Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.
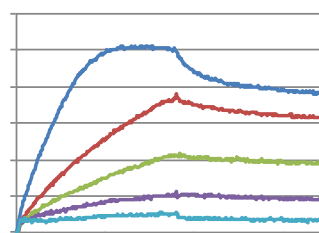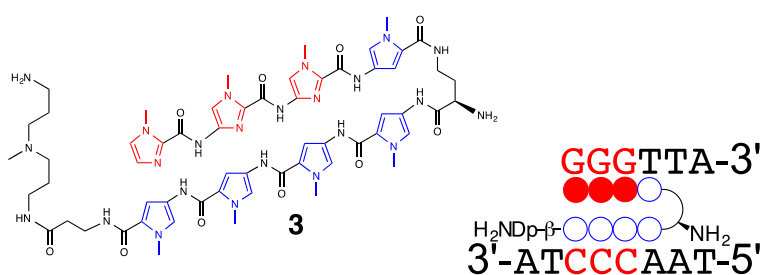
Rational Design of Specific Binding Hairpin Py–Im Polyamides Targeting Human Telomere Sequences

Leave this area blank for abstract info.

Chuanxin Guo[a], Yusuke Kawamoto[a], Sefan Asamitsu[a], Yoshito Sawatani[a], Kaori Hashiya[a], Toshikazu Bando[a,*], Hiroshi Sugiyama[a,b,*]

[a]*Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan*
[b]*Institute for Integrated Cell-Materials Science (iCeMS), Kyoto University, Sakyo, Kyoto 606-8501, Japan*

Rational design of specific binding hairpin Py–Im polyamides targeting human telomere sequences

Chuanxin Guo[a], Yusuke Kawamoto[a], Sefan Asamitsu[a], Yoshito Sawatani[a], Kaori Hashiya[a], Toshikazu Bando[a,*], Hiroshi Sugiyama[a,b,*]

[a]Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan
[b]Institute for Integrated Cell-Materials Science (iCeMS), Kyoto University, Sakyo, Kyoto 606-8501, Japan

ARTICLE INFO

ABSTRACT

N-Methylpyrrole (Py)–N-methylimidazole (Im) polyamides are organic molecules that can recognize predetermined DNA sequences in a sequence-specific manner. Human telomeres contain regions of (TTAGGG)$_n$ repetitive nucleotide sequences at each end of chromosomes, and these regions protect the chromosome from deterioration or from fusion with neighboring chromosomes. The telomeres are disposable buffers at the ends of chromosomes that are truncated during cell division. Tandem hairpin Py–Im polyamide TH59, which recognizes human telomere sequences, was reported by Laemmli's group in 2001. Here, we synthesized three types of Py–Im polyamides **1**–**3** based on TH59 for specific recognition of human telomere repeat sequences. Thermal melting temperature (T_m) measurements and surface plasmon resonance analysis were used to evaluate the abilities of the three types of Py–Im polyamides to discriminate between three kinds of DNA sequences. Significantly, the results showed that polyamides **1** and **2** have better affinities to TTAAGG than to TTAGGG. In contrast, polyamide **3** displayed good specificity to human telomere sequence, TTAGGG, as expected on the basis of Py–Im binding rules

1. Introduction

[1]Telomeres are located at each end of chromosomes and these regions have important roles relating to life processes.[1] In the human genome, telomeres are composed of double-stranded d(TTAGGG)/d(CCCTAA) repeats and single-stranded d(TTAGGG) regions running from the 5'- to the 3'-end. As part of the cell aging process, the lengths of telomeres decrease by 50–150 bp with each cell division.[2,3]

Polyamides containing N-methylpyrrole (Py) and N-methylimidazole (Im) amino acids bind double-stranded DNA in the minor groove and uniquely recognize each of the four Watson–Crick base pairs.[4] The manner of recognition is regulated by general rules: antiparallel pairing of Im opposite Py recognizes a G•C base pair and Py/Im pairing recognizes a C•G pair, whereas Py/Py pairing recognizes an A•T or T•A base pair.[5a] It was reported that replacement of Py with an aliphatic β-alanine (β) increases its selectivity due to the flexibility of the β component.[5b] Furthermore, the inclusion of C-terminal β-alanine and the turn moieties of γ-amino butyric acid and 2,4-diaminobutyric acid (Dab) residues lead to recognition of A•T or T•A base pairs.[5c,d]

