for 3 min at 95 °C and [1 min at 94 °C, 1min at 63 °C and 1 min at 72 °C] for 35 cycles. The PCR products are purified by 2% agarose gel electrophoresis. Finally DNA was isolated by using the Mermaid Kit (Q-biogene) according the manufacturer's instructions.

DNase I digestions were conducted in a total volume of 8 μL. The labeled DNA fragment (200 μL, 200 count.sec^{-1}) was incubated 30 min in 4 μL TN binding buffer (10 mM Tris Base 10 mM NaCl, pH 7) containing the desired drug concentration. Cleavage by DNase I was initiated by addition of 2 μL DNase I solution (2 μL, 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, DNase I 0.1 unit.mL^{-1}, pH 8). After 3 min, the digestion was stopped by transferring the tube to dry ice and then the samples were lyophilized. DNA was resuspended in 4 μL formamide loading dye (95% formamide, 20 mM EDTA, 005% bromophenol blue and 0.05% cyanol blue) and denatured by heating the sample to 90 °C and cooled on ice prior to loading onto a conventional denaturing polyacrylamide (10%) gel containing urea (7.5 M). Electrophoresis was performed for 2.5 hours (70W, 50 °C) in TBE Buffer (89 mM Tri base, 89 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3). The gel was transferred onto whatman 3MM paper and dried under vacuum at 80 °C for 45 min. The gel was exposed overnight with a phosphorimager screen and picture was acquired using a phosphorimager scanner (STORM 840).

ACKNOWLEDGEMENTS

The authors thank NSF (CHE-0550992), Medical Research Council UK (G0000168) and Cancer Research UK (C2259/A3083) for support.

REFERENCES

[1] (a) Pelton, J.G.; Wemmer, D.E. Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci.,* **1989**, *86*, 5723-7; (b) Pelton, J.G.; Wemmer, D.E. Binding modes of distamycin A with d(CGCAAAT TTGCG)$_2$ determined by two-dimensional NMR. *J. Am. Chem. Soc*., **1990**, *112*, 1393-99.

[2] (a) Dervan, P.B.; Edelson, B.S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol*., **2003**, *13*, 284-99; (b) Kopka, M.L.; Goodsell, D.S.; Hun, G.W.; Chin, T.K.; Lown, J.W.; Dickerson, R.E. Defining GC-specificity in the minor groove: side-by-side binding of the di-imidazole lexitropsin to C-A-T-G-G-C-C-A-T-G. *Structure,* **1997**, *5*, 1033-46; (c) Yang, X.L.; Hubbara, R.B.; Lee, M.; Tao, Z.F.; Sugiyama, H.; Wang, A.H.J. Imidazole-imidazole pair as a minor groove recognition motif for T:G mismatched base pairs. *Nucl. Acids Res.,* **1999**, *27*, 4187-90.

[3] (a) Dervan, P.B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem*., **2001**, *9*, 2215-35; (b) Gottesfeld, J.M. Small molecules affecting transcription in Friedreich ataxia. *Pharmacol. Ther.,* **2007**, *116*, 236-48; (c) Melander, C.; Burnett, R.; Gottesfeld, J.M. Regulation of gene expression with pyrrole-imidazole polyamides. *J. Biotechnol.,* **2004**, *112*, 195-220; (d) Dervan, P.B.; Poulin-Kerstien, A.T.; Fechter, E.J.; Edelson, B.S. Regulation of gene expression by synthetic DNA-binding ligands. *Top. Curr. Chem.,* **2005**, *253*, 1-31; (e) Matsuda, H.; Fukuda, N.; Ueno, T.; Tahira, Y.; Ayame, H.; Zhang, W.; Bando, T.; Sugiyama, H.; Saito, S.; Matsumoto, K.; Mugishima, H.; Serie, K. Development of gene silencing pyrrole-imidazole polyamide targeting the TGF-beta1 promoter for treatment of progressive renal diseases. *J. Am. Soc. Nephrol*., **2006**, *17*, 422-32.

