Pergamon

Bioorganic & Medicinal Chemistry Letters 12 (2002) 2245–2248

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Sequence Selective Recognition of DNA by Hairpin Conjugates of a Racemic *seco*-Cyclopropaneindoline-2-benzofurancarboxamide and Polyamides

James L. Toth,[a] Carly A. Price,[a] Erik C. Madsen,[a] Heather L. Handl,[a] Stephen J. Hudson,[b] Richard B. Hubbard, III,[a,c] J. Phillip Bowen,[c] Konstantinos Kiakos,[d] John A. Hartley[d] and Moses Lee[a,*]

[a]*Department of Chemistry, Furman University, Greenville, SC 29613, USA*
[b]*Department of Biology, Furman University, Greenville, SC 29613, USA*
[c]*Department of Chemistry, University of Georgia, Athens, GA 30602, USA*
[d]*Department of Oncology, Royal Free University College Medical School, London W1W 7BS, UK*

Received 15 January 2002; accepted 27 March 2002

This paper is dedicated to Professor J. William Lown on the occasion of his retirement

Abstract—Conjugates of racemic *seco*-cyclopropaneindoline-2-benzofurancarboxamide (CI-Bf) and four diamides (ImIm **1**, ImPy **2**, PyIm **3**, and PyPy **4**, where Py is pyrrole, and Im is imidazole), linked by a γ-aminobutyrate group were synthesized. In addition to alkylating at adenine-N3 positions within an A_5 sequence, the imidazole-containing compounds **1** and **2** were found to also alkylate purine-N3 positions within a sequence 3′-GGGGGGA(888)CTGCTC(894)-5′. A model for the binding of hairpin conjugates **1** and **2** with the 3′-GACT-5′ sequence is proposed.

The human genome project, which has sequenced the genome of 3 billion base pairs, has led to the identification of many new genes and their associated control sequences.[1] Small molecules that are tailored to bind to predetermined DNA sequences, can permeate cell membranes, interfere with DNA replication and transcription are potentially powerful tools for molecular biology[2,3] as well as diagnosis and treatment of diseases at the DNA level.[2,4]

The DNA minor groove is an attractive target for the design and development of sequence-specific ligands for biological activity.[5] This is due partly to the observation that covalent adducts in the minor groove are generally poorly repaired compared to those in the major groove.[6] The minor groove binding agent distamycin,[7] a polyamide, and its imidazole-containing analogues have been extensively studied for their ability to bind to specific DNA sequences in a 2:1 or side-by-side binding motif. A set of rules for their sequence specific recognition have been reported.[8–10] In addition, polyamides that fold into a hairpin conformation thereby creating an 'internal side-by-side' binding motif have been shown to be able to inhibit the expression of the 5S RNA genes by RNA polymerase III[11] and the Ets-1, LEF-1 and TBP in the HIV-1 enhancer/promoter element in cells.[12] Even though the non-covalently reactive polyamide hairpin compounds were biologically active, the question arises whether a covalently reactive hairpin would be endowed with improved biological properties. In this regard, a heterodimer of duocarmycin A and distamycin was found to alkylate within the minor groove of DNA in a side-by-side manner.[13] The dimer bound preferentially to a sequence different from either one of the compounds alone. While both distamycin and duocarmycin A bind strictly to AT-rich sequences, the dimer alkylates GN3 position in GC-rich sequences, such as 5′-AGGTG-3′. More recently, we demonstrated that imidazole-containing triamides, such as compound **5**, dramatically modulated the sequence-specificity of duocarmycin A alkylating guanine-N3 atoms at GC sequences.[14] These findings indicate that hairpin conjugates of polyamides and duocarmycin A might provide an attractive direction for the discovery of a new

*Corresponding author. Tel.: +1-864-294-3368; fax: +1-864-294-3559; e-mail: moses.lee@furman.edu

0960-894X/02/$ - see front matter
PII: S0960-894X(02)00341-4

class of gene specific 'knockout' agents. To date, several conjugates of the duocarmycin alkylating pharmacophore and its analogues with polyamides, including hairpin polyamides, have been reported.[8c,15] However, the sequence specific alkylation exhibited by these conjugates were directed by the imidazole/pyrrole pairings and side-by-side binding motif of the polyamides.[8–10] Conjugates of CC-1065 analogues with oligodeoxynucleotides have also been reported.[16] In this paper, we report the synthesis and sequence selectivity of a series of conjugates of racemic *seco*-cyclopropaneindoline-2-benzofurancarboxamide (CI-Bf) with four different polyamides (ImIm **1**, PyIm **2**, ImPy **3**, and PyPy **4**), in which the DNA sequence recognition of the side-by-side CI-Bf/polyamide pairing will be investigated (Fig. 1).