The ability of the tandem hairpin Py–Im polyamide shown in Figure 1 to recognize human telomere sequences was reported by Laemmli's group in 2001, and the fluorescence of Texas Red-labeled TH59 was observed in the human telomere region of human chromosome.[6] Recently, facile synthetic methods for and functional studies of fluorescent tandem hairpin Py–Im polyamide probes targeting human telomeres were reported.[7] However, given that G•C base pairs were typically recognized by Im/Py pairing, the use of the rare binding motif indicated by boxes in Figure 1 in the tandem hairpin design is intriguing. Indeed, in this probe, the indicated Im groups enclosed by squares would be expected to recognize and bind with C without an antiparallel pairing.

Unclear recognition by tandem motifs has been a concern that has restricted the application of such probes for the analysis of general gene sequences. To address this issue, we synthesized three types of single hairpin Py–Im polyamides **1**–**3** targeting human telomere sequences. We anticipated that by using the new

∗ Corresponding author. Tel.: +81-75-753-4002; fax: +81-75-753-3670; e-mail: bando@kuchem.kyoto-u.ac.jp (T. Bando)

∗ Corresponding author. Tel.: +81-75-753-4002; fax: +81-75-753-3670; e-mail: hs@kuchem.kyoto-u.ac.jp (H. Sugiyama)
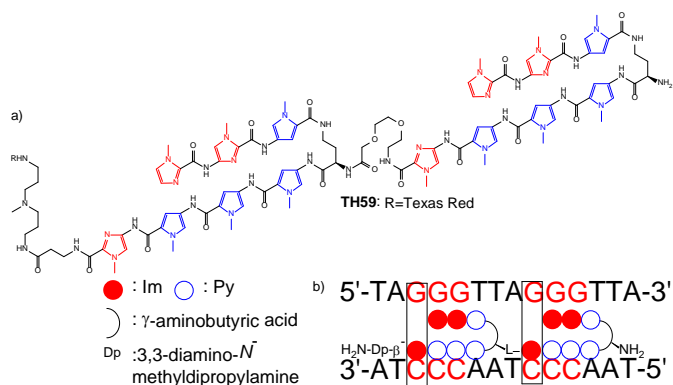
Figure 1. a) Chemical structure of tandem hairpin Py–Im polyamide (TH59). b) Schematic representation of sequence-specific binding in the human telomere sequence by TH59. The two boxes indicate the unusual recognition site by Py/Im pairing.

polyamides, the resulting Py/Im pairing would improve both binding affinities and base pair discrimination. To test this hypothesis, we conducted thermal melting temperature (T_m) measurements to investigate whether these polyamides would bind to target DNA sequences.[8] Surface plasmon resonance (SPR) binding analyses for these polyamides were also investigated to check the effect of Py/Im pairing.[9]

2. Result and discussion

2.1. Synthesis of hairpin Py-Im polyamides *1–3*

We designed three types of Py–Im polyamides (**1–3**; Figure 2) to evaluate their abilities to recognize human telomere sequences. Py–Im polyamide moieties on β-Wang resin were synthesized by using Fmoc solid-phase synthesis and then cleaved with 3,3'-diamino-*N*-methyldipropylamine.[7a,10] The synthesized Py–Im polyamides **1–3** were then purified by reversed-phase HPLC and their identities were confirmed by ESI–TOFMS analysis.

