Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 53–61

DNA sequence-specific recognition of peptides incorporating the HPRK and polyamide motifs

Jung-Cheng Chang,[a] Chia-Hung Yang,[a] Ping-Yen Chou,[a] Wan-Hsu Yang,[a] I.-Chun Chou,[a] Ching-Tai Lu,[a] Pei-Hsuan Lin,[a] Rolis Chien-Wei Hou,[b] Kee-Ching G. Jeng,[b] Chien-Chung Cheng[c] and Leung Sheh[a,*]

[a]*Department of Chemistry, Tunghai Christian University, Taichung, Taiwan 407, ROC*
[b]*Department of Medical Research, Taichung Veterans General Hospital, Taiwan 405, ROC*
[c]*Department of Chemistry, Tamkang University, Taichung, Taiwan 251, ROC*

Received 3 September 2003; accepted 20 October 2003

Abstract—Three peptide amides, HPRK(Py)$_4$HPRK-NH$_2$ (PyH-12), HPRK(Py)$_3$HPRK-NH$_2$ (PyH-11) and HPRK(Py)$_2$HPRK-NH$_2$ (PyH-10), incorporating two HPRK motifs and various 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) were synthesized by solid-phase peptide methodology. The binding of these three peptides to a 5′-^{32}P-labeled 158-mer DNA duplex (Watson fragment) and to a 5′-^{32}P-labeled 135-mer DNA duplex (complementary Crick fragment) was investigated by quantitative DNase I footprinting. On the 158-mer Watson strand, the most distinctive DNase I blockages seen with all three peptides occur around positions 105–112 and 76–79, corresponding to the sequences 5′-GAGAAAAT-3′ and 5′-CGGT-3′, respectively. However, on the complementary Crick strand, only PyH-12 strongly discriminates the 5′-TTT-3′ site around positions 108–110 whereas both PyH-11 and PyH-10 have moderate binding around positions 102–112 comprising the sequence 5′-ATTTTCTCCTT-3′. Possible bidentate and single interactions of the side-chain functions and α-amino protons of the peptides with DNA bases are discussed.

1. Introduction

The studies of synthetic DNA binding and cleavage agents are considered to be of growing importance in the research of gene regulation. DNA sequence-specific binding agents which chemists studied extensively over the past three decades can be roughly divided into two categories: naturally occurring antibiotics such as quinoxaline antibiotics,[1–4] anthraquinones,[5–7] netropsin and distamycin;[8–11] and synthetic polymers[12–15] related to netropsin and distamycin which bind specifically to the DNA minor groove. The advances in the understanding of sequence-specific recognition and mechanisms of synthetic drugs would be expected to contribute significantly in future gene therapy of many diseases. For example, sequence-specific binding drugs can be targeted to bind to various sites of specific genes and switch off the expression of functional proteins, thus aid in the control of malignant diseases.

On the other hand, synthetic DNA cleavage agents are mostly peptide conjugates of drugs, metal ions or chromophores that can cleave DNA and have been extensively investigated.[16–21] Recently, we have been studying peptides which can nick plasmid DNA without conjugation to drugs or metal ions.[22] As part of our ongoing interest in the search for new DNA binding and cleavage agents, we designed and synthesized three polyamide–peptide conjugates incorporating 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) with the HPRK motif (C-terminal in amide form): HPRK(Py)$_4$HPRK-NH$_2$ (PyH-12), HPRK(Py)$_3$HPRK-NH$_2$ (PyH-11), and HPRK(Py)$_2$HPRK-NH$_2$ (PyH-10). These peptide amides are designed by the rationale that both the XPRK motif[23] and the polyamide motif[8–15] are considered to bind in sequence-specific or sequence-selective manners to the DNA minor groove. We reasoned that the conjugation of the XPRK motif with the polyamide motif in a peptide might afford new DNA sequence-specific agents. The design of the new XPRK motif is prompted by the observation of Suzuki that a SPXX motif[24–27] is often found in repeating sequences in histones, steroid hormone receptors, various segmentation gene products and some oncogene products. The

Keywords: HPRK motif; Polyamide conjugates; Quantitative DNase I footprinting.
* Correspondence author. Tel.: +886-4-2359-0248; fax: +886-4-2359-0426; e-mail: lsheh@mail.thu.edu.tw

0968-0896/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2003.10.049

SPXX motif assumes a β-turn stabilized by two hydrogen bonds, and the side chains of the two basic residues engage in salt bridges with the DNA phosphate groups.[24–26] In this work, all three new peptides displayed significant agarose gel electrophoretic retardation against supercoiled pBR322 DNA, indicating that tight binding of the peptides to DNA had occurred.

Agarose gel electrophoresis showed that PyH-12 induced significant nicking of pBR322 plasmid DNA at peptide concentrations as low as 2 μM at 65 °C, whereas $(HPRK)_3$-NH_2 showed much less DNA nicking effect. The 10-mer PyH-10, the 11-mer PyH-11 also produced much less nicking effect whereas $YPRK(Py)_4YPRK$-NH_2 had no detectable DNA nicking effect and no agarose gel electrophoretic retardation effect under identical conditions (data and figures not shown). To gain insight into the recognition of PyH-12, PyH-11 and PyH-10 to specific binding sites of the DNA double helix quantitative DNase I footprinting was carried out on a 158-mer DNA fragment and on a complementary 135-mer DNA fragment.

2. Materials and methods

2.1. Chemicals

All of the protected amino acid derivatives were purchased from Bachem California (Torrance, CA, USA) and AnaSpec, Inc. (San Jose, CA, USA) or synthesized in our own laboratory according to published procedures. All other analytical reagents were purchased from Acros, Tedia or Sigma.

2.2. Chemical methods

Melting points were determined on a Mel-Temp apparatus (Cambridge, MA, USA) and are uncorrected. Optical rotations were determined on a Rudolph Autopol II instrument. Semi-preparative and analytical HPLC (Vydac reversed-phase columns, TP201; column 1, 1×25 cm; column 2, 0.4×25 cm) were performed using a Hitachi L-7100 pump equipped with a gradient elution device and a Soma S-3702 variable wavelength UV detector which is connected to a PC computer installed with Hitachi HPLC analytical software. Mass spectra were determined with a Finnigan/Thermo Quest MAT 95XL instrument operating in the electrospray ionization (ESI) mode in Chung-Hsing University.