The synthetic strategy for preparation of the target hairpin molecules **1**–**4** is depicted in Scheme 1. Using published methods,[17] the diheterocyclic polyamides **6**–**9** were readily prepared from *N*-butoxycarbonyl (BOC) protected diaminobutane with the acid chlorides of 1-methyl-4-nitroimidazole-2-carboxylic acid or 1-methyl-4-nitropyrrole-2-carboxylic acid.[18] Reduction of the nitro group of compounds **6**–**9** gave the corresponding amines, which were reacted with glutaric anhydride in the presence of *N*,*N*-dimethylaminopyridine (DMAP) in pyridine. The corresponding acids **10**–**13** were obtained in 11–98% yields after silica gel column chromatography. Synthesis of the racemic *seco*-cyclopropaneindoline-2-benzofurancarboxamido (CI-Bf) moiety, a duocarmycin analogue, started from dichloride **14** that was prepared from the corresponding diol[19] in quantitative yield with triphenylphosphine and CCl_4. Using a method that was recently reported from our laboratory, the dichloride was reduced with stannous chloride in hydrochloric acid, which after work up with sodium hydroxide gave an indoline intermediate.[20] Coupling of the indoline with 5-nitrobenzofuran-2-carboxylic acid in the presence EDCI in DMF gave the desired amide **15** in 71% yield. Selective reduction of the nitro moiety in compound **15** was readily achieved using Adams catalyst and hydrogenation at a pressure of 55 PSI. Subsequently, reaction of the amine with carboxylic acids **10**–**13** gave the protected hairpin compounds. The phenolic moieties in compounds **17**–**20** were unmasked by hydrogenation over 10% palladium on carbon, and the products were obtained in 30–78% yields after silica gel column chromatography. The BOC protecting group was removed using 3 M hydrochloric acid in dry ethyl acetate. The desired off-white precipitates of products **1**–**4** were collected, washed with dry ethyl acetate, and dried in vacuo at room temperature. All the compounds prepared in this study were characterized by 500 MHz ^{1}H NMR, FT-IR, FAB mass spectrometry and accurate mass measurements.

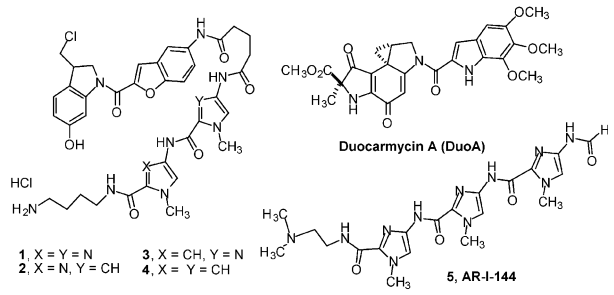
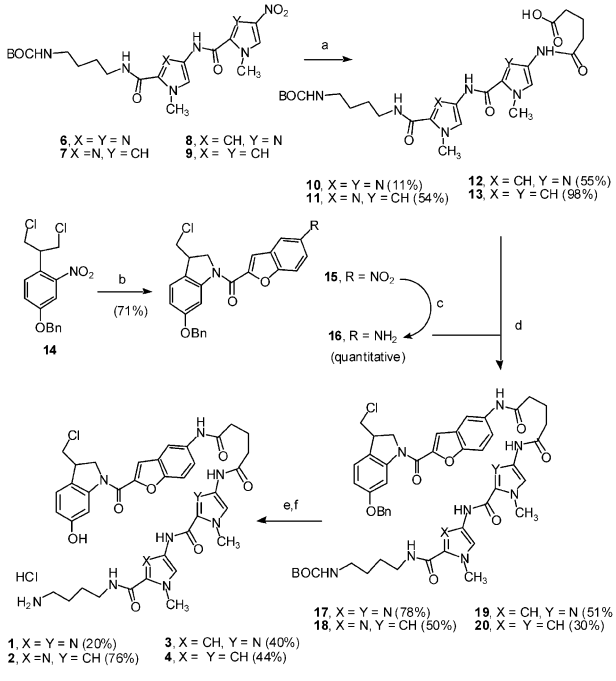


Figure 1. Structures of the target compounds **1**–**4**, duocarmycin A (DuoA), and AR-1-144.



