

Sequence-specific DNA alkylation by tandem Py-Im polyamides conjugates

Rhys Dylan Taylor^a, Yusuke Kawamoto^a, Kaori Hashiya^a, Toshikazu Bando^{*, a} and Hiroshi Sugiyama^{*, a, b, c}

Abstract: Tandem *N*-methylpyrrole–*N*-methylimidazole (Py–Im) polyamides were designed, and synthesized with good sequence-specific DNA alkylating activities. Three alkylating tandem Py–Im polyamides with different linkers, which have the same 10 bp DNA sequence recognition moiety, were evaluated for their reactivity and

selectivity on DNA alkylation using high-resolution denaturing gel electrophoresis. All three conjugates displayed high reactivity for the target sequence. Especially, conjugate **1** with β -alanine linker displayed the most selective sequence-specific alkylation toward the target 10-bp DNA sequences. The tandem Py–Im polyamides conjugates displayed

greater sequence specific DNA alkylation than conventional hairpin Py–Im polyamides conjugates (**4** and **5**). For further research, the design of tandem Py–Im polyamides conjugates would play the important roles for targeting specific gene sequences.

Keywords: Pyrrole-imidazole polyamide • DNA recognition • sequence-specific DNA alkylation

Introduction

N-methylpyrrole–*N*-methylimidazole (Py–Im) polyamides are small synthetic molecules based on the natural products distamycin A and netropsin that can bind to the minor groove of predetermined sequences of DNA.^[3–5] The antiparallel pairing of Im opposite Py (Im/Py) recognizes a G–C base pair, and Py opposite Py (Py/Py) recognizes A–T or T–A base pairs.

We have developed functionalized Py–Im polyamides targeting specific DNA sequences, and evaluated their biological properties.^[6] Py–Im polyamides conjugates with the alkylating moiety 1,2,9,9a-tetra-hydrocyclopropa[1,2-c]benz[1,2-e]indol-4-one (CBI)^[7–9] have DNA sequence-specific alkylating activities to alkylate at the N3 position of adenine at a predetermined site.^[10] Especially, we have

successfully targeted and induced sequence-specific alkylation at the Kras codon 13 mutation site.^[11] We have demonstrated to downregulate specific genes in mammalian cells by *seco*-CBI Py–Im polyamide conjugates.^[12] In contrast, some Py–Im polyamides conjugates with suberoylanilide hydroxamic acid (SAHA) as HDAC inhibitor successfully induced to upregulate multiple genes.^[13–15]

Dervan's group evaluated the binding affinity of several different tandem hairpin polyamides with various linkers that recognize 10, 11, and 12 bp sequences.^[26,27] Tandem hairpin Py–Im polyamides with a polyethylene glycol (PEG) linker have been used to bind to insect and vertebrate telomeric repeats,^[28] and to fluorescently label human telomeric repeats.^[1] Tandem hairpin Py–Im polyamides have also been shown to inhibit the human papilloma virus (HPV) by preventing the virus from binding to its target sequence.^[30] Recently, we demonstrated human telomere sequence-specific DNA alkylation by alkylating tandem hairpin Py–Im polyamides.^[XX]

To explore further the use of alkylating tandem hairpin Py–Im polyamides, we used high-resolution denaturing polyacrylamide gel electrophoresis (PAGE) to investigate the reactivity and selectivity of conjugates **1–3** containing the indole-*seco*-CBI^[25] alkylating moiety, and we compared their activities with those of two corresponding hairpin Py–Im polyamides, **4** and **5**. Here, we demonstrate that tandem hairpin Py–Im polyamides are more sequence specific and reactive than the corresponding hairpin Py–Im polyamides designed to target the same match site sequence.

[a] R. D. Taylor, Y. Kawamoto, K. Hashiya, Dr. T. Bando, Prof. Dr. H. Sugiyama

Department of Chemistry, Kyoto University,
Kitashirakawa-Oiwaketyo, Sakyo, Kyoto, 606-8502 (Japan)
Fax: (+) (+81) 75-753-3670
E-mail: hs@kuchem.kyoto-u.ac.jp

[b] Prof. Dr. H. Sugiyama
Institute for Integrated Cell-Materials Science (iCeMS) Kyoto University
Sakyo, Kyoto, 606-8502 (Japan)

[c] Prof. Dr. H. Sugiyama
CREST, Japan Science and Technology Corporation (JST)
Sanbancho, Chiyoda-ku, Tokyo 102-0075 (Japan)

Supporting information for this article is available on the WWW under <http://www.chemeurj.org/> or from the author.

Results and Discussion

Molecular design: Compared with our previous study,^[29] alkylating tandem hairpin Py-Im polyamides **1-3** were designed to recognize targeting 10-bp sequence, 5'-AGCCTTTCCA-3' (bold A representing the alkylation site). There are linker units as the only variation moieties in their structures, as shown in Figure 1. We decided to evaluate three linker units: β -alanine for **1**, which is often used in place of Py to facilitate binding of longer hairpins;^[24] PEG (n=2) for **2**, which has been reported to offer optimum binding for tandem hairpin Py-Im polyamides targeting telomeres;^[1,28] and a β -alanine dimer for **3** with an alkyl chain length between the other two linkers.

Hairpin Py-Im polyamide **4** was designed to demonstrate 5-bp sequence-specific DNA alkylation, 5'-TTCCA-3'. This would allow us to evaluate whether binding of the tandem hairpin was important for efficient DNA alkylation. The long hairpin Py-Im polyamide **5** was designed to recognize with the same 10 bp DNA sequence following to Py-Im recognition rules. This would allow a direct comparison between the reactivity and selectivity of the tandem polyamides against hairpin Py-Im polyamides.