2.2. Binding affinities and specificities of Py–Im polyamides *1–3*

The thermal stabilization of the Py–Im polyamide–DNA complex was estimated by thermal melting temperature (T_m) analysis. By comparing these values for Py–Im polyamides **1–3**, the relative binding affinities and specificities could be evaluated for each matched and mismatched sequences.[7–9]

The T_m measurements were summarized in Table 1. The ΔT_m values obtained from the difference between the T_m of the telomere sequence with polyamides **1–3** and the T_m of the native telomere sequence (ODN-1/2) were 27.8, 27.4, and 27.7 °C, respectively. These ΔT_m values demonstrated that these Py–Im polyamides had good binding affinities that matched those of cyclic Py–Im polyamides.[11] In contrast, the ΔT_m (match) – ΔT_m (mismatch) ($\Delta\Delta T_m$) values of ODN-3/4-**1**, -**2**, and -**3** were –1.5, –4.2, and 2.9 °C, respectively, indicating that ODN-3/4-**1** and -**2** were more stable than ODN-1/2-**1** and -**2**. These results suggested that whereas **1** and **2** may not be useful for discriminating between TTAGGG and TTAAGG, polyamide **3** could discriminate between these telomere sequences. In contrast, the complexes of **1**, **2**, and **3** with ODN-5/6, which also had a 1-bp mismatch (TTAGAG), exhibited good $\Delta\Delta T_m$ values of 5.4, 5.8, and 9.4 °C, respectively. These results suggested that polyamides **1**, **2**, and **3** could discriminate between TTAGGG and TTAGAG, and **3** would likely exhibit good specificity against human telomere sequence, TTAGGG.

2.3 SPR analysis of Py–Im polyamides *1–3*
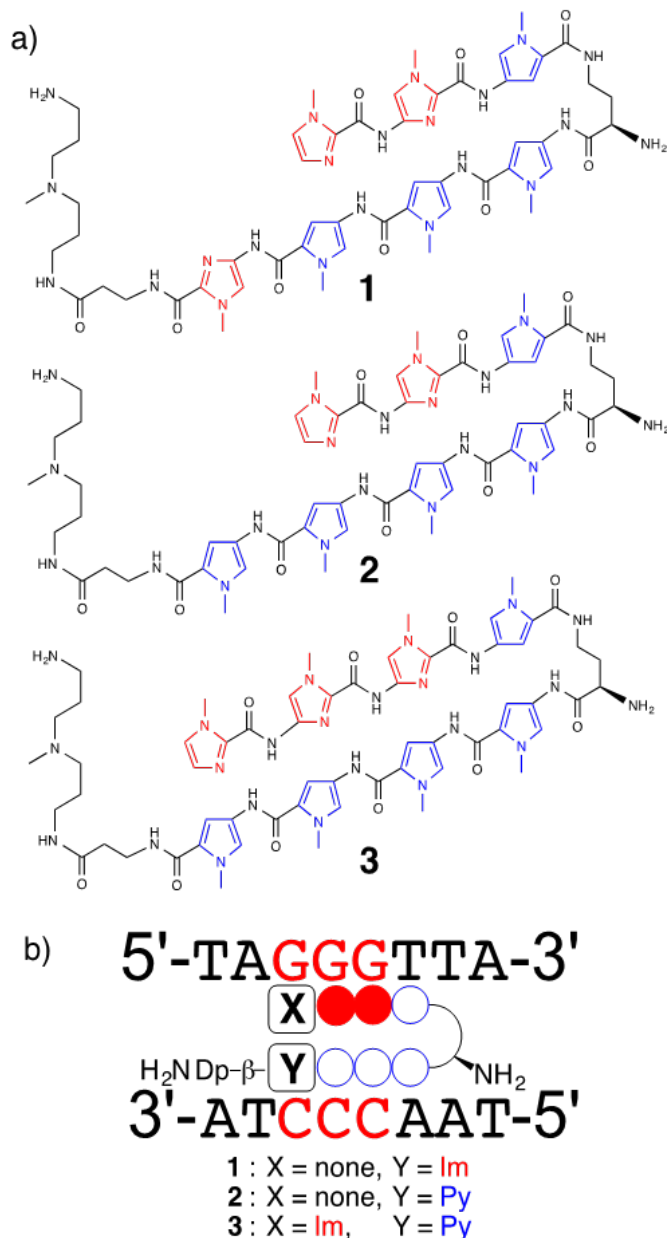
Figure 2. a) Chemical structures of Py–Im hairpin polyamides **1–3** targeting the human telomere sequence. b) Schematic representation of binding by polyamides **1–3**.