2.3. Biochemicals

Radiolabeled nucleoside triphosphates [γ-^{32}P]ATP were obtained from NEN Life Science Products at a specific activity of 6000 Ci/mmol. *Taq* polymerase, T4 polynucleotide kinase, and DNase I were purchased from Promega. All of the biochemicals were used according to the supplier's recommended protocol in the activity buffer provided. All other chemicals were analytical grade reagents, and all solutions were prepared using doubly deionized, Millipore filtered water.

2.3.1. His-Pro-Arg-Lys-Py-Py-Py-Py-His-Pro-Arg-Lys-NH_2 (PyH-12). This peptide was synthesized using solid-phase methodology by manual operation of a Protein Technology PS3 peptide synthesizer. The first Fmoc-protected amino acid was coupled to the Nova Rink amide AM resin using benzotriazole-1-yloxy-trispyrrolidinophosphoniumhexafluorophosphonate/*N*-methylmorpholine (PyBOP/NMM) in DMF. All of the N_α-Fmoc-protected amino acids (in 4 equivalent ratio excess to the resin) were coupled in a stepwise fashion using PyBOP/NMM in DMF after deprotection of the *N*-Fmoc group by piperidine. The side chains of Arg, Lys, Tyr and His are protected by the pentamethylchroman-6-sulfonyl (Pmc), Boc and trityl (Trt) groups, respectively. After coupling the last N-terminal Fmoc-amino acid residue, the resin was treated with the cleavage reagent (0.75 g phenol, 10 mL TFA, 0.5mL thioanisole, 0.25 mL EDT) for 1.5 h, and then lyophilized. The resin was washed with dry ether (2×30 mL), filtered, and then washed with 5% acetic acid (200 mL). The combined filtrate was lyophilized and the product was purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution. Eluent A: 5% MeCN, 95% H_2O, 0.1% TFA; Eluent B, 95% MeCN, 5% H_2O, 0.1% TFA. A linear gradient was achieved by increasing the MeCN content from eluent A to eluent B in 30 min. R_t (column 2), 11.77 min, mp 157–158 °C, $[\alpha]_D^{28}$ −41.3 (*c* 0.0444, H_2O); ESIMS requires: 1542.76, found: 1542.

2.3.2. His-Pro-Arg-Lys-Py-Py-Py-His-Pro-Arg-Lys-NH_2 (PyH-11). This peptide was synthesized using a similar procedure as for PyH-12. The crude product was purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution as described for the purification of PyH-12. R_t (column 2), 11.41 min, mp 152–153 °C, $[\alpha]_D^{28}$ −75.8 (*c* 0.0528, H_2O); ESIMS requires: 1420.63, found: 1421.9.

2.3.3. His-Pro-Arg-Lys-Py-Py-His-Pro-Arg-Lys-NH_2 (PyH-10). This peptide was synthesized using a similar procedure as for PyH-12. The crude product was purified by semi-preparative reversed-phase HPLC (column 1) using gradient elution as described for the purification of PyH-12. R_t (column 2), 10.033 min, mp 149–152 °C, $[\alpha]_D^{28}$ −76.0 (*c* 0.051, H_2O); ESIMS requires: 1298.51, found: 1300.

2.4. Polymerase chain reaction (PCR) and end-labeling of PCR products

pBR322 plasmid was subjected to PCR amplification in a thermal cycler (Perkin-Elmer DNA thermal cycler 480). Briefly, for the preparation of the 135-mer Crick fragment, the primers 5′-ACGTAGCGATAGCGGAGTG (Watson/pBR322 sequencing primer) and 5′-AGCGGAAGAGCGCCTGATG (Crick/pBR322 sequencing primer) bearing a 5′-OH were used and obtained from ScinoPharm Biotech Ltd. The 5′-OH Crick primer was labeled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase. The Watson primer corresponds to pBR322 nucleotide sequence of 2226–2244 whereas the Crick primer corresponds to the sequence of 2360–2342.

250 μM of each appropriate dNTP and 1 unit of Taq polymerase were mixed in a volume of 100 μL containing 50 mM KCl, 10 mM Tris–HCl, pH 8.3, 0.1% Triton X-100, and 1.5 mM $MgCl_2$. The reaction mixture was heated to 60 °C before the addition of Taq polymerase and mineral oil was added to prevent evaporation. Following an initial step of 3 min at 94 °C, 30 amplification cycles were performed, each consists of 94 °C denaturation for 50 s, 55 °C primer annealing for 50 s, and polymerization at 72 °C for 50 s. After the last amplification cycle, the PCR samples were heated to 72 °C for another 10 min to complete unfinished polynucleotide chains. The reaction mixtures were loaded to PCR columns (QIAGEN QIAquick PCR purification kit), washed with PB buffer (500 μL), centrifuged for 60 s and the lower layer discarded. 0.75 mL of PE buffer was added and the mixture centrifuged for 60 s and the lower layer discarded. The column was centrifuged for another min and the lower layer again discarded. The column was adopted in an Eppendorf tube and stood with 50 μL of deionized water for 1 min and then centrifuged for 1 min. The purified DNA fragment was diluted to 50 cps/μL. In a similar procedure, the 158-mer Watson fragment was prepared using the Watson primer 5′-ATTGTACTGAGAGTGCACC-3′ (labeled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase) and the Crick primer: 5′-ATTACCGCCTTTGAGTGAG-3′.

2.5. DNase I footprinting

Reactions were conducted in a total volume of 10 μL. Radiolabeled DNA (2 μL) was mixed with varying concentrations of peptide (2 μL) dissolved in 5 mM sodium cacodylate buffer, pH 6.5) and equilibrated at room temperature for 60 min. DNase I (2 μL) was added and the reaction allowed to proceed at 37 °C for 10 min. The DNase I solution (in 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$) was adjusted to yield a final concentration of 0.009 unit/mL so as to limit the digestion to less than 30% of the starting material in order to minimize the incidence of multiple cleavages in any strand. The digestion was stopped by adding stop solution (4 μL) containing 80% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol. Samples were heated at 90 °C for 4 min and chilled in ice for 4 min prior to electrophoresis.