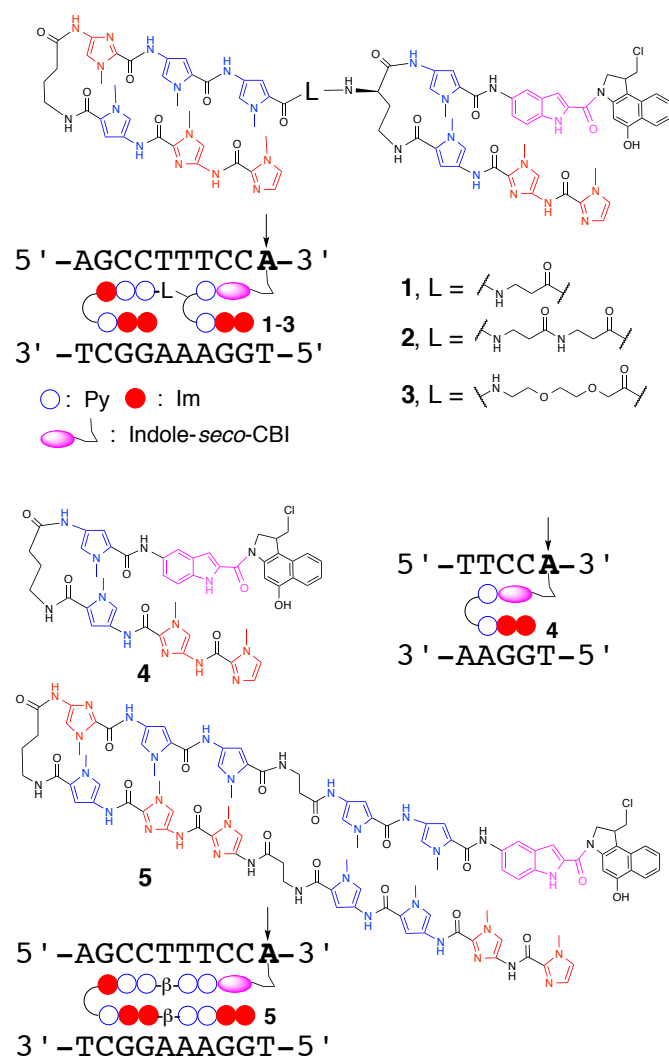


Figure 1. Chemical structures of tandem hairpin Py-Im polyamides **1-3** and hairpin Py-Im polyamides **4** and **5**. The Arrow and bold font of A in Schematic representation indicated DNA sequence-specific alkylation to target sites by each conjugates.

Synthesis: Following to the previous reported method for synthesizing all Py-Im polyamides conjugates,^[26,27,29] tandem hairpin Py-Im polyamides **1-3**, and hairpin Py-Im polyamides **4** and **5** were verified using reverse-phase HPLC and ESI-TOF mass spectrometry. These conjugates were then used in the DNA alkylation reactions with 205 bp DNA fragments containing target DNA sequences.

Evaluation of DNA alkylation: The alkylation strength of conjugates **1-5** was evaluated using 5'- Texas Red-labeled 205 bp DNA fragments and an automated DNA sequencer. DNA alkylation was conducted for 18 h at 25°C, after which the reaction was quenched with calf thymus DNA and then heated at 95°C for 5 min to cleave the DNA at the alkylation sites. Two DNA fragments were used for the DNA alkylation reactions for schematic representations; see Figure 2. DNA fragment 1 contained the site 1, 5'-AGCCTTTCCA-3' for target match sequence against conjugate **1-3** and **5**, and site 2, 5'-TATTTTCCA-3' for 3-bp mismatch sequence against conjugate **1-3** and **5**. DNA fragment 2 contained the site 1 and site 3, 5'-TGTCTTTCCA-3' for 1-bp mismatch sequence against conjugate **1-3** and **5**. Conjugate **4** was able to alkylate at site 1-3 for the target match sequences.



Figure 2. Schematic representation at the site 1-3 containing DNA fragments 1 and 2. Italic font of A-T base pairs indicated mismatch base pairs for Im/Py.

The alkylation activity of the tandem hairpin Py-Im polyamides **1-3** was evaluated using DNA fragment 1 (Figure 3). A small hairpin Py-Im polyamide **4** was used as a positive control with match sites at sites 1 and 2 (5'-TTCCA-3'). Alkylation by conjugate **1** (lanes 2-6) was observed from 100 nM (lane 4) and was clearly visible at concentrations of 500 nM and 1 μ M (lanes 5 and 6). Alkylation by conjugate **2** (lanes 7-11) occurred at a concentration as low as 50 nM and up to 500 nM (lanes 8-10) for site 1. At 1 μ M, nonspecific DNA alkylation was observed near the 5'-end with the DNA consumed completely at the 3'-end (lane 11). Conjugate **3** (lanes 12-16) induced alkylation from 50 nM up to 1 μ M (lanes 13-16) with the DNA consumed at the 3'-end at 500 nM and 1 μ M. conjugate **4** (lanes 17-21) displayed sequence-specific alkylation at both sites 1 and 2, at 50 nM and 100 nM (lanes 18 and 19), but only weak alkylation was observed at site 2. Minor alkylation also occurred at a 2 bp mismatch site, 5'-GTCGA-3' (site 4) at 50 nM and 100 nM. Nonspecific DNA alkylation was induced near the 5'-end at 500 nM and 1 μ M (lanes 20 and 21). Interestingly, conjugate **4** showed preference for site 1 over site 2 even though the binding sequence, 5'-TTCCA-3', was the same. One possible explanation is that an A-T-rich sequence preceded site 2, which may have induced some conformational change to the DNA, thereby inhibiting the efficient binding of **4**.