Sequence-specific binding of **1–3** was also evaluated by SPR experiments, and the results were compared with the data obtained by T_m analysis. The same three DNA sequences (5'-TTAGGG-3', 5'-TTAAGG-3', and 5'-TTAGAG-3') were used. The SPR profiles and K_D, k_a and k_d values for **1–3** are summarized in Figures 3–5 and Table 2, respectively. The SPR sensorgrams of **1**, the structure of which was based on TH59, are shown in Figure 3. Efficient DNA binding by **1** occurred to the TTAGGG sequence, with a K_D, k_a and k_d values of 1.76×10^{-9} M, 1.89×10^6 M^{-1}s^{-1} and 3.33×10^{-3} s^{-1} being measured (Figure 3a and Table 2). Similarly, the K_D, k_a and k_d values measured for binding to TTAGGG of **2** (4.31×10^{-10} M, 1.03×10^7 M^{-1}s^{-1} and 4.44×10^{-3} s^{-1}) and **3** (1.56×10^{-9} M, 1.32×10^6 M^{-1}s^{-1} and 1.27×10^{-3} s^{-1}) indicated that Py–Im polyamide **2** had the highest association speed and **3** was the slowest one (Figure 4a and 5a and Table 2). As the k_d values, Py–Im polyamide **3** had the lowest dissociation speed, **2** was the fastest polyamide. In summarize, as K_D values shown, all of the **1–3** had good DNA binding affinity against the human telomere sequence (Figures 3a, 4a, 5a and Table 2). Surprisingly, K_D values indicated that binding to TTAAGG by **1**

Table 1. Results of T_m analyses. $\Delta\Delta T_m = \Delta T_m$ (match) – ΔT_m (mismatch).