Scheme 1. Synthesis of the target compounds **1**–**4**: (a) (i) hydrogen, 5% Pd/C, MeOH; (ii) glutaric anhydride, DMAP, pyridine; (b) (i) $SnCl_2$, HCl; (ii) NaOH; (iii) 5-nitrobenzofuran-2-carboxylic acid, EDCI, DMAP, DMF; (c) hydrogen, PtO_2, 55PSI; (d) polyamide acid, EDCI, HOBt, DMF; (e) hydrogen, 10% Pd/C, THF; (f) 3 M HCl, EtOAc.

The covalent sequence specificity of compounds **1**–**4** was assessed by a thermally induced DNA strand cleavage experiment, which is commonly used to probe sequence specific covalent interactions with purine-N3 in the minor groove.[21] The DNA fragment used in these studies was obtained from PCR amplification of base pairs 741–940 of the pUC18 plasmid that was linearized with *Hind III*. A 5′-^{32}P labeled 5′-CTGTCGGGTTT-3′ primer was used as the forward primer so that each final probe copy was singly end-labeled. Results from the thermally induced DNA strand break experiment, as depicted in Figure 2, indicate that compounds **1**–**4** retain a memory for alkylating adenine-N3 within an A_5 sequence, a site that is preferred by CC-1065, adozelesin, and *seco*-CI-trimethoxyindole (or *seco*-CI-TMI) (data not shown). Interestingly, minor, but unique, alkylation bands at 3′-GGGGGGA(888)CTGCTC(894)-5′ (alkylation at the purine-N3 of the underlined bases) were also be observed for compounds **1** and **2**, indicating their ability to recognize a GC-containing sequence. To do so, the *seco*-CI-Bf and polyamide subunits fold to form a hairpin structure, and the resulting pairing of an imidazole unit side-by-side with the benzofuran moiety then interacts with a GC base pair. These results are consistent with our earlier findings that imidazole-containing polyamides are capable of altering

the sequence selectivity of duocarmycin A from AT to GC-rich sequences.

In a proposed model of the hairpin complex of compound (*S*)-**1** with the 3′-GGG*A(888)CT*GCTC-5′ sequence, given in Figure 3, the imidazole would stack with the benzofuran to form a complex similar to the pyrrole/imidazole pairing that is capable of recognizing a C/G base pair.[8,9] It is not likely that compounds **1** and **2** would bind to this sequence in an extended conformation,

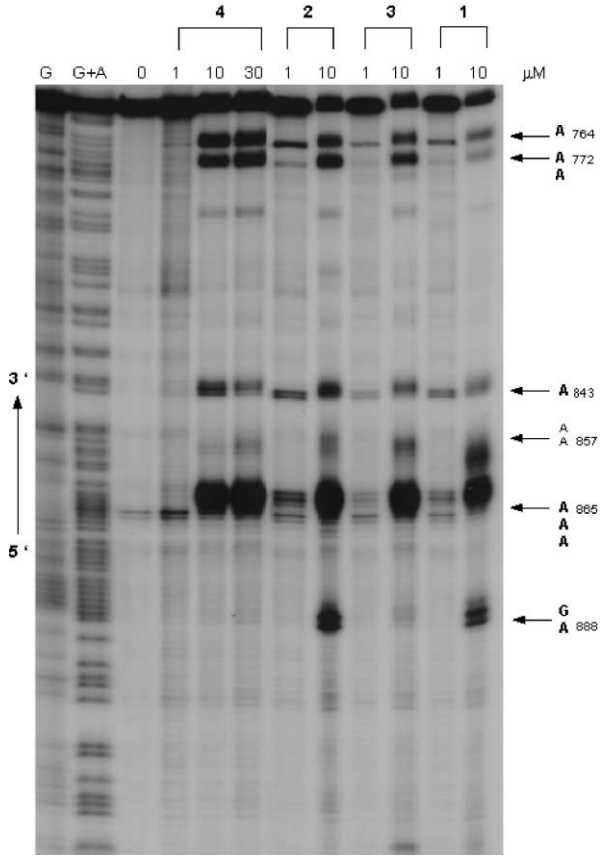


Figure 2. Thermally induced and sequence selective DNA strand break study on compounds **1**–**4**.