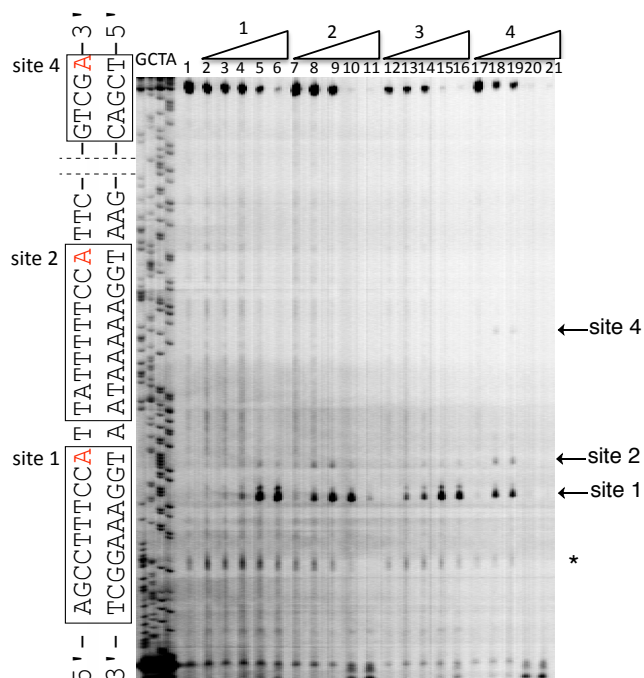


Figure 3. Thermally induced strand cleavage of the 5-Texas Red- labeled 205-bp DNA fragment 1 (10 nM) by conjugates 1, 2, 3, and 4, incubated for 18 h at 23°C at the following Py-Im polyamide concentrations; lane 1 = DNA control; lanes 2–6 = 10, 50, 100, 500, and 1000 nM of 1; lanes 7–11 = 10, 50, 100, 500, and 1000 nM of 2; lanes 12–16 = 10, 50, 100, 500, and 1000 nM of 3; lanes 17–21 = 10, 50, 100, 500, and 1000 nM of 4. * is not alkylation as the band is also visible in lane 1 (control).

Conjugates 1–3 displayed DNA alkylation at site 1 confirming that the sequence-specific DNA alkylation probably occurred with recognition by tandem hairpin Py-Im polyamides. However, the observation that the conjugate 4 also predominantly alkylate at site 1 indicated inadequate proof for the differences in specific DNA alkylation among tandem Py-Im polyamide conjugates 1–3.

To further explore the sequence selectivity of 1–4 for confirming clear differences among conjugate 1–3, DNA alkylation reactions were carried out with DNA fragment 2 (Figure 4). Similar results were observed at site 1 of DNA fragment 2, with all conjugates alkylating at the match site as expected. Conjugate 1 (lanes 2–6) induced efficient sequence-specific DNA alkylation from 50 nM up to 1 μ M (lanes 3–6), with strong alkylation occurring at 500 nM and 1 μ M (lanes 5 and 6). Weak alkylation was observed at site 3 at concentrations of 500 nM and 1 μ M, but this was much weaker than that at match site 1. No mismatch alkylation was observed at 50 nM or 100 nM. Conjugate 2 (lanes 7–11) induced alkylation at site 1 from 50 nM to 1 μ M (lanes 6–11); the strongest alkylation occurred between 100 nM and 500 nM. At site 3, strong alkylation was also induced from 50 nM to 500 nM, but there was no mismatch binding at 1 μ M. However nonspecific alkylation could be observed at the 5'-end from 500 nM, and the DNA was consumed almost completely at the 3'-end, especially at 1 μ M. Conjugate 3 (lanes 12–16) alkylated site 1 from 100 nM, with strong alkylation at 500 nM and 1 μ M (lanes 14–16). Strong alkylation at site 3 was observed from 100 nM to 1 μ M, the alkylation band intensity was indistinguishable from site 1. Conjugate 4 (lanes 17–21) yielded strong alkylation from 50 nM to 500 nM (lanes 18–20) at site 1, with weaker alkylation at 1 μ M (lane 21). Nonspecific alkylation occurred at the 5'-end at 500 nM and 1 μ M. Much stronger alkylation was observed at site 3 than site 2 of DNA fragment 1 (Figure 3) at 50 nM and 100 nM. Weak alkylation was again observed at the 2 bp mismatch site 4 at these concentrations. Above 100 nM, no alkylation was observed at site 3 or site 4.

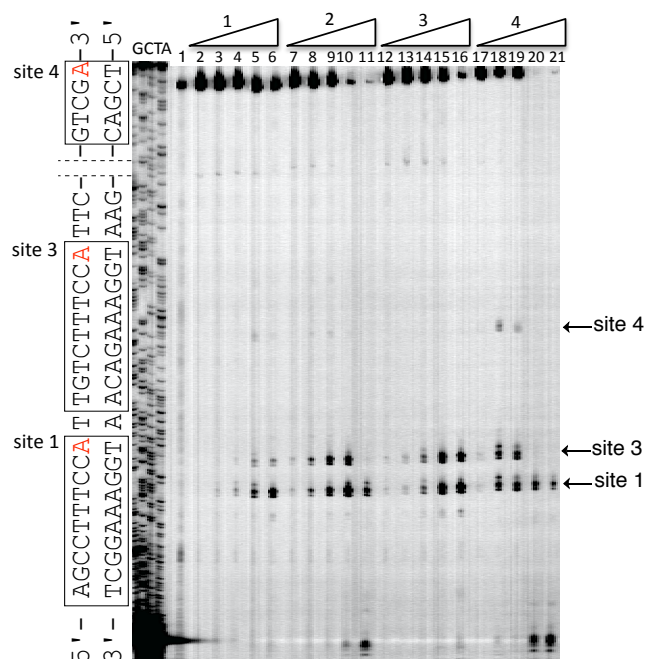


Figure 4. Thermally induced strand cleavage of the 5-Texas Red- labeled 205-bp DNA fragment 2 (10 nM) by conjugates 1, 2, 3, and 4, incubated for 18 h at 23°C at the following Py-Im polyamide concentrations; lane 1 = DNA control; lanes 2–6 = 10, 50, 100, 500, and 1000 nM of 1; lanes 7–11 = 10, 50, 100, 500, and 1000 nM of 2; lanes 12–16 = 10, 50, 100, 500, and 1000 nM of 3; lanes 17–21 = 10, 50, 100, 500, and 1000 nM of 4.