2.6. Electrophoresis and autoradiography

The products of DNA cleavage were resolved by polyacrylamide gel electrophoresis under denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea). G + A and G tracks were generated by treatment of the radiolabeled DNA with formic acid and dimethyl sulfate, respectively. Briefly, to 15 μL of the untreated DNA were added 15 μL of water and 30 μL of formic acid or 2 μL of dimethyl sulfate. For the formic acid reaction, the reaction mixture was incubated at 20 °C for 10 min. For the dimethyl sulfate reaction, the reaction was allowed to proceed at room temperature for 20 s. The reaction was quenched by adding unlabeled DNA (2 μL, 3 μg/μL) 3 M sodium acetate (6 μL), 2-mercaptoethanol (5 μL, for DMS reaction only) and 800 μL of 95% ethanol at −20 °C. The mixture was centrifuged at 12,000*g* for 15 min at 4 °C and the supernatant discarded. The residue was washed twice with ethanol, mixed with 50 μL of 0.7 M piperidine and the reaction was allowed to proceed at 90 °C for 30 min. The mixture was dried under vacuum and loaded with tracking dyes to the gel. The various samples after DNase I treatment are also loaded to the gel in a total volume of 10 μL/well. After electrophoresis (about 1.45 h at 70 W, 1800 V in TBE buffer, BRL sequencer model S2), gels were soaked in 10% acetic acid/10% methanol for 15 min, transferred to Whatman 3MM paper, dried under vacuum at 80 °C for 45 min, and placed against a Molecular Dynamics PhosphorImager storage screen overnight. The electrophoretic band areas were analyzed by a PC computer installed with Viber Lourmat BIO-ID software (Marne La Valle, Cedex 1, France). Footprinting data are presented in the form $\ln(f_a/f_c)$ representing the differential cleavage at each site relative to that in the control (f_a is the fractional cleavage at any band in the presence of the peptide and f_c is the fractional cleavage in the control, i.e., absence of peptide). The results are displayed on a natural logarithmic scale. Positive values indicate enhanced cleavage whereas negative values indicate blockage. The C_{50} values of the peptides were determined using the method of Sayers and Waring.[28]

3. Results and discussion

Initial DNA nicking experiments were performed by a standard procedure[22] of incubating pBR322 plasmid with the peptides at 65 °C for 25 min. Figure 1 shows that the 12-mer PyH-12 effectively nicked the plasmid at concentration as low as 2 μM whereas PyH-11, PyH-10 and $(HPRK)_3$-NH_2 showed much less nicking ability at similar concentrations.

At present, we are unable to locate the cleavage sites in the pBR322 plasmid induced by these peptides. However, the unusual DNA nicking property of these pep-

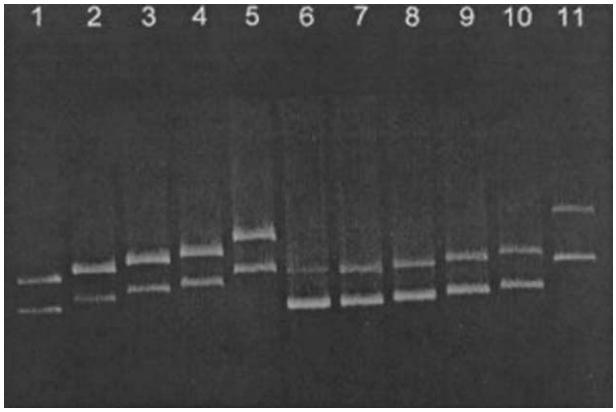

Figure 1. Nicking of supercoiled pBR322 DNA induced by PyH-12 and $(HPRK)_3$-NH_2 as monitored by agarose electrophoresis. The DNA was incubated with each peptide at 65 °C for 25 min. The upper bands represent the open-circular form, and the lower bands the supercoiled form. Lane 6, pBR322 DNA as received. Lanes 1–5, DNA incubated with PyH-12 at 0.5, 2, 5, 10 and 25 μM, respectively. Lanes 7–11, DNA incubated with $(HPRK)_3$-NH_2 at 0.5, 2, 5, 10 and 25 μM, respectively.

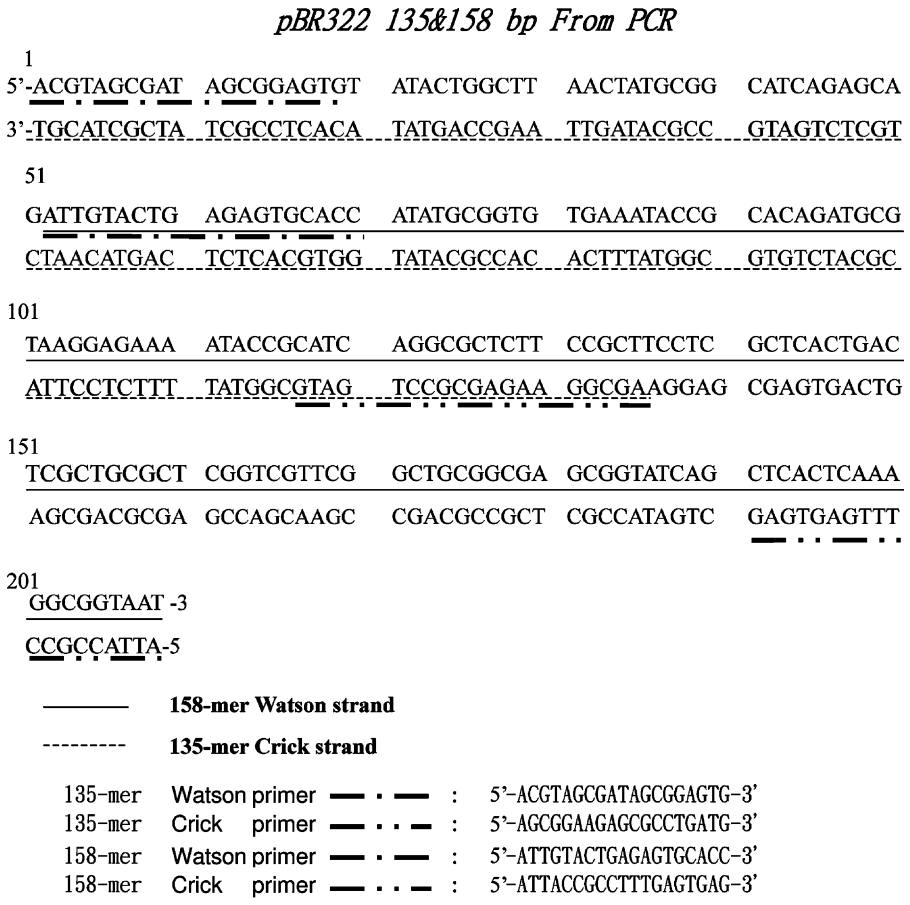

Figure 2. Nucleotide sequence of the 5′-^{32}P-labeled 158-mer DNA duplex (the corresponding 5′-^{32}P-labeled strand is referred to as the Watson strand) and a complementary 5′-^{32}P-labeled 135-mer DNA duplex (the corresponding 5′-^{32}P-labeled strand is referred to as the Crick strand). The sequences of the primers for PCR are shown.

