

# NOVEL PICOLINIC ACID-CONTAINING PYRROLE-IMIDAZOLE POLYAMIDES: SYNTHESIS AND T•G MISMATCHED BASE PAIR RECOGNITION

Peter B. Uthe, Andrew M. Staples, Mark Turlington, Justin B. Jones, Kevin N. Blackmon, Suzanna L. Bailey, Karen L. Buchmueller, Moses Lee\*

Moses.Lee@furman.edu

Department of Chemistry, Furman University, Greenville, SC 29613

**Abstract:** Pyrrole (Py) – Imidazole (Im) polyamides bind as stacked side-by-side dimers in the minor groove of DNA with strong affinity and specificity. When two imidazole-containing polyamides are stacked, the overlapping imidazole units bind tightly to T•G or G•T mismatched base pairs. In our design of polyamide analogs capable of discriminating a T•G from G•T site, two picolinic acid-containing polyamides have been synthesized and the base pair preference of the stacked picolinate/imidazole pairing was studied. The compounds contain a picolinic acid (Pic) or a 6-methylpicolinic acid (MePic) group at the N-terminus, and their structures are PicPyIm and MePicPyIm. Results from circular dichroism (CD) studies showed that the stacked Pic/Im and MePic/Im pairing bound to deoxynucleotide including T•G mismatched base pairs more strongly over the corresponding G•T and Watson-Crick G•C base pairs.

## Introduction

v

One of the most exciting discoveries from the human genome project is the recognition that the genome is a dynamic structure, and its composition is continuously changing. One mechanism that leads to such changes is single nucleotide polymorphism (SNP),<sup>1</sup> in which mismatched DNA base pairs, including T•G mismatches, play an important role in their formation.<sup>2</sup> SNPs are being investigated for many applications, including population genetics and pharmacogenomics.<sup>3</sup> The understanding of T•G mismatched base pairs in DNA are also important because they are responsible for most of the common mutations leading to formation of tumors in humans. For example, in human bladder carcinoma, a G•C to A•T transition at the 3'-G of the GG doublet in codon 12 converts the Ha-*ras* and Ki-*ras* proto-oncogenes into oncogenes.<sup>4</sup> Because of the biological importance of the T•G mismatch, its structural and physical properties have been extensively investigated. NMR analysis of duplex 5'-CGTGAATTCGCG-3', containing symmetrical T•G mismatches, indicates that the mismatches adopt a wobble conformation and structural perturbations are mainly localized in the vicinity of the mismatch and the nearest