Conjugate 1 effectively alkylated at the target site, with only minor mismatch binding observed at higher concentrations of 500 nM and 1 μ M. However, conjugates 2 and 3 both induced specific and nonspecific DNA alkylation at sites 1 and 3, respectively. The alkylation intensities appeared almost identical for sites 1 and 3 for both conjugates, although at 1 μ M, conjugate 2 displayed no mismatch alkylation at site 3. However, nonspecific alkylation was clearly visible at the 5'-end, and the DNA was consumed almost completely at the 3'-end. Conjugates 2 and 3, with the β -alanine dimer and PEG linkers, respectively, were unable to differentiate between the target match site and 1 bp mismatch site. The longer linker lengths of the β -alanine dimer and PEG linkers in conjugates 2 and 3 likely allowed for more flexibility and conformational freedom, which might allow the alkylating hairpin to alkylate effectively at the N3 of adenine at lower concentrations than with 1. The shorter β -alanine linker between the two hairpins in conjugate 1 would allow for less conformational freedom, thereby restricting its alkylation; however, this linker also inhibited binding to the 1 bp mismatch observed for 2 and 3. Conjugate 4 induced sequence-specific DNA alkylation at match sites 1 and 3 but also at the 2 bp mismatch binding at site 4. Alkylation intensities were stronger in DNA fragment 2 at sites 1 and 3 than in DNA fragment 1 at sites 1 and 2. By comparison the intensity at site 4 remained the same in both DNA fragments. It should be noted that at site 1 multiple bands are visible on the gel for the same alkylation site. Thermal degradation at the abasic site induces β -elimination, which can in turn give rise to modified sugar moieties. The predominant product formed at the 3'- terminus is the trans- α , β -unsaturated aldose residue, however under heating conditions this can isomerize to the cis isomer through its hydrate. Thus the two isomers traverse through the gel at slightly different rates with a minor band appearing slightly above the main site on the gel. [38]

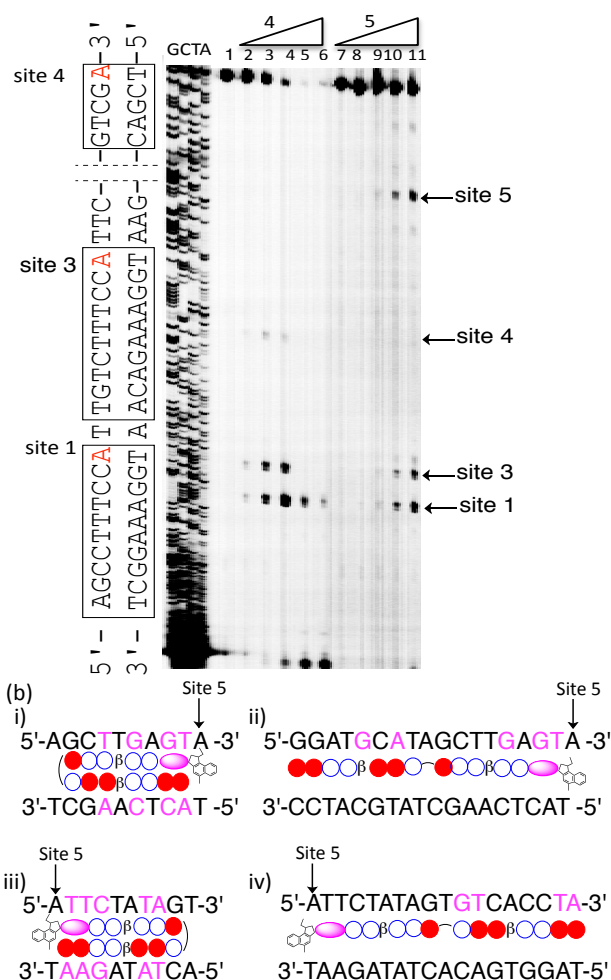


Figure 5. (a) Thermally induced strand cleavage of the 5-Texas Red- labeled 205-bp DNA fragment 2 (10 nM) by conjugates 4 and 5, incubated for 18 h at 23°C at the following Py-Im polyamide concentrations; lane 1 = DNA control; lanes 2–6 = 10, 50, 100, 500, and 1000 nM of 4; lanes 7–11 = 10, 50, 100, 500, and 1000 nM of 5. Unexpected alkylation was observed at site 5 for hairpin 5. It likely bound to the minor groove of DNA in a linear conformation. (b) Representation of possible binding orientation of hairpin 5 to the DNA at site 5. i) Forward binding hairpin conformation; ii) Forward binding linear conformation; iii) Reverse binding hairpin conformation; iv) Reverse binding linear conformation.

Conjugate 5 was designed with a β - β pairing at the same location as the linker in 1–3. Figure 4 shows the alkylation of DNA fragment 2 by hairpins 4 and 5. Conjugate 5 alkylated at the match site 1 at a concentration of 500 nM to 1 μ M (lanes 10 and 11). However, identical alkylation was also observed at the 1 bp mismatch site 3. Unexpected alkylation was also observed in the same concentration range at site 5 (Figure 5a). Site 5 represents a mismatch site for conjugate 5, if 5 was orientated in the forward N–C direction with regard to the 5'–3' direction of the DNA duplex, 5'-AGCTTGAGTA-3' would correspond to a 4 bp mismatch site. However, if orientated in the reverse direction, C–N for the 5'–3' direction, 5 would correspond to the 5 bp mismatch site 5'-ATTCTATAGT-3'. Conversely, it has been reported that hairpin Py-Im polyamides can bind in a reverse linear conformation.^[40] If 5 were to bind to site 5 in this orientation the first 10 bp would constitute a match sequence with the indole linker degenerately recognizing AT (Figure 5b). Though if the second half of the polyamide also bound to the minor groove, a 4 bp mismatch would be observed. In addition the reverse linear conformation is the only one from the four possibilities in which the indole-*seco*-CBI unit is not involved in mismatch binding.

We have demonstrated that the tandem hairpin conjugates 1–3 plays an important role in sequence-specific DNA alkylation. Compared with hairpin Py-Im polyamides 4 and 5, tandem hairpin Py-Im polyamides 1 demonstrated superior reactivity and selectivity for the target sequence; the presence of β -alanine linker in 1 caused the good sequence-specific DNA alkylation.