Py-Im polyamide	human telomere sequence ODN 1: 5'-GGTTAGGGTTAGG-3' ODN 2: 3'-CCAATCCCAATCC-5' T_m = 34.4 °C (±0.1)			1 bp mismatch ODN 3: 5'-GGTTAAGGTTAGG-3' ODN 4: 3'-CCAATTCCAATCC-5' T_m = 30.6 °C (±0.1)			1 bp mismatch ODN 5: 5'-GGTTAGAGTTAGG-3' ODN 6: 3'-CCAATCTCAATCC-5' T_m = 30.4 °C (±0.1)		
	T_m /°C	ΔT_m/°C	$\Delta\Delta T_m$/°C	T_m /°C	ΔT_m/°C	$\Delta\Delta T_m$/°C	T_m/°C	ΔT_m/°C	$\Delta\Delta T_m$/°C
1	62.2±0.1	27.8	-	60.1±0.1	29.3	-1.5	53.0±0.4	22.4	5.4
2	61.5±0.1	27.4	-	62.7±0.1	31.6	-4.2	53.2±0.3	21.6	5.8
3	62.2±0.1	27.7	-	57.1±0.4	24.8	2.9	48.8±0.1	18.3	9.4
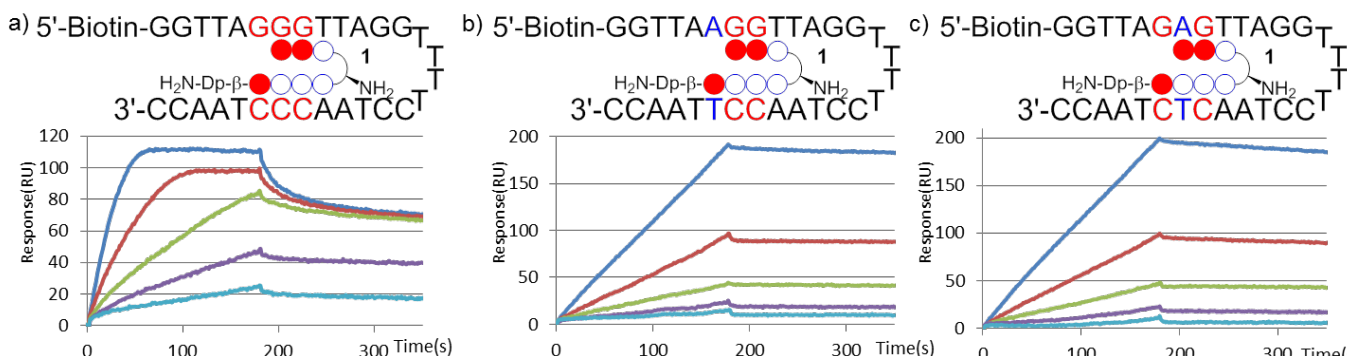
Figure 3. SPR sensorgrams and DNA binding affinities for the interactions of Py-Im polyamide **1** with each DNA sequences. The concentrations of **1** were a) 125 nM (blue), 62.5 nM (red), 31.3 nM (green), 15.6 nM (purple), 7.8 nM (azure); b) and c) 62.5 nM (blue), 31.3 nM (red), 15.6 nM (green), 7.8 nM (purple), 3.9 nM (azure). Specificity is calculated by dividing the K_D of the 1-bp mutated DNA by that of the TTAGGG sequence
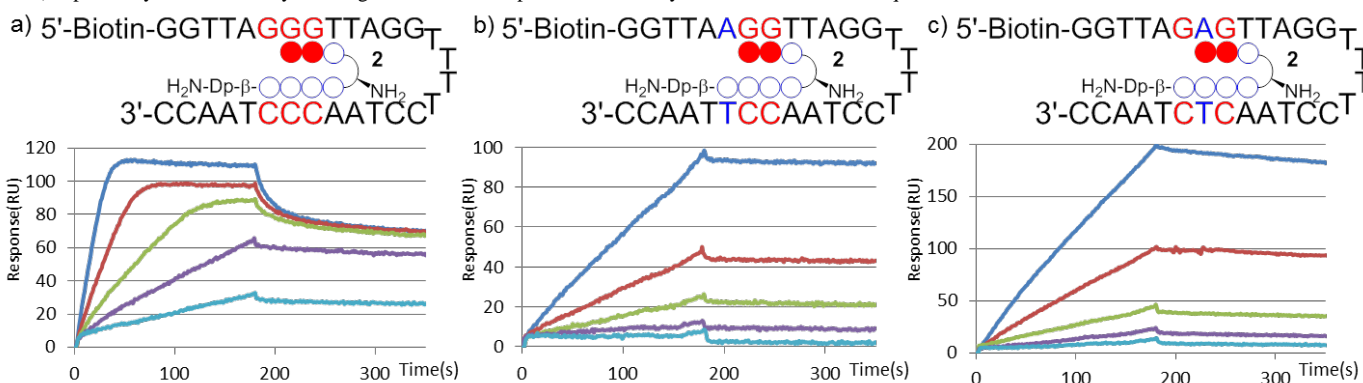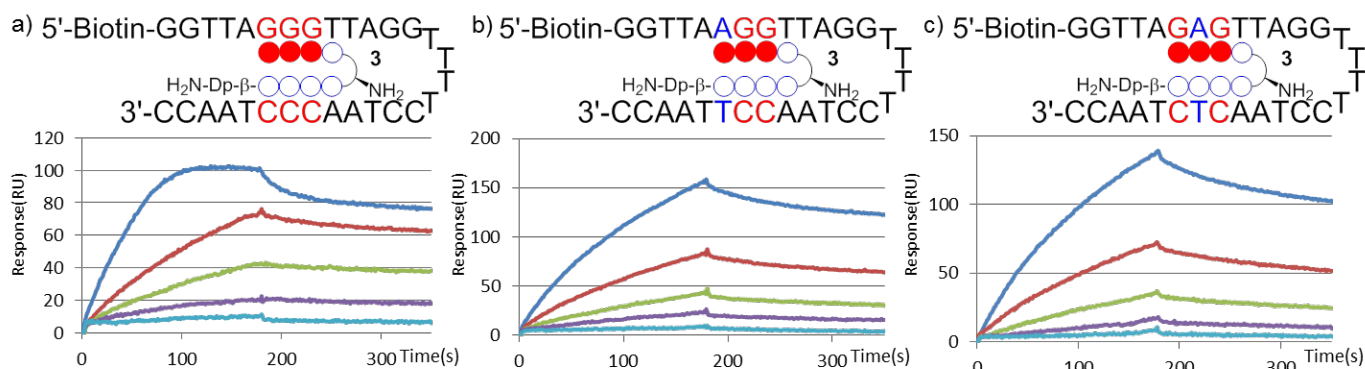