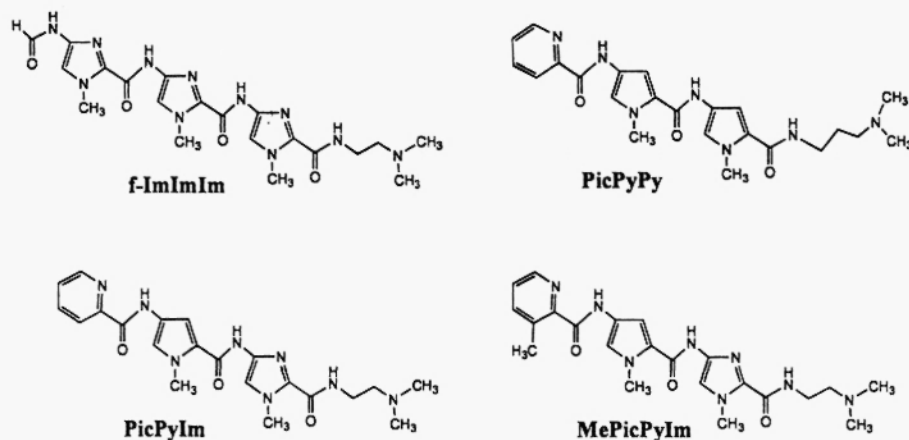


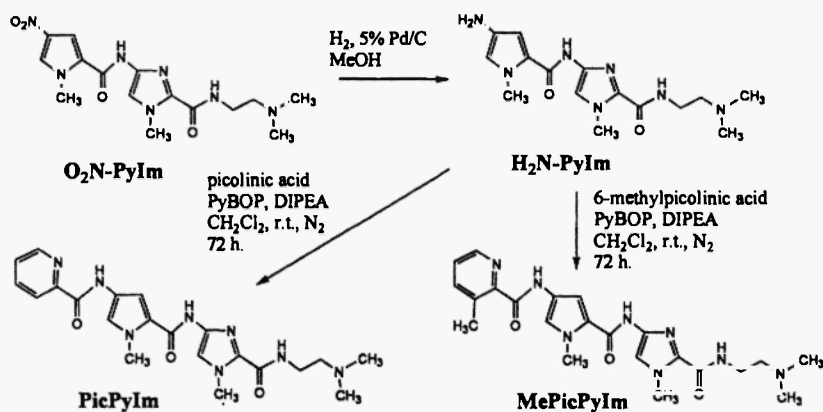
Figure 1. Structures of polyamides, f-ImImIm, PicPyPy, PicPyIm, and MePicPyIm.

neighbor.<sup>5</sup> These localized perturbations in T•G mismatch-containing oligonucleotides are supported by an x-ray analysis of 5'-CGCGAATTTGCG-3'.<sup>6</sup> Based on these structural subtleties, our group discovered that stacked imidazole-containing dimers of polyamides were able to recognize T•G or G•T sites with unexpected affinity.<sup>7</sup> According to the NMR structure of a stacked f-ImImIm (Figure 1) dimer bound to a T•G-containing oligonucleotide, Lee and collaborators proposed that the Im/Im pairing could be a good motif for recognition of a T•G mismatch.<sup>7a</sup> The free amino group projects into the minor groove and it can form two separate hydrogen bonds to a side-by-side Im/Im pairing. This molecular recognition motif was further characterized by a combination of molecular biology (DNase I footprinting)<sup>7b</sup> and biophysical analyses (surface plasmon resonance,<sup>7b</sup> circular dichroism,<sup>7a</sup> and isothermal titration calorimetry).<sup>7c</sup> The unique recognition of T•G mismatched base pairs by the Im/Im pairing was corroborated in a study reported by the Dervan group, in which one of their hairpin polyamide compounds showed some selectivity for a T•G site over a G•T base pair.<sup>8</sup> However, from our studies on f-ImImIm the preference for either a T•G or G•T site was not apparent, which we had attributed to the symmetrical nature of the Im/Im pairing.

As part of our program aimed at designing compounds capable of discriminating a T•G from a G•T site, we have synthesized two picolinic acid-containing polyamide analogs, PicPyIm and MePicPyIm (Figure 1). The picolinic acid (Pic) or 6-methylpicolinic acid (MePic) group was attached to the N-terminus of the polyamide, such that their corresponding side-by-side stacked dimer would provide the necessary unsymmetrical Pic/Im or MePic/Im pairing. These pairings should not only mimic the recognition of T•G or G•T mismatched sites of the Im/Im pairing, but also to discriminate between them. The Pic moiety was chosen as an analog of imidazole because the Dervan group had shown the stacked dimer of PicPyPy (Figure 1) to recognize GC containing sequences, indicating that the Pic/Py pairing behaved in a similar manner to Im/Py.<sup>9</sup> However, PicPyPy retained a strong memory for AT-rich sequences. Dervan's group suggested that the dual recognition of PicPyPy was due to rotation about the bond between the pyridyl and carboxamide moieties. We have consequently designed polyamide-containing analogs that bear a 6-methylpicolinic acid (MePic) group at the N-terminus where the 6-methyl group should restrict the rotation about the pyridyl-carboxamide bond, much like a methylimidazole unit, forcing the N-atom in the pyridyl group into the DNA interacting concave face of the molecule.