Molecular modeling studies of conjugate 1: To gain insight into the alkylating reactivity of conjugate 1, we performed molecular modeling studies of the 5'-dTCGATAGCCTTCCATTGTC3'-5'-dGACAATGGAAAG-GCTATCGA-3'-1 complex, as shown in Figure 6. After minimization of the complex and the water solvent, The energy minimized structure showed that 1 could efficiently bind in the DNA minor groove, allowing for effective DNA alkylation to occur. The distance between N3 of adenine and C9 of the cyclopropane unit of CBI was measured as only 3.23Å, and the angle between the N3–C9–C8b was 139.1°. The N-terminal Im–Im–Py of both the hairpins in 1 retained a planar conformation, allowing for efficient binding between the polyamide and DNA. This planar conformation would also contribute to stabilizing the CBI unit in the leading hairpin for efficient alkylation. From the minimized structure, we also observed that the β -alanine linker worked well to span the 1 bp distance required between the hairpins and to keep both hairpins bound tightly to their corresponding base pairs in the minor groove.

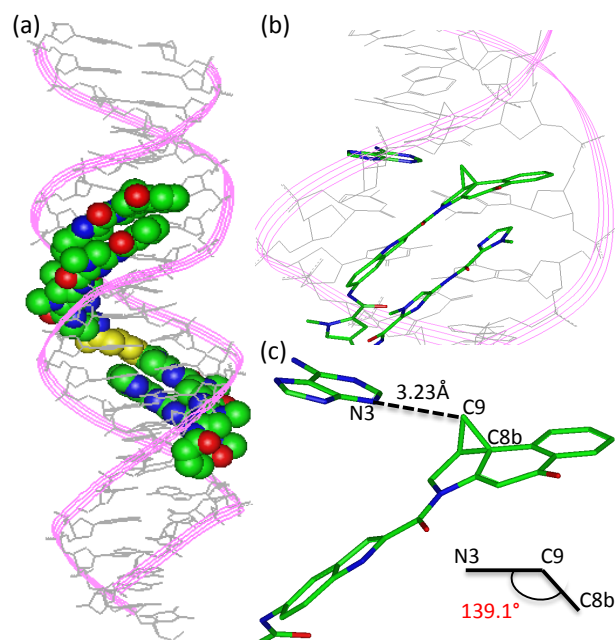


Figure 6. (a) Energy-minimized structure of the 5'-dAGCCTTCCATTGTCTTCCA-3' / 5'-dTGGAAAGACAATGGAAAGGCT-3' - 1 complex. The β linker is highlighted in yellow color. The water layer, Hydrogen atoms, and Na ions have been removed. (b) Close up. The indole-*seco*-CBI and adenine have been highlighted and the N-terminal tail is also visible. (c) The distance between the indole-*seco*-CBI moiety and adenine and the angle between the C8b, C9 of the cyclopropane and N3 of adenine was measured.

Conclusion

Alkylating tandem hairpin Py-Im polyamides was designed to synthesized as sequence-specific DNA alkylating conjugates. We have shown that tandem Py-Im polyamides conjugate 1 can target a 10 bp sequence, 5'-AGCCTTCCA-3', with good reactivity and

selectivity compared with regular hairpin Py–Im polyamides conjugate **4** and **5**. The shorter β -alanine linker had superior selectivity compared with the other long linkers. This confirms earlier reports that excessive linker length should not be used for efficient binding.^[26] Our findings would suggest that alkylating tandem Py–Im polyamides would apply for targeting specific gene sequences on further research.

Experimental Section

General: Reagents and solvents were purchased from standard suppliers and used without further purification. Solid phase peptide synthesis was carried out using PSSM-8 peptide synthesizer (Shimadzu, Kyoto), with a computer-assisted operating system. The oxime resin (200–400 mesh) was purchased from Novabiochem. Water was purified with a Millipore Direct-Q™ (MQ). HPLC was performed with a Jasco PU-2089 HPLC pump, a UV-2075 HPLC UV/ VIS detector, a Chemcobond 5-ODS-H 4.6 x 150mm column (Chemco Scientific) was used for analysis with 0.1% TFA in water and CH₃CN as eluent at a flow rate of 1.0 mL/min, and a linear gradient elution of 0–100% CH₃CN over 20 min with detection at 254 nm. A Chemcobond 5-ODS-H 10 x 150mm column (Chemco Scientific) was used for purification with 0.1% TFA in water and CH₃CN as eluent at a flow rate of 3.0 mL/min, with a linear gradient elution of 0–100% CH₃CN over 20 min and with detection at 254 nm. Flash chromatography purification was conducted using a Combiflash RF by Teledyne Isco Inc. using C18 silica as the stationary phase with 0.1% TFA in water and CH₃CN as eluent at a flow rate of 18.0 mL/min with an overall gradient elution of 0–100% CH₃CN over 60 min, (with an elution of 0 for first 5 min, 0–50% from 5 min–50 min, 50% from 50 min–55 min, 100% from 55 min–60 min). ESI-TOFMS was produced on a BioTOF II (Bruker Daltonics) mass spectrometer using a positive ionization mode. UV spectra were measured on a Nanodrop ND-1000 spectrophotometer. The DNA oligonucleotides and the 5'-Texas Red-labeled primers were purchased from Sigma Aldrich. A GenElute™ plasmid mini prep kit and Genelute™ PCR clean-up kit was purchased from Sigma-Aldrich. A Thermo Sequence core sequencing kit and loading dye (formamide with fuchsin red) were purchased from GE Healthcare. Polymerase chain reaction (PCR) was performed using an iCycler (Bio-Rad). Polyacrylamide gel electrophoresis was performed using a HITACHI SQ5500-E DNA sequencer and data were analyzed by FLAGLYS version 2 software (HITACHI). NMR spectra were recorded using a JEOL JNM ECA-600 nuclear magnetic resonance spectrometer with tetramethylsilane as internal standard. Proton NMR spectra were recorded at 600MHz in parts per million (ppm) downfield relative to tetramethylsilane. The following abbreviations apply to spin multiplicity: s (singlet), d (doublet), t (triplet), q (quartet), qu (quintet), m (multiplet) and brd (broad). NH₂-Indole-*seco*-CBI was synthesized by reported procedures.^[7–9,41] Synthesis of intermediates **6** and **7**: were synthesized from NO₂-ImPy-CO₂Me as previously reported.^[11,34]