[4] (a) Hochhauser, D.; Kotecha, M.; O'hare, C.; Morris, P.J.; Hartley, J.M.; Taherbhai, Z.; Harris, D.; Forni, C.; Mantovani, R.; Lee, M.; Hartley, J.A. Modulation of topoisomerase IIα expression by a DNA sequence-specific polyamide. *Mol. Cancer Ther*., **2007**, *6*, 346-54; (b) Le, N.M.; Sielaff, A.M.; Cooper, A.J.; Mackay, H.; Brown, T.; Kotecha, M.; O'Hare, C.; Hochhauser, D.; Lee, M.; Hartley, J.A. Binding of f-PIP, a pyrrole- and imidazole-containing triamide, to the inverted CCAAT box-2 of the topoisomerase IIα promoter and modulation of gene expression in cells. *Bioorg. Med. Chem. Lett.*, **2006**, *16*, 6161-4.

[5] (a) Nguyen, B.; Tardy, C.; Bailly, C.; Colson, P.; Houssier, C.; Kumar, A.; Boykin, D.W.; Wilson, W.D. Influence of compound structure on affinity, sequence selectivity, and mode of binding to DNA for unfused aromatic dications related to furamidine. *Biopolymers*, ***2002,*** *63,* 281-97; (b) Nguyen, B.; Hamelberg, D.; Bailly, C.; Colson, P.; Stanek, J.; Brun, R.; Neidle, S.; Wilson, W.D. Characterization of a novel DNA minor-groove complex. *Biophys. J*., **2004***, 86,* 1028-41.

[6] (a) Ngyuen, B.; Lee, M.P.H.; Hamelberg, D.; Joubert, A.; Bailly, C.; Brun, R.; Neidle, S.; Wilson, W.D. Strong binding in the DNA minor groove by an aromatic diamidine with a shape that does not match the curvature of the groove. *J. Amer. Chem. Soc.,* **2002**, *124*, 13680-1; (b) Miao, Y.; Lee, M.P.H.; Parkinson, G.N.; Batista-Parra, A.; Ismail, M.A.; Neidle, S.; Boykin, D.W.; Wilson, W. D. Out-of-shape DNA minor groove binders: induced fit interactions of heterocyclic dications with the DNA minor groove. *Biochemistry,* **2005**, *44*, 14701-8; (c) Ismail, M.A.; Batista-Parra, A.; Miao, Y.; Wilson, W.D.; Wenzler, T.; Brun, R.; Boykin, D.W. Dicationic near-linear biphenyl benzimidazole derivatives as DNA-targeted antiprotozoal agents. *Bioorg. Med. Chem.,* **2005**, *13*, 6718-26.

[7] (a) Fairley, T.A.; Tidwell, R.R.; Donkor, I.; Naiman, N.A.; Ohemeng, K.A.; Lombardy, R.J.; Bentley, J.A.; Cory, M. Structure, DNA minor groove binding, and base pair specificity of alkyl- and aryl-linked bis(amidinobenzimidazoles) and bis(amidinoindoles). *J. Med. Chem.,* **1993**, *36*, 1746-53; (b) Cory, M.; Tidwell, R.R.; Fairley, T.A. Structure and DNA binding activity of analogues of 1,5-bis(4-amidinophenoxy)pentane (pentamidine). *J. Med. Chem*., **1992**, *35*, 431-8; (c) Tidwell, R.R.; Jones, S.K.; Geratz, J.D.; Ohemeng, K.A.; Cory, M.; Hall, J.E. Analogues of 1,5-bis(4-amidinophenoxy)pentane (pentamidine) in the treatment of experimental *Pneumocystis carinii* pneumonia. *J. Med. Chem*., **1990**, *33*, 1252-7.