Figure 4. SPR sensorgrams and DNA binding affinities for the interactions of Py–Im polyamide **2** with each DNA sequences. The concentrations of **2** were a) and c) 62.5 nM (blue), 31.3 nM (red), 15.6 nM (green), 7.8 nM (purple), 3.9 nM (azure); b) 31.3 nM (blue), 15.6 nM (red), 7.8 nM (green), 3.9 nM (purple), 2.0 nM (azure). Specificity is calculated by dividing the K_D of the 1-bp mutated DNA by that of the TTAGGG sequence.

(6.27×10^{-10} M) and **2** (5.61×10^{-12} M) were both better than to TTAGGG (Figures 3b, 4b and Table 2). The k_a values of **1-3** to TTAAGG were lower than the values to TTAGGG, which were 4.62×10^5 M^{-1}s^{-1}, 7.61×10^4 M^{-1}s^{-1} and 6.12×10^4 M^{-1}s^{-1}. But polyamides **1** and **2**, whose k_d values were 2.89×10^{-4} s^{-1} and 4.21×10^{-7} s^{-1}, respectively, had lower dissociation speed than TTAGGG. The excepted polyamide was **3** whose k_d value was 1.27×10^{-3} s^{-1} (Figures 3b, 4b, 5b and Table 2). Given that the position of the A·T base pair in the TTAAGG sequence was optimal for **1** and **2** (especially for **2**), the highest DNA binding affinities would be expected for these Py–Im polyamides. In contrast, the K_D value for binding of **3** to TTAGGG (1.56×10^{-9} M) indicated good specificity (13.3) compared with that for the 1 bp mismatch sequence, TTAAGG (2.07×10^{-8} M; Figure 5b and Table 2). In T_m analysis (Table 1), the polyamide **3** also showed higher ΔT_m for TTAGGG (27.7 °C) compared with that for TTAAGG (24.8 °C). As the same of T_m analysis, the 3 kinds of polyamides' affinities to 1 bp mismatch-TTAGAG had been detected. In TTAGAG sequence, all of the polyamides **1-3** had lower association speed (1.06×10^4 M^{-1}s^{-1}, 4.03×10^4 M^{-1}s^{-1} and 4.02×10^4 M^{-1}s^{-1}) but higher dissociation speed was demonstrated only from polyamide **3** (2.05×10^{-3} s^{-1}), other than polyamides **1-2**

(3.24×10^{-4} s^{-1} and 4.27×10^{-4} s^{-1}). As expected, good specificities (17.3, 24.6, and 32.7) by **1–3** to 1-bp mismatch sequence, TTAGAG, were demonstrated by the respective K_D values (**1**: 3.04×10^{-8} M, **2**: 1.06×10^{-8} M, **3**: 5.11×10^{-8} M) (Figures 3c, 4c, 5c and Table 2). These data were consistent with T_m analysis data (Table 1, ODN-5/6).