tides has prompted us to investigate their aspects of DNA sequence-specific recognition. Thus, quantitative DNase I footprinting experiments were designed to determine the preferred binding sites by the three peptides. Footprinting experiments were carried out on a 158-mer DNA duplex (referred to as the Watson fragment and the corresponding 5′-^{32}P-labeled strand the Watson strand) and on a 135-mer DNA duplex (referred to as the Crick fragment and the corresponding 5′-^{32}P-labeled strand the Crick strand whose sequence is complementary to the Watson strand) that were prepared by the PCR technique. For experimental convenience we prefer to label the 5′-end of both Watson and Crick strands with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase. The sequences of the Watson and Crick fragments are copied from fragments of the pBR322 plasmid and are shown in Figure 2. The 5′-^{32}P-labeled 135-mer Crick fragment was subjected to DNase I footprinting and the peptide binding sites were located first. The primers for the preparation of the 5′-^{32}P-labeled 158-mer Watson fragment was then designed in a way that the proposed binding sites are located as close to the 5′-^{32}P-labeled end as possible for better visualization of the electrophoretic bands.

On the 158-mer Watson strand, PyH-12 binds strongly at positions 105–112, revealed as a broad DNase I blockage around the sequence 5′-GAGAAAAT-3′ (Figs 3 and 4a). Other significant blockages occur at positions 84–88, and 76–79, comprising the sequences 5′-AATAC-3′ and 5′-CGGT-3′, respectively. On the complementary 135-mer Crick strand, there is only a major sharp blockage at positions 108–110 around the sequence 5′-TTT-3′, centered at T110 and T109 and flanked on the 5′-side by another T residue and on the 3′-side by a C residue. Another small blockage occurs at positions 83–86 round the sequence 5′-ATTT-3′.

It is apparent that PyH-11 and PyH-10 displayed much less binding affinity (expressed as C_{50} values, Table 1)

Table 1. C_{50} values (μM) of peptides on various binding sites of the 158-mer (5′-^{32}P-labeled Watson strand) and 135-mer (5′-^{32}P-labeled Crick strand) DNA fragments

	PyH-12	PyH-11	PyH-10
5′-GAGAAAAT-3′	0.03		0.52
5′-AGGAGAAAAT-3′		1.16	
5′-CGGT-3′	0.01		
5′-AATAC-3′	0.01		0.5
5′-AATACCG-3′	1.38		
5′-TTT-3′	0.13		
5′-ATTTTCTC-3′		0.78	
5′-ATTTTCTCCTT-3′			3.31

C_{50} values are defined as the peptide concentration at which half of the maximal reduction in electrophoretic band intensity is observed and are computed according to the method of Sayers and Waring.[28]