[8] (a) 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. Physical and structural basis for the strong interactions of the -ImPy- central pairing motif in the polyamide f-ImPyIm. *Biochemistry,* **2006**, *45*, 13551-65; (b) Buchmueller, K.L.; Staples, A.M.; Howard, C.M.; Horick, S.M.; Uthe, P.B.; Le, N.M.; Cox, K.K.; Nguyen, B.; Pacheco, K.A.; Wilson, W.D.; Lee, M. Extending the language of DNA molecular recognition by polyamides: unexpected influence of imidazole and pyrrole arrangement on binding affinity and specificity. *J. Am. Chem. Soc.,* **2005**, *127*, 742-50.

[9] (a) Verma, R.; Patapoutian, A.; Gordon, C. B.; Campbell, J. L. Identification and purification of a factor that binds to the Mlu I cell cycle box of yeast DNA replication genes. *Proc. Natl. Acad. Sci. USA,* **1991**, *88*, 7155-9; (b) Wu, X.; Lee, H. Human Dbf4/ASK promoter is activated through the Sp1 and MluI cell-cycle box (MCB) transcription elements. *Oncogene,* **2002**, *21*, 7786-96; (c) Yamada, M.; Sato, N.; Taniyama, C.; Ohtani, K.; Arai, K.; Masai, H. A 63-base pair DNA segment containing an Sp1 site but not a canonical E2F site can confer growth-dependent and E2F-mediated transcriptional stimulation of the human ASK gene encoding the regulatory subunit for human Cdc7-related kinase. *J. Biol. Chem*., **2002**, *277*, 27668-81; (d) Hess, G. F.; Drong, R. F.; Weiland, K. L.; Slightom, J. L.; Sclafani, R. A.; Hollingsworth, R. E. A human homolog of the yeast CDC7 gene is overexpressed in some tumors and transformed cell lines. *Gene,* **1998**, *28*, 133-40.

[10] (a) Lee, M.; Rhodes, A.L.; Wyatt, M.D.; Forrow, S.; Hartley, J.A. GC base sequence recognition by oligo(imidazolecarboxamide) and C-terminus-modified analogues of distamycin deduced from circular dichroism, proton nuclear magnetic resonance, and methidium-propylethylenediaminetetraacetate-iron(II) footprinting studies. *Biochemistry,* **1993**, *32*, 4237-45; (b) Brown, T.; Mackay, H.; Turlington, M.; Sutterfield, A.; Smith, T.; Sielaff, A.; Westrate, L.; Bruce, C.; Kluza, J.; O'Hare, C.; Hartley, J.A.; Nguyen, B.; Wilson, W.D.; Lee, M. Modifying the N-terminus of polyamides: PyIm-PyIm has improved sequence specificity over f-ImPyIm. *Bioorg. Med. Chem*., **2008**, *16*, 5266-76.

[11] March, J. *In Advanced Organic Chemistry: Reactions, Mechanisms and Structure*, 4th ed., Wiley-Interscience, John Wiley & Sons, **1992**, p. 392.

[12] (a) Lacy, E.R.; Le, N.M.; Price, C.A.; Lee, M.; Wilson, W.D. Influence of a terminal formamido group on the sequence recognition of DNA by polyamides. *J. Am. Chem. Soc*., **2002**, *124*, 2153-63; (b) Lyng, R.; Rodger, A.; Norden, B. The CD of ligand-DNA systems. 2. Poly(dA-dT) B-DNA. *Biopolymers,* **1992**, *32*, 1201-14; (c) Lyng, R.; Rodger, A.; Norden, B. The CD of ligand-DNA systems. I. Poly(dG-dC) B-DNA. *Biopolymers,* **1991**, *31*, 1709.

[13] Tanious, F.A.; Nguyen, B.; Wilson, W.D. Biosensor-surface plasmon resonance methods for quantitative analysis of biomolecular interactions. *Methods Cell Biol.,* **2008**, *84*, 53-77.

[14] Indyk, L.; Fisher, H.F. Theoretical aspects of isothermal titration calorimetry. *Methods Enzymol.,* **1998**, *295*, 350-64.

Received: 22 December, 2008 Revised: 24 February, 2009 Accepted: 25 February, 2009