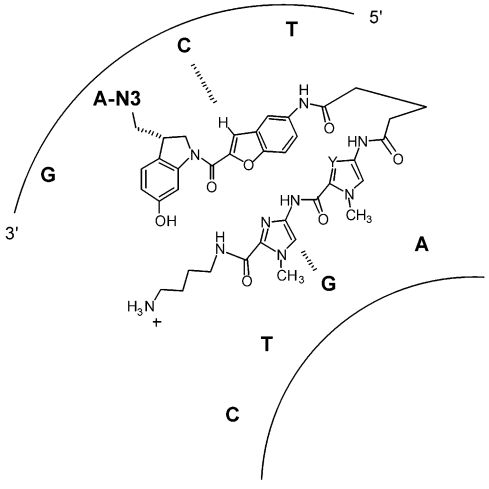


Figure 3. Model of molecular interactions between the natural (*S*)-enantiomers of compounds **1** and **2** with 3′-GACT-5′.

because in that case the pyrrole moiety in compound **2** would have to be unfavorably placed next to a G(891)/C base pair.[22,23] In addition, an extended conformation of compounds **1** and **2** would require the benzofuran group to bind unfavorably to a C(889)/G base pair. We have found that like *seco*-CI-TMI, *seco*-CI-Bf also bound to the A(865) cluster and not to GC-containing sites (data not shown). Further support of the hairpin conformation of compounds **1** and **2** came from a report that an aliphatic dicarboxamide of polyamides, such as a bis-linked netropsin with pimelic acid, was capable of forming a hairpin structure.[24] Structural analyses of the binding of compound **1** using NMR and molecular modeling studies are on-going and the results will be reported in due course.

The cytotoxicity of the hairpin compounds **1**–**4** was determined against L1210 (murine leukemia) and P815 (murine mastocytoma) cell lines using a MTT assay.[22] Following a 3-day continuous exposure of compounds **1**–**4**, the experimental IC_{50} values against the growth of L1210 cells were determined to be 6.8, 6.7, 45, and 28 μM, respectively. It is worthy to note that compounds **1** and **2** were more cytotoxic to the L1210 cells than compounds **3** and **4**, and that might be related to their different DNA sequence selectivity profiles. For P815 cells, the IC_{50} values are 52, 9.1, 77, and 41 μM, respectively.

In summary, the studies described in this communication show that conjugates of duocarmycins and polyamides that can fold into a hairpin conformation are capable of recognizing specific sequences of DNA. They provide a feasible platform for the design of sequence specific probes for therapeutic and diagnostic applications. Studies aimed at improving the sequence specificity and increasing the DNA binding site size are currently underway in the investigators' laboratories and the results will be reported in due course.

Acknowledgements

Financial support from Research Corporation, NSF-REU (to M.L.), the Arnold and Mabel Arnold Beckman Foundation (to M.L. and E.C.M.), and the Cancer Research Campaign (to J.A.H.) is gratefully acknowledged. The authors also thank Dr. Robert Kelly at Pharmacia & Upjohn for a generous gift of CC-1065.

References and Notes

1. (a) Brennan, M. *Chem. Eng. News* **2000**, *78*, 4. (b) Borman, S. *Chem. Eng. News* **2000**, *78*, 14. (c) Venter, J. C. et al. *Science* **2001**, *291*, 1304. (d) Lander, E. S., et al. *Nature* **2001**, *409*, 860.
2. (a) Neidle, S. *DNA Structure and Recognition*; IRL: Oxford, 1994. (b) Neidle, S., Waring, M., Ed. *Molecular Aspects of Anticancer Drug-DNA Interaction*; CRC: Boca Raton, 1993; Vol. 1. (c) *Molecular Aspects of Anticancer Drug-DNA Interaction*; Neidle, S., Waring, M., Eds.; CRC: Boca Raton, 1994; Vol. 2. (d) Propst, C. L., Perun, T. J., Eds., *Nucleic Acid Targeted Drug Design*; Marcel Dekker: New