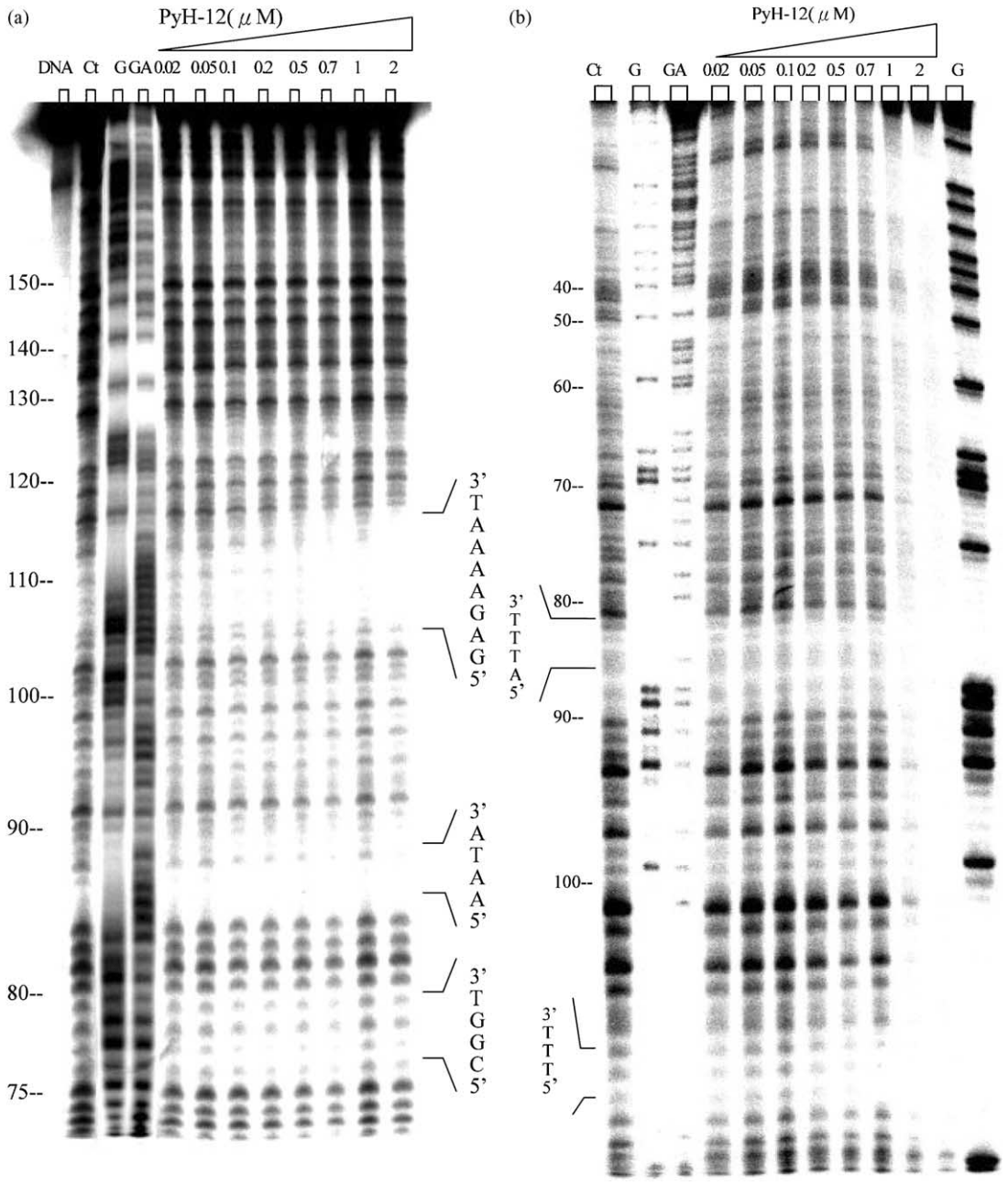

Figure 3. (a) Autoradiograph showing DNase I footprinting of PyH-12 bound to the 158-mer DNA duplex (Watson fragment) labeled with [γ-^{32}P]ATP on the 5′-end (Watson strand) in the presence of polynucleotide kinase. PyH-12 was incubated with DNA in 5 mM sodium cacodylate buffer (pH 6.5) at room temperature for 60 min before DNase I cleavage. G and GA represent Maxam-Gilbert purine sequencing tracks and Ct represents a DNase I digestion control lane. (b) Autoradiograph showing DNase I footprinting of PyH-12 bound to the 135-mer DNA duplex (Crick fragment) labeled with [γ-^{32}P]dATP on the 5′-end (Crick strand) in the presence of polynucleotide kinase. Reaction conditions as indicated in the legend of (a).

than PyH-12 at various preferred binding sites on both Watson and Crick strands. On the 158-mer Watson strand, PyH-11 showed less sequence-specific binding than PyH-12, with a broad DNase I blockage around positions 103–112 corresponding to the sequence 5′-AGGAGAAAAT-3′ (Figs 4b and 5a). Another major blockage by PyH-11 occurs around positions 84–90, comprising the sequence 5′-AATACCG-3′. Smaller but sharp blockages occur around positions 92–96 and 98–101, corresponding to the sequences 5′-ACAGA-3′ and 5′-GCGT-3′, respectively. On the 135-mer Crick strand, PyH-11 showed only one weak and broad DNase I blockage around positions 105–112, corresponding to the sequence 5′-ATTTTCTC-3′ (Figs 4b and 5b).

On the 158-mer Watson strand, the preferred binding sites of PyH-10 are similar to those of PyH-12 and spread around positions 105–112 comprising the sequence 5′-GAGAAAAT-3′ (Figs 4c and 5a). Another major blockage of PyH-10 occurs at positions 84–90 corresponding to the sequence 5′-AATACCG-3′. However, unlike PyH-12, no blockage at positions 76–79 around sequence 5′-CGGT-3′ is observed for PyH-10 or