The fact that the Py–Im polyamide **2** had the strongest affinity to TTAAGG sequence was a problem for targeting human telomere sequence, because the lack of an *N*-terminal Im would lead to stronger DNA binding to the TTAAGG DNA sequence and reduce the specific binding to the telomere sequence. However Py–Im polyamide **1** also showed the specificity (0.36) to TTAAGG, which explained the good human telomere localization in the design of TH59. In the present evaluation, Py–Im polyamide **3** would have good DNA binding affinity to human telomere sequence, TTAGGG, and have good specificity for the human telomere sequence. As the previous research, the ΔT_m value of the telomere sequence with TH59 was 26 °C which was a little lower than polyamide **1**.[7a] For the *SPR* analysis, k_a value of TH59 to 5'-TTAGGGTTAGGGTT-3' is 1.7×10^6 M^{-1}s^{-1}, which was almost same as polyamide **1**. Because the TH59 had more hydrogen bonds with DNA than polyamide **1**,

Figure 5. SPR sensorgrams and DNA binding affinities for the interactions of Py–Im polyamide **3** with each DNA sequences. The concentrations of **3** were a) 62.5 nM (blue), 31.3 nM (red), 15.6 nM (green), 7.8 nM (purple), 3.9 nM (azure); b) and c) 125 nM (blue), 62.5 nM (red), 31.3 nM (green), 15.6 nM (purple), 7.8 nM (azure). Specificity is calculated by dividing the K_D of the 1-bp mutated DNA by that of the TTAGGG sequence.

Table 2. Results of SPR analysis. Specificity is calculated by dividing the K_D of the 1-bp mutated DNA by that of the TTAGGG sequence.

Py-Im polyamide	Human telomere sequence 5'-Biotin-GGTTAGGGTTAGG 3'-CCAATCCCAATCC				1 bp mismatch 5'-Biotin-GGTTAAGGTTAGG 3'-CCAATTCCAATCC				1 bp mismatch 5'-Biotin-GGTTAGAGTTAGG 3'-CCAATCTCAATCC			
	K_D(M)	specificity	k_a(M^{-1}s^{-1})	k_d(s^{-1})	K_D(M)	specificity	k_a(M^{-1}s^{-1})	k_d(s^{-1})	K_D(M)	specificity	k_a(M^{-1}s^{-1})	k_d(s^{-1})
1	1.76×10^{-9}	1	1.89×10^{6}	3.33×10^{-3}	6.27×10^{-10}	0.36	4.62×10^{5}	2.89×10^{-4}	3.04×10^{-8}	17.3	1.06×10^{4}	3.24×10^{-4}
2	4.31×10^{-10}	1	1.03×10^{7}	4.44×10^{-3}	5.61×10^{-12}	0.013	7.61×10^{4}	4.21×10^{-7}	1.06×10^{-8}	24.6	4.03×10^{4}	4.27×10^{-4}
3	1.56×10^{-9}	1	1.32×10^{6}	1.27×10^{-3}	2.07×10^{-8}	13.3	6.12×10^{4}	1.27×10^{-3}	5.11×10^{-8}	32.8	4.02×10^{4}	2.05×10^{-3}

it was more difficult for TH59 to dissociate from DNA, and thus the k_d value of TH59 was 1.6×10^{-3} s^{-1}, which was lower than polyamide **1**. The K_D value for TH59 was 9.3×10^{-10} M, which showed a little better affinity than polyamide **1**.[7b] The polyamide **1** showed the highest affinity in the T_m analyses, although TH59 showed higher affinity than **1** in the SPR measurements, where DNAs were immobilized and influenced in this experiment. The polyamide **2** showed low specificity for the human telomere sequence, but high affinity to 1 bp mismatch TTAAGG, thus evaluation specificity of the tandem hairpin Py-Im polyamides including the hairpin moiety of **2** to human telomere sequence will not be necessary in the future. This kind of tandem hairpin Py-Im polyamides including the hairpin moiety of **2** was expected to have a good specificity to mutant human telomere sequence TTAAGG. The present data suggested that the addition of an N-terminal Im to the Py–Im hairpin polyamide design would increase the specificity against the human telomere sequence. The tandem polyamides including **3** should be supposed to increase the specificity against the human telomere sequence.

3. Conclusion

We have compared three types of hairpin Py–Im polyamides targeting the human telomere sequence. We have demonstrated that the compound (*R*)-Dab (ImImImPy)-PyPyPyPy-β-Dp-NH$_2$ (**3**) has the chemical potential to become a good probe for human telomere staining, particularly based on the discrimination between TTAGGG and TTAAGG. The previously reported synthetic route to TH59 could be reproduced consistently. The utility of the design was expected to be demonstrated by the application to tandem hairpin Py–Im polyamides. Tandem hairpin Py–Im polyamides having an ImImIm moiety such as **3** were expected to be used to target human telomeres in further studies.