Solid-phase synthesis of Py-Im polyamides: Py–Im polyamides supported by oxime resin (for the synthesis of the conjugates 1–5) were prepared in a stepwise reaction using reported Fmoc solid-phase protocol^[35–37] utilizing compounds **6** and **7** and FmocPyCO₂H, FmocIm-CO₂H, FmocPyIm-CO₂H, Fmoc-b-CO₂H, HCTU, DIEA, DMF. Cleavage from the oxime resin was performed with a 1:1 mixture of 2M NaOH and 1,4-Dioxane at 55°C for 3 hours. The resulting Py-Im polyamide carboxylic acids were acidified with CH₃COOH to pH 6, precipitated with Et₂O. The corresponding carboxylic acids of **1**, **2**, **4**, and **5**, were washed with milli-Q water and lyophilized. The corresponding crudes were used as coupling partners with NH₂-Indole-*seco*-CBI without further purification. The carboxylic acid precursor for conjugate **3** was purified prior to the coupling step using flash chromatography purification (0.1% TFA/CH₃CN with an overall gradient of 0–100% 0–60min) and the peak fraction containing the product was collected. Py–Im polyamide carboxylic acid (for the synthesis of **1**): ¹H NMR (600 MHz, DMSO-*d*₆) δ 12.17 (brd, 1H; OH), 10.26 (s, 2H; NH), 10.25 (s, 1H; NH), 9.99 (s, 1H; NH), 9.94 (s, 1H; NH), 9.89 (s, 1H; NH), 9.71 (s, 2H; NH), 8.27 (d, J=8.4 Hz, 1H; NH), 8.05 (m, 2H; NH), 7.97 (t, J=4.8 Hz, 1H; NH), 7.57 (s, 1H, CH), 7.56 (d, 1H, CH), 7.45 (s, 2H; CH), 7.31 (d, J=2.4 Hz, 1H; CH), 7.26 (d, J=1.2Hz, 1H; CH), 7.23 (d, J=2Hz, 1H;CH), 7.22(d, J=2Hz, 1H; CH), 7.19 (s, 1H; CH), 7.17 (s, 1H; CH), 7.11 (s, 1H; CH), 7.08 (s, 1H; CH), 7.07 (s, 1H; CH), 7.00 (s, 1H; CH), 6.97 (d, J=1.2Hz, 1H; CH), 6.86 (s, 1H; CH), 6.74 (d, J=2.4 Hz, 1H; CH), 4.41 (m, 1H; CH), 4.01 (s, 6H; NCH₃), 4.00 (s, 3H; NCH₃), 3.99 (s, 3H; NCH₃), 3.95 (s, 3H; NCH₃), 3.84 (s, 3H; NCH₃), 3.81(s, 6H; NCH₃), 3.79 (s, 6H; NCH₃), 3.21 (m, 4H; CH₂), 3.00 (m, 2H; CH₂), 2.45 (m, 2H;CH₂), 2.36 (m, 2H;CH₂), 1.78 (m, 4H, CH₂). ESI-TOF-MS m/e for C₆₆H₇₆N₂₇O₁₄ [M +H]⁺ calcd 1470.6065, found 1470.6072. Analytical HPLC: t_R = 9.5 min (0.1% TFA/CH₃CN 0–100% linear gradient, 0–20 min).

Synthesis of Conjugate 1 by coupling with NH₂-indole *seco*-CBI: A solution of Py–Im polyamide carboxylic acid **1** (1.0 mg, 0.68 μ mol), DMF (100 μ L), iPr₂NEt (0.5 μ L, 2.86 mmol) and PyBOP (1.0 mg, 1.95 mmol) were stirred for 30 min at room temperature. After the formation of the activated 1-hydroxybenzotriazole ester was checked by analytical HPLC: t_R = 12.0 min (0.1% TFA/CH₃CN 0–100% linear gradient,

0–20 min), NH₂-indole *seco*-CBI (0.6 mg, 1.58 mmol) to stir overnight at room temperature under N₂ atmosphere. Et₂O was added to the mixture, and the precipitate was collected by centrifugation and washed by Et₂O and CH₂Cl₂, followed by HPLC purification (0.1% TFA/CH₃CN 20–65% linear gradient, 0–30 min). After collecting the peak and lyophilization, conjugate **1** was obtained as a yellow solid (1.0 mg, 0.54mmol, 80%). ESI-TOF-MS m/z calcd for Chemical Formula: C₈₈H₉₁CIN₃₀O₁₅ [M+2H]²⁺ 922.3576; found 922.3564. Analytical HPLC: t_R = 12.9 min (0.1% TFA/CH₃CN 0–100% linear gradient, 0–20 min). According to the synthetic procedure for conjugate **1**, conjugates **2**, **3**, **4**, and **5** were synthesized from their corresponding carboxylic acids. Conjugate **2** was obtained as a yellow powder. ESI-TOF-MS m/z calcd for Chemical Formula: C₉₁H₉₆CIN₃₁O₁₆, [M+2H]²⁺ 957.8765; found 957.8769. Analytical HPLC: t_R = 12.7 min (0.1% TFA/CH₃CN 0–100% linear gradient, 0–20 min). Conjugate **3** was obtained as a brown powder. ESI-TOF-MS m/z calcd for Chemical Formula: C₉₁H₉₇CIN₃₀O₁₇ [M+2H]²⁺ 959.3746; found 959.3763. Analytical HPLC: t_R = 12.9 min (0.1% TFA/CH₃CN 0–100% linear gradient, 0–20 min). Conjugate **4** was obtained as a yellow powder. ESI-TOF-MS m/z calcd for Chemical Formula: C₄₈H₄₆CIN₁₅O₇ [M+H]⁺ 9952.3410; found 952.3440. Analytical HPLC: t_R = 13.6 min (0.1% TFA/CH₃CN 0–100% linear gradient, 0–20 min). Conjugate **5** was obtained as a yellow powder. ESI-TOF-MS m/z calcd for Chemical Formula: C₉₉H₁₀₁CIN₃₄O₁₇ [M+2H]²⁺ 1037.3964; found 1037.3942. Analytical HPLC: t_R = 13.7 min (0.1% TFA/CH₃CN 0–100% linear gradient, 0–20 min). All spectra can be found in the supporting information.