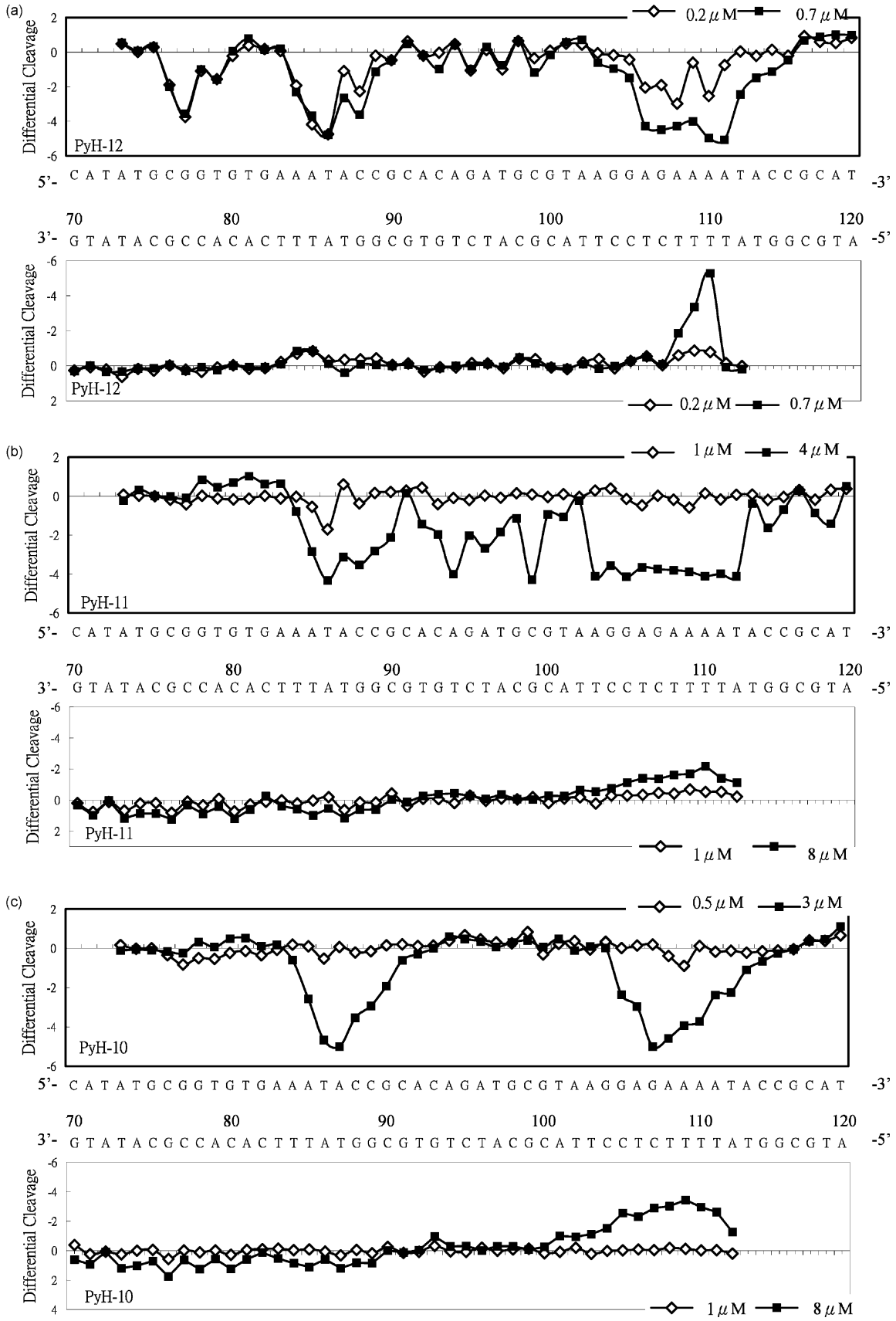

Figure 4. Differential cleavage plots comparing the susceptibility of the DNA fragment to DNase I after incubation with the peptide in 5 mM sodium cacodylate buffer (pH 6.5) at room temperature for 60 min. For each panel (a), (b) and (c), the upper plot represents the differential cleavage plot of PyH-12, PyH-11 and PyH-10 on the 158-mer duplex (5′-^{32}P-labeled Watson strand), respectively; the lower plot represents the differential cleavage plot of PyH-12, PyH-11 and PyH-10 on the 135-mer duplex (5′-^{32}P-labeled Crick strand), respectively.

PyH-11. Similar to the binding of PyH-11 on the 135-mer Crick strand PyH-10 shows only one broad blockage around positions 102–112, corresponding to the sequence 5′-ATTTTCTCCTT-3′.

The footprinting results indicate that PyH-12, PyH-11 and PyH-10 highly prefer sequences comprising four consecutive A residues and the binding of a molecule appears to cover both the Watson strand (consisting of four A's) and also the complementary Crick strand consisting of three or four T's. The preference of these peptides for A/T rich sites in the minor groove is likely due to hydrogen bonds directed from the protons of N_α, N_ε, or N_ω of arginyl residues, or N_α and N_ε of Lysyl residues to O2 of T and N3 of A.[29] Hydrogen bond formation may also be directed from the amide protons of 4-amino-1-methylpyrrole-2-carboxylic acid residues (Py) of the peptide to O2 of T and N3 of A. In many examples, the hydrogen donors of the peptide are unable to bond to the N3 group of G due to steric effect imposed by the exocyclic guanine 2-amino group.[30] Thus, the unusual strong preference of PyH-12 to the 5′-CGGT-3′ site is attributed to the result of hydrogen bonds between the imidazole imino group of histidine residues and guanine 2-amino protons. This kind of hydrogen bond is identical to that between the imid-

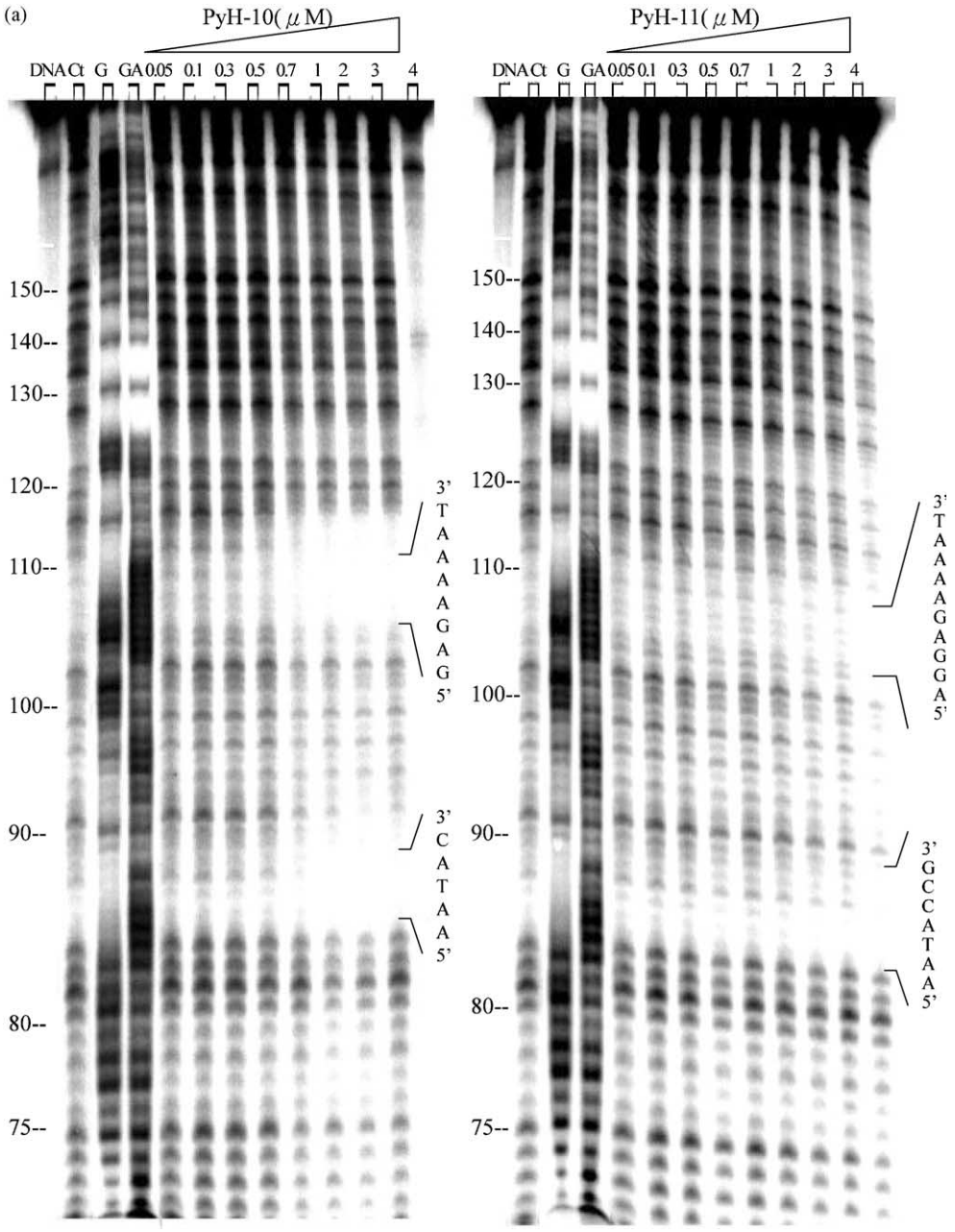

Figure 5. (a) Autoradiograph showing DNase I footprinting of PyH-11 (right panel) and of PyH-10 (left panel) bound to the 158-mer DNA duplex labeled with [γ-^{32}P]ATP on the 5′-end (Watson strand) in the presence of polynucleotide kinase. Reaction conditions as indicated in the legend of Figure 2(a). (b) Autoradiograph showing DNase I footprinting of PyH-11 (right panel) and of PyH-10 (left panel) bound to the 135-mer DNA duplex labeled with [γ-^{32}P]ATP on the 5′-end (Crick strand) in the presence of polynucleotide kinase. Reaction conditions as indicated in the legend of Figure 2a.