4. Experiments

General. Reagents and solvents were purchased from standard suppliers and used without further purification. ^{1}H NMR spectra were recorded on a JEOL JNM ECA-600 spectrometer (600 MHz for ^{1}H), with chemical shifts reported in parts per million

relative to residual solvent and coupling constants in Hertz. The following abbreviations were applied to spin multiplicity: s (singlet), d (doublet), t (triplet), m (multiplet) and dd (doubledoublet). High-performance liquid chromatography (HPLC) analysis and purification were performed with a JASCO PU-2080 Plus HPLC pump, a JASCO 807-IT HPLC UV/VIS detector and a Chemcobond 5-ODS-H reversed phase column (4.6×150 mm) in 0.1% TFA in water with acetonitrile as eluent at a flow rate of 1.0 mL/min, and a linear gradient elution of 0-50% acetonitrile over 20 min with detection at 254 nm. Electrospray ionization time-of-flight mass spectrometry (ESI-TOF-MS) was performed on a BioTOF II (Bruker Daltonics) mass spectrometer using a positive ionization mode. Machine-assisted polyamide syntheses were performed on a PSSM-8 Peptide Synthesizer Simultaneous Multiple (Shimazu) with computer-assisted operation system at a 0.10 mmol scale by using Fmoc chemistry. UV spectra were measured on a NanoDrop ND-1000 spectrophotometer (Thermo Scientific., Inc). All DNA fragments were purchased from Sigma-Aldrich. SPR assays were performed with a Biacore X system (GE Healthcare), and processing of data was carried out by using BIA evaluation, version 4.1. 3,3′-diamino-*N*-methyldipropylamine (DMDPA) was from Tokyo Chemical Industry Co., LTD. Boc-D-Dab(Fmoc)-OH, Fmoc-β-Wang resin (0.55 mmolg^{-1}) and O-(1*H*-6-Chlorobenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) were purchased from Peptide International. Diisopropylethylamine (DIEA) was purchased from Nacalai Tesque, Inc. Fmoc-Py-CO$_2$H, Fmoc-PyIm-CO$_2$H, Fmoc-Im-CO$_2$H, *N,N*-dimethylformamide (DMF), 1-methyl-2-pyrrolidone (NMP), trifluoroacetic acid (TFA), and piperidine were from Wako. 1-methyl-1*H*-imidazole-2-carboxylic acid was from Sigma-Aldrich. Dichloromethane (DCM) was purchased from Sasaki chemical co., Ltd.

4.1. General Procedures of Fmoc Solid-phase Peptide Synthesis

Synthesis of each polyamide was performed on a PSSM-8 (Shimadzu) computer-assisted operation system on a 0.03 mmol scale by using Fmoc Chemistry. An Fmoc building block (0.20 mmol) in each steps were setted up to solve by NMP on the synthetic line. The synthetic procedure of all Py-Im polyamides were as follows; twice deblocking for 4 min with 20%

piperidine/NMP (0.6 mL), activating for 2 min with HCTU (88 mg, 0.21 mmol) in NMP (1 mL) and 10% DIEA/NMP (0.4 mL), coupling for 60 min, and washing with DMF. All coupling were carried out with a single-coupling cycle.

(*R*)-Dab(ImImPy)-PyPyPyIm-β-Dp-NH₂ (1) and (*R*)-Dab(ImImPy)-PyPyPyPy-β-Dp-NH₂ (2). Py-Im polyamides on Wang resin were synthesized in a stepwise reaction with previous Fmoc solid phase protocol.[7a] A sample of resin was cleaved with DMDPA (3 h, 55 °C). The resin was removed by filtration and washed thoroughly with DCM, and the filtrate was concentrated in vacuo. The residue was washed with diethyl ether, triturated