York, 1992. (e) Hurley, L. H., Ed. *Advances in DNA Sequence-Specific Agents*; JAI: Greenwich, 1992; Vol. 1. (f) Jones, G. B., Palumbo, M., Eds. *Advances in DNA Sequence-Specific Agents*; JAI: Greenwich, 1998; Vol. 3. (g) Neidle, S.; Thurston, D. E. In *New Molecular Targets for Cancer Chemotherapy*; Kerr, D. J., Workman, P., Eds.; CRC: Boca Raton, 1994; p. 159.
3. Dervan, P. B. *Science* **1986**, *232*, 464.
4. (a) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *Chem. Eur. J.* **1997**, *3*, 1600. (b) Hurley, L. H. *J. Med. Chem.* **1989**, *32*, 2027.
5. (a) Lown, J. W. *Anti-Cancer Drug Des.* **1988**, *3*, 25. (b) Krowicki, K.; Lee, M.; Hartley, J. A.; Ward, B.; Kissinger, K.; Skorobogaty, A.; Dabrowiak, J. C.; Lown, J. W. *Struct. Expression* **1988**, *2*, 251. (c) Helene, C. *Nature* **1998**, *391*, 436. (d) Weisz, K. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2592. (e) Nielsen, P. E. *Chem. Eur. J.* **1997**, *3*, 505.
6. (a) Brooks, N.; McHugh, P. J.; Lee, M.; Hartley, J. A. *Chem. Biol.* **2000**, *7*, 659. (b) Smellie, M.; Kelland, L. R.; Thurston, D. E.; Souhami, R. L.; Hartley, J. A. *Br. J. Cancer* **1994**, *70*, 48.
7. (a) Wemmer, D. E. *Biopolymers* **2001**, *52*, 197. (b) Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439.
8. (a) Dervan, P. B.; Burlii, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688. (b) Wemmer, D. E.; Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997**, *7*, 355. (c) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
9. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 194.
10. Kopka, M. L.; Goodsell, D. S.; Han, G. W.; Chiu, T. K.; Lown, J. W.; Dickerson, R. E. *Structure* **1997**, *5*, 1033.
11. Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202.
12. Dickinson, L. A.; Gilizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12890.
13. Sugiyama, H.; Lian, C.; Isomura, M.; Saito, S.; Wang, A. H.-J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14405.
14. Fujiwara, T.; Tao, Z.-H.; Ozeki, Y.; Saito, I.; Wang, A. H.-J.; Lee, M.; Sugiyama, H. *J. Am. Chem. Soc.* **1999**, *121*, 7706.
15. (a) Chang, A. Y.; Dervan, P. B. *J. Am. Chem. Soc.* **2000**, *122*, 4856. (b) Tao, Z.-F.; Fujiwara, T.; Saito, I. *J. Am. Chem. Soc.* **1999**, *121*, 4961. (c) Tao, Z.-F.; Saito, I.; Sugiyama, H. *J. Am. Chem. Soc.* **2000**, *122*, 1602. (d) Tao, Z.-F.; Fujiwara, T.; Saito, I.; Sugiyama, H. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 650.
16. Kutyavin, I. V.; Afonina, I. A.; Mills, A.; Gorn, V. V.; Lukhtanov, E. A.; Belousov, E. S.; Singer, M. J.; Walburge, D. K.; Lokhov, S. G.; Gall, A. A.; Dempcy, R.; Reed, M. W.; Meyer, R. B.; Hedgpeth, J. *Nucl. Acids Res.* **2000**, *28*, 655.
17. Brooks, N.; Hartley, J. A.; Simpson, J. E., Jr.; Wright, S. R.; Woo, S.; Centioni, S.; Fontaine, M. D.; McIntyre, T. E.; Lee, M. *Bioorg. Med. Chem.* **1997**, *5*, 1497.
18. Lown, J. W.; Krowicki, K. *J. Org. Chem.* **1985**, *50*, 3774.
19. Warpehoski, M. A.; Gebhard, I.; Kelly, R. C.; Krueger, W. C.; Li, L. J.; McGovern, J. P.; Prairie, M. D.; Wicnienski, N.; Wierenga, W. *J. Med. Chem.* **1988**, *31*, 590.
20. Jennings, S. A.; Toth, J. L.; Roller, S. G.; Brooks, N.; Kiakos, K.; Hartley, J. A.; Burke, P. J.; Lee, M. *Heterocyl. Commun.* **2001**, *7*, 7.
21. (a) Boger, D. L.; Johnson, D. S. *Angew. Chem., Int. Ed.* **1996**, *35*, 1438. (b) Hartley, J. A.; Wyatt, M. D. In *Drug-DNA Interaction Protocols*; Fox, K. R., Ed.; Humana: Totowa, NJ, 1997; p 147.
22. Lee, M.; Rhodes, A. L.; Wyatt, M. D.; D'Incalci, M.; Forrow, S.; Hartley, J. A. *J. Med. Chem.* **1993**, *36*, 863.
23. Urbach, A. R.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4343.
24. Surovaya, A. N.; Burckhardt, G.; Birch-Hirschfeld, E. B.; Nikitin, A. M.; Frizsche, H.; Zimmer, C.; Gursky, G. V. *J. Biomol. Struct. Dyn.* **2001**, *18*, 689.