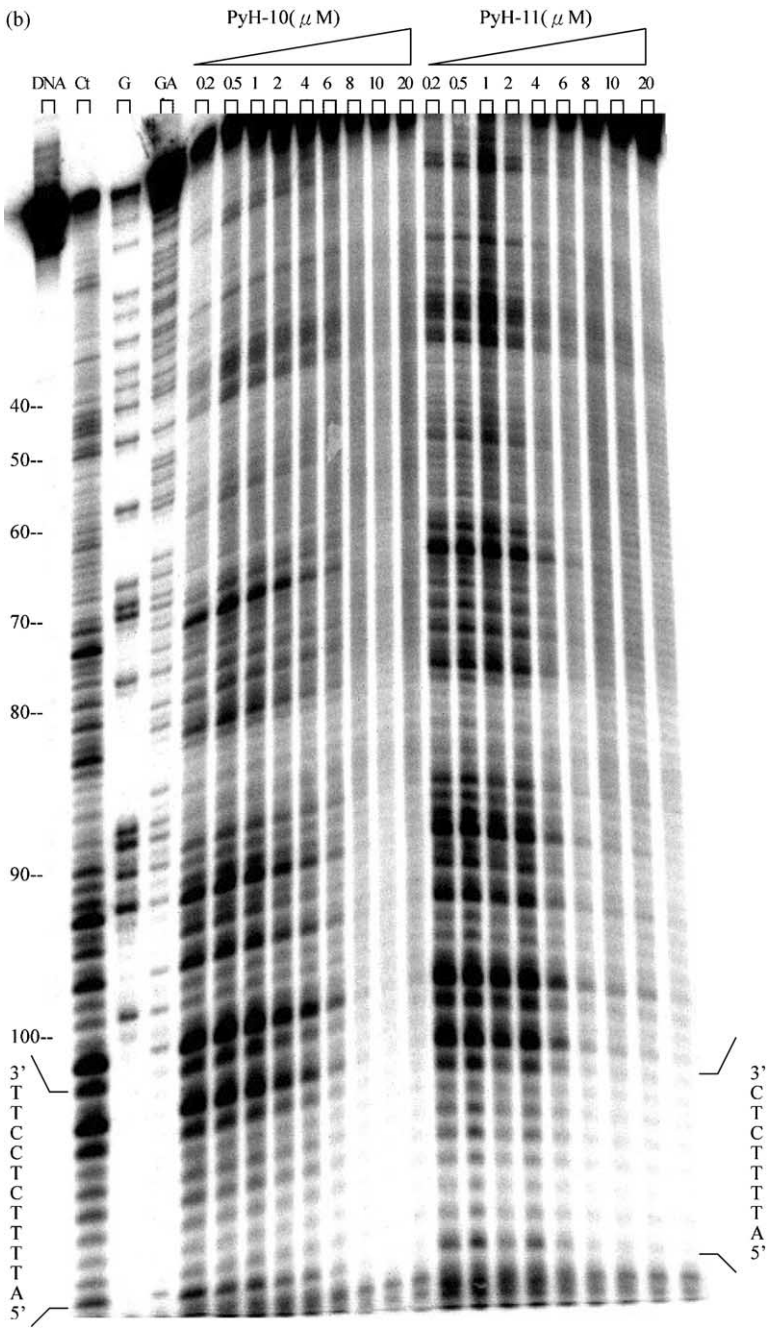

Figure 5. *Continued*

azole imino group of *N*-methylimidazole amino acid residues and guanine 2-amino protons in various synthetic polyamides.[13–15]

Variation in the number of Py residues in the peptide motif appears to induce some effects on sequence-specific binding to DNA. For example, all of the three peptides have strong preference to the 5′-GAGAAAAT-3′ site and to the 5′-AATAC-3′ site on the Watson strand. Of the three peptides, only PyH-12 has strong preference for the 5′-CGGT-3′ site (Fig. 4a). On the complementary Crick strand, only PyH-12 strongly discriminates the 5′-TTT-3′ site whereas both PyH-11 and PyH-10 have moderate and broad preference around the sequence 5′-ATTTTCTCCTT-3′ (Fig. 4b and c). Since PyH-11 has less specific binding around positions 91–100 of the Watson strand, and rather weak binding are observed around positions 105–112 of the Crick strand, the overall binding of PyH-11 to the Watson and Crick strands can only be termed sequence-selective.

It is noteworthy that all of the three peptides have preferential binding to more sites in the Watson strand (2–4 sites) than the Crick strand (one site). In the case of PyH-12, almost equal magnitude in the binding of the peptide is observed on the Watson strand around positions 108–111 of the sequence 5′-AAAA-3′ and on the complementary 5′-TTT-3′ sequence around positions 108–110 on the Crick strand (Fig. 4a). This strongly suggests that PyH-12 is involved in two to three biden-

tate interactions (two or more hydrogen bonds formed between a hydrogen donor of an amino acid residue and a base pair in the minor groove)[29] between the arginyl α-, ε- or ω-amino protons of the peptide with O2 of T and N3 of A at this binding site in the minor groove. In the case of PyH-11, bidentate interactions between the peptide with A/T base pairs is also possible but the blockage induced by this peptide is much stronger on the Watson strand around positions 103–112 corresponding to the sequence 5′-AGGAGAAAAT-3′ than on the complementary Crick strand corresponding to the sequence 5′-ATTTTCTC-3′ at the same site (Fig. 4b). That is, at this binding site the side-chain functions and α-amino protons of Arg and Lys (afford bidentate interactions) and the Py amide protons (afford single interactions)[29] of the peptide backbone have more hydrogen bond interactions with the N3 of A's on the Watson strand than with the O2 of T's on the Crick strand in the minor groove. Similar bidentate interactions are also identified in the binding of PyH-10 to the Watson and Crick strands around positions 102–112 (Fig. 4c). Complex interactions[29] between basic residues of the peptide with more than one base step are also possible yet we envisage in this study that bidentate interactions predominant in A/T base pairs in the minor groove.