## Results and Discussion

The target compounds, MePicPyIm and PicPyIm, were synthesized using a strategy depicted in Scheme 1. Reduction of O<sub>2</sub>N-PyIm<sup>10</sup> by catalytic hydrogenation with 5% Pd/C provided H<sub>2</sub>N-PyIm, an unstable compound, which was used immediately in the subsequent step. A solution of H<sub>2</sub>N-PyIm in CH<sub>2</sub>Cl<sub>2</sub> was coupled to picolinic acid or 6-methylpicolinic acid in the presence of benzotriazole-1-yloxy-trisphosphonium hexafluorophosphate (PyBOP) and N,N-diisopropylethylamine (DIPEA). The reaction mixtures were stirred under a N<sub>2</sub> atmosphere and at room temperature for 72 hours. MePicPyIm and PicPyIm were isolated as white solids, after purification on a MeOH/CHCl<sub>3</sub> silica gel column, and in yields of 70% and 86%, respectively. The compounds were characterized by proton NMR, infrared spectroscopy, mass spectrometry, and accurate mass measurements.



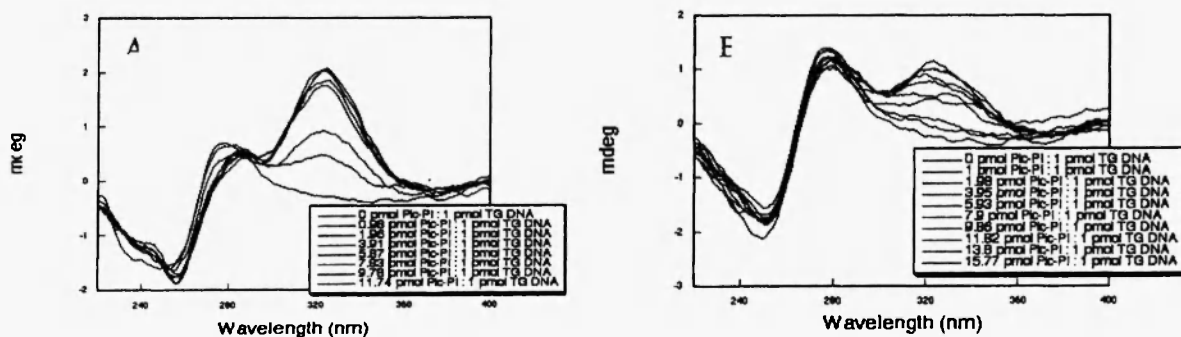
The ability of PicPyIm and MePicPyIm to interact with Watson-Crick and T•G-containing mismatched oligonucleotide sequences was examined by thermal melts ( $T_M$ ) and by circular dichroism (CD) studies. The oligonucleotides used in these studies are given in Figure 2.

G•C, 5'-CGT AGA CTA CGG TTTT CCG TAG TCT ACG-3'  
 C•G, 5'-CGT ACA GTA CGG TTTT CCG TAC TGT ACG-3'  
 A•T, 5'-CGC AAA TTC CGG TTTT CCG GAA TTT GCG-3'  
 T•A, 5'-CGC ATA ATC CGG TTTT CCG GAT TAT GCG-3'  
 G•T, 5'-CGT AGA TTA CGG TTTT CCG TAG TTT ACG-3'  
 T•G, 5'-CGT ATA GTA CGG TTTT CCG TAT TGT ACG-3'

Figure 2: Oligonucleotides used in the DNA binding studies.