Cloning of 205 base pair DNA fragment: All oligonucleotides were purchased from Sigma-Aldrich, annealed at a concentration of 10 μ M and ligated into the pGEM–T Easy vector (Promega). Escherichia coli DH5a competent cells (TOYOBO) were transformed and cultured on an LB plate with 100 μ g/mL ampicillin and 32 μ g of X-gal (20 mg/mL), 25 μ L IPTG 100mM, 25 μ L overnight at 37 °C. White colonies were identified by colony PCR in 20 μ L of the reaction mixtures containing 250nM of each primer (T7 primer, 5'-TAATACGACTCACTATAGGG-3'; SP6 primer, 5'-TATTTAGGTGACACTATAG-3'), 200 μ M dNTPs (Sigma Aldrich), 2 units of Taq polymerase, and 1 \times ThermoPol reaction buffer (New England Bio Labs). Amplification of the DNA fragments was carried out incubating at 95 °C for 5 min, followed by 35 cycles of 95 °C for 35 s, 50 °C for 35 s, 72 °C for 30 s, with a final extension step of 72 °C for 7 min. The appropriate colony was selected for transfer to 5 mL of LB medium with 100 μ g/mL ampicillin and cultured overnight at 37 °C. The plasmids were extracted using a GenElute™ Plasmid Miniprep Kit (Sigma Aldrich) and identified by PCR (program and reaction mixtures were the same as above). The 5'-Texas Red-modified 205 bp DNA fragment was prepared by PCR with 5'-Texas Red-modified T7 primer from 1ng/ L of the 45bp fragments inserted into the pGEM-T Easy vector (program and reaction mixtures were the same as above). Fragments were purified by GenElute™ PCR Clean-up Kit (Sigma Aldrich), and their concentrations were determined by UV absorption.

High-Resolution Gel Electrophoresis: The 5'-Texas Red-labeled 205bp DNA fragment (6nM) was alkylated by the PI polyamide conjugates at various concentrations ranging from 10nM to 1 μ M in DMF with 5 μ M sodium phosphate buffer (pH 7.0) at 23 °C for 18 h. After incubation, the reaction mixture was quenched by calf thymus DNA and heated for 5 min at 95 °C. The solution was concentrated by vacuum centrifugation. The pellet was dissolved with 6 μ L of loading dye (formamide with fuchsin red), heated at 95 °C for 20 min, and then immediately cooled to 0 °C. A 1.2 μ L aliquot was subjected to electrophoresis on a 6% denaturing polyacrylamide gel using a Hitachi DNA sequencer.

Molecular Modeling Studies on the DNA Complex of conjugate 1: Minimizations were performed with the Discover (MSI, San Diego, CA) program using CVFF force-field parameters. The starting structure was constructed using builder module of the program Insight II using standard bond lengths and angles. Where the three upper and lower sides of Watson–Crick base pairs were fixed, conjugate **1** was inserted in 5'-d TCGATAGCCTTTCCATTGTC-3'/5'-dGACAATGGAAAG-GCTATCGA-3'. Thirty-eight Na cations were placed at the bifurcating position of the O–P–O angle at a distance of 2.21 Å from the phosphorus atom. The resulting complex was soaked in a 15Å layer of water. The layer of water was minimized without constraints to the stage where the rms was less than 0.001 kcal/mol·Å using the steepest then the conjugate algorithm; successively the whole complex was also minimized in the same way.

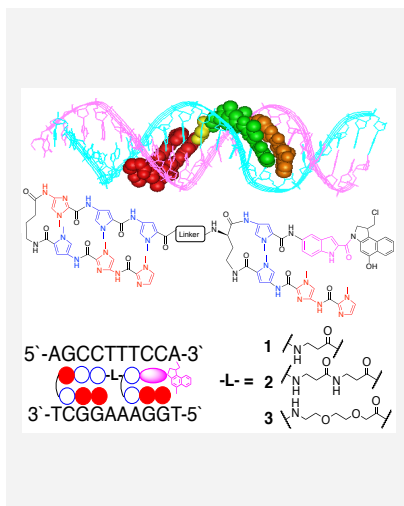
Acknowledgements

This work was supported by a Grant-in-Aid for Priority Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and Core Research for Evolutional Science and Technology (CREST) from Japan Science and Technology (JST).

- [1] Y. Kawamoto, T. Bando, F. Kamada, Y. Li, K. Hashiya, K. Maeshima, H. Sugiyama, *J. Am. Chem. Soc.* **2013**, *135*, 16468–16477.
- [2] W. a Denny, *Curr. Med. Chem.* **2001**, *8*, 533–44.