Single interactions[29] (single hydrogen bond) directed from the hydrogen donors of basic amino acid residues or Py amide protons to DNA bases occur mostly on the Watson strand. For PyH-12, single interactions are observed around positions 76–79, comprising the sequence 5′-CGGT-3′. For PyH-10, single interactions are found at positions 85–90, around the sequence 5′-ATACCG-3′. The less specific bindings around positions 91–100 of the Watson strand by PyH-11 are also interpreted as single interactions of the peptide with the bases.

The higher sequence-specificity of PyH-12 over PyH-11 and PyH-10 is attributed to the significantly higher DNA binding affinity (expressed as C_{50} values)[28] of PyH-12 to preferred sites over its analogues (Table 1). The C_{50} value is equivalent to the apparent dissociation constant of the peptide at a specific site on the assumption that no significant cooperativity of binding between adjacent sites has occurred.[31]

This work reveals that the sequence-specific recognition of these minor groove binding agents is significantly dependent on bidentate and single interactions of the hydrogen donor or hydrogen acceptor of the peptide binding moieties (side-chain functions and α-amino protons) with DNA bases. The rather high DNA sequence-specificity of PyH-12 and PyH-10 suggests that these peptides can be chosen as potential agents for future research in the modulation of gene expression.

Acknowledgements

We thank Prof. Michael J. Waring for discussion and Miss. L. M. Hsu, Department of Chemistry, Chung-Hsing University for ESI mass analyses. This work was supported by grant NSC 91-2113-M029-003 to L.S. from the National Science Council, ROC, and by grant TCVGHT-T-927806 to L.S. and K.-C.G.J. from Taichung Veterans General Hospital.

References and notes

1. Katagiri, K.; Yoshida, T.; Sato, K. In *Antibiotics*; Corcoran, J. W., Hahn, F. E., Eds.; Springer: Berlin, 1975; Vol. 3, p 234.
2. Waring, M. J. In *Antibiotics*; Hahn, F. E., Eds.; Springer: Berlin, 1979; Vol. 5, p 173.
3. Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Hakoshima, T.; van der Marel, G. A.; van Boom, J. H.; Rich, A. *Science* **1984**, *225*, 1115.
4. Boger, D. L.; Chen, J.-H.; Saionz, K. W. *J. Am. Chem. Soc.* **1996**, *118*, 1629.
5. Chaires, J. B. In *Advances in DNA Sequence Specific Agents*; Hurley, L. H., Chaires, J. B., Eds.; Elsevier Health Sciences, 1996; Vol. 2, p 141.
6. Chaires, J. B.; Fox, K. R.; Herrera, J. E.; Britt, M.; Waring, M. J. *Biochemistry* **1987**, *26*, 8227.
7. Bailly, C.; Suh, D.; Waring, M. J.; Chaires, J. B. *Biochemistry* **1998**, *37*, 1033.
8. Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513.
9. Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *79*, 5470.
10. Portugal, J.; Waring, M. J. *Eur. J. Biochem.* **1987**, *167*, 281.
11. Abu-Daya, A.; Brown, F. M.; Fox, K. R. *Nucleic. Acids Res.* **1995**, *23*, 3385.
12. Bailly, C.; Colson, P.; Houssier, C.; Houssin, R.; Mrani, D.; Gosselin, G.; Imbach, J. L.; Waring, M. J.; Lown, J. W.; Heinichart, J. P. *Biochemistry* **1992**, *31*, 8349.
13. Urbach, A. R.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4343.
14. Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Science* **1994**, *266*, 646.
15. Chen, Y.-H.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 6995.
16. Sigman, D. S.; Mazumder, A.; Perrin, D. M. *Chem. Rev.* **1993**, *93*, 2295.
17. Helissey, P.; Bailly, C.; Vishwakarma, J. N.; Auclair, C.; Waring, M. J.; Giorgi-Renault, S. *Anti-Cancer Drug Des.* **1996**, *11*, 527.
18. Meunier, B. *Chem. Rev.* **1992**, *92*, 141.
19. Abraham, A. T.; Zhou, X.; Hecht, S. M. *J. Am. Chem. Soc.* **1999**, *121*, 1982.
20. Hashimoto, S.; Wang, B.; Hecht, S. M. *J. Am. Chem. Soc.* **2001**, *123*, 7437.
21. Minnock, A.; Lin, L. S.; Morgan, J.; Crow, S. D. G.; Waring, M. J.; Sheh, L. *Bioconjugate Chem.* **2001**, *12*, 870.
22. Cheng, C.-T.; Lo, V.; Chen, J.; Chen, W.-C.; Lin, C.-Y; Lin, H.-C.; Yang, C.-H.; Sheh, L. *Bioorg. Med. Chem.* **2001**, *9*, 1493.
23. Yang, C.-H.; Chou, P.-J.; Luo, Z.-L.; Chou, I.-C.; Chang, J.-C.; Cheng, C.-C.; Martin, C. R. H.; Waring, M. J.; Sheh, L. *Bioorg. Med. Chem.* **2003**, *11*, 3279.
24. Churchill, M. E. A.; Suzuki, M. *EMBO J.* **1989**, *8*, 4189.
25. Suzuki, M. *EMBO J.* **1989**, *8*, 797.
26. Suzuki, M. *J. Mol. Biol.* **1989**, *207*, 61.
27. Suzuki, M. *Nature* **1990**, *344*, 562.
28. Sayers, E. W.; Waring, M. J. *Biochemistry* **1993**, *32*, 9094.
29. Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. *Nucleic Acids Res.* **2001**, *29*, 2860.
30. Bailly, C.; Perrine, D.; Lancelot, J.-C.; Saturnino, C.; Robba, M.; Waring, M. J. *Biochem. J.* **1997**, *323*, 23, and references quoted therein.
31. Waring, M. J. Personal communications.