Using a 3:1 mole ratio of compound to DNA, and a sodium phosphate buffer (10 mM sodium phosphate, 1 mM EDTA, pH 6.2), both compounds produced negligible increases in the melting temperature for all the DNAs depicted in Figure 2. However, using a CD titration experiment, the mole ratio of compound to DNA was gradually raised from 0 to 11. Consistent with the DNA melt study, at a mole ratio of 3 the DNA induced ligand band at 320 nm was weak, less than a mdeg. However, when the mole ratio was raised further ( $> 5$ ) for the T•G and G•T oligonucleotides, a clear positive Cotton effect was seen, indicating that the compounds were bound to the minor groove.<sup>11</sup> As depicted in Figure 3A, titration of MePicPyIm to the T•G containing DNA produced an isodichroic point and a larger DNA induced band (about 2 mdeg) when compared to the titration of MePicPyIm to the G•T containing DNA (less than 1 mdeg, Figure 3B). These results suggest that the MePic/Im pairing formed in the stacked dimer demonstrates some preference for a T•G mismatched pair over G•T.

Figure 3: A. CD Studies on the titration of MePicPyIm to T•G-containing DNA. B. Titration of MePicPyIm to G•T-containing DNA.



Moreover, according to CD studies, MePicPyIm did not bind to the AT sequences, and binding to the corresponding GC sequences were similar to that observed for the G•T sequence, thereby supporting the preference of MePic/Im pairing for T•G sites. A similar T•G preference was seen for PicPyIm. In conclusion, we have discovered a novel DNA binding motif for the recognition of T•G mismatched DNA base pairs.

### Acknowledgements

The authors acknowledge support from the NSF-REU, NIH (RR16461-10) from the BRIN Program, and Furman University.

### References

1. (a) Brennan, M. : *Chem & Eng. News* July 3, **78**, 4 (2000). (b) Borman, S.: *Chem & Eng. News* May 29, **78**, 14 (2000), and references given therein.
2. Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E. et al. (The international SNP map working group) *Nature*, **409**, 928 (2001).
3. (a) Evans, W.E., Relling, M.V. : *Science*, **286**, 487 (1999). (b) Liggett, S.B. : *Nature Med.*, **7**, 281 (2001).
4. (a) Land, H., Parada, L.F., Weinberg, R.A. : *Science*, **222**, 771 (1983). (b) Almoguera, C., Shibata, D., Forrester, K., Martin, J., Arnheim, N., Perucho, M. : *Cell*, **53**, 549 (1988). (c) Watanabe, H., Ha, A., Hu, Y.X., Ohtsubo, K., Yamaguchi, Y., Motoo, Y., Okai, T., Toya, D., Tanaka, N., Sawabu, N. : *Cancer*, **86**, 1441 (1999).
5. Hare, D., Shapiro, L., Patel, D.J. : *Biochemistry*, **25**, 7445 (1986).
6. Hunter, W.N., Brown, T., Kneale, G., Anand, N.N., Rabinovitch, D., Kennard, O. : *J. Biol. Chem.*, **262**, 9962 (1987).
7. (a) Yang, X-L., Hubbard IV, R.B., Lee, M., Tao, Z-H., Sugiyama, H., Wang, A.H-J. : *Nucl. Acids Res.*, **27**, 4183 (1999). (b) Lacy, E.R., Cox, K.K., Wilson, W.D., Lee, M. : *Nucl. Acids Res.*, **30**, 1834 (2002). (c) Lacy, E.R., Nguyen, B., Le, N., Cox, K.K., Lee, M., O'Hare, C., Hartley, J A., Wilson, W.D. : *Nucl. Acids Res.*, **32**, 2000 (2003).
8. Rucker, V.C., Foister, S., Melander, C. and Dervan, P.B. : *J. Am. Chem. Soc.*, **125**, 1195 (2003).
9. Wade, W. S., Mrksich, M., Dervan, Peter B. : *J. Am. Chem. Soc.*, **114**, 8783 (1992).
10. Lacy, E.R., Le, N.M., Price, C.A., Lee, M., Wilson, W.D. : *J. Am. Chem. Soc.*, **124**, 2153 (2002).
11. Rodger, L.R., Norden, B. : *Biopolymers*, **32**, 1201 (1992), and references given therein.

**Received on October 10, 2004**