- [3] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–35.
- [4] P. B. Dervan, B. S. Edelson, *Curr. Opin. Struct. Biol.* **2003**, *13*, 284–299.
- [5] P. B. Dervan, R. M. Doss, M. A. Marques, *Curr. Med. Chem. Anticancer. Agents* **2005**, *5*, 373–87.
- [6] G. Kashiwazaki, T. Bando, T. Yoshidome, S. Masui, T. Takagaki, K. Hashiya, G. N. Pandian, J. Yasuoka, K. Akiyoshi, H. Sugiyama, *J. Med. Chem.* **2012**, *55*, 2057–66.
- [7] D. L. Boger, T. Ishizaki, P. A. Kitos, O. Suntornwat, *J. Org. Chem.* **1990**, *55*, 5823–5832.
- [8] D. L. Boger, W. Yun, B. R. Teegarden, *J. Org. Chem.* **1992**, *57*, 2813–2876.
- [9] D. L. Boger, J. A. McKie, *J. Org. Chem.* **1995**, *60*, 1271–1275.
- [10] T. Bando, H. Sugiyama, *Acc. Chem. Res.* **2006**, *39*, 935–44.
- [11] R. D. Taylor, S. Asamitsu, T. Takenaka, M. Yamamoto, K. Hashiya, Y. Kawamoto, T. Bando, H. Nagase, H. Sugiyama, *Chem. - A Eur. J.* **2013**, *19*, 1–9.
- [12] K. Shinohara, A. Narita, T. Oyoshi, T. Bando, H. Teraoka, H. Sugiyama, *J. Am. Chem. Soc.* **2004**, *126*, 5113–5118.
- [13] G. N. Pandian, Y. Nakano, S. Sato, H. Morinaga, T. Bando, H. Nagase, H. Sugiyama, *Sci. Rep.* **2012**, *2*, 544.
- [14] G. N. Pandian, H. Sugiyama, *Biotechnol. J.* **2012**, *7*, 798–809.
- [15] L. Han, G. N. Pandian, S. Junetha, S. Sato, C. Anandhakumar, J. Taniguchi, A. Saha, T. Bando, H. Nagase, H. Sugiyama, *Angew. Chem. Int. Ed. Engl.* **2013**, 1–5.
- [16] J. J. Kelly, E. E. Baird, P. B. Dervan, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 6981–5.
- [17] M. E. Parks, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6153–6159.
- [18] J. M. Turner, S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 6219–6226.
- [19] C. C. C. Wang, U. Ellervik, P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 653–657.
- [20] C. R. Woods, T. Ishii, B. Wu, K. W. Bair, D. L. Boger, *J. Am. Chem. Soc.* **2002**, *124*, 2148–52.
- [21] P. B. Dervan, A. R. Urbach, *Essays in Contemporary Chemistry*, Verlag Helvetica Chimica Acta, Zürich, **2001**.
- [22] S. Wang, R. Nanjunda, K. Aston, J. K. Bashkin, W. D. Wilson, *Biochemistry* **2012**, *51*, 9796–806.
- [23] J. K. Bashkin, K. Aston, J. P. Ramos, K. J. Koeller, R. Nanjunda, G. He, C. M. Dupreux, W. David Wilson, *Biochimie* **2013**, *95*, 271–9.
- [24] T. Bando, M. Minoshima, G. Kashiwazaki, K. Shinohara, S. Sasaki, J. Fujimoto, A. Ohtsuki, M. Murakami, S. Nakazono, H. Sugiyama, *Bioorg. Med. Chem.* **2008**, *16*, 2286–91.
- [25] T. Bando, S. Sasaki, M. Minoshima, C. Dohno, K. Shinohara, A. Narita, H. Sugiyama, *Bioconjug. Chem.* **2006**, *17*, 715–20.
- [26] I. Kers, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 3339–49.
- [27] D. M. Herman, E. E. Baird, P. B. Dervan, *Chem. - A Eur. J.* **1999**, *5*, 975–983.
- [28] K. Maeshima, S. Janssen, U. K. Laemmli, *EMBO J.* **2001**, *20*, 3218–28.
- [29] S. Sasaki, T. Bando, M. Minoshima, K. Shinohara, H. Sugiyama, *Chem. - A Eur. J.* **2008**, *14*, 864–70.
- [30] T. D. Schaal, W. G. Mallet, D. L. McMinn, N. V. Nguyen, M. M. Sopko, S. John, B. S. Parekh, *Nucleic Acids Res.* **2003**, *31*, 1282–1291.
- [31] W. Su, S. J. Gray, R. Dondi, G. A. Burley, *Org. Lett.* **2009**, *11*, 3910–3.
- [32] D. M. Chenoweth, D. A. Harki, P. B. Dervan, *J. Am. Chem. Soc.* **2009**, *131*, 7175–81.
- [33] M. Wetzler, D. E. Wemmer, *Org. Lett.* **2010**, *12*, 3488–90.
- [34] T. Bando, A. Narita, I. Saito, H. Sugiyama, *J. Am. Chem. Soc.* **2003**, *125*, 3471–85.
- [35] E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6141–6146.
- [36] N. R. Wurtz, J. M. Turner, E. E. Baird, P. B. Dervan, *Org. Lett.* **2001**, *3*, 1201–3.
- [37] J. M. Belitsky, D. H. Nguyen, N. R. Wurtz, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 2767–74.
- [38] H. Sugiyama, T. Fujiwara, A. Ura, *Chem. Res. Toxicol.* **1994**, 673–683.
- [39] M. Minoshima, T. Bando, S. Sasaki, K. Shinohara, T. Shimizu, J. Fujimoto, H. Sugiyama, *J. Am. Chem. Soc.* **2007**, *129*, 5384–90.
- [40] J. L. Meier, A. S. Yu, I. Korf, D. J. Segal, P. B. Dervan, *J. Am. Chem. Soc.* **2012**, *134*, 17814–22.
- [41] T. Bando, A. Narita, S. Sasaki, H. Sugiyama, *J. Am. Chem. Soc.* **2005**, *127*, 13890–5.

Rhys Dylan. Taylor, Yusuke Kawamoto, Kaori Hashiya, Toshikazu Bando* and Hiroshi Sugiyama*. Page – Page



Tandem hairpin *N*-methylpyrrole (Py)-*N*-methylimidazole (Im) polyamide *seco*-CBI conjugates were designed with various linker units and were compared against hairpin Py-Im polyamides. High-resolution denaturing gel electrophoresis using 205 base pair (bp) DNA fragments was used to compare their alkylating reactivity and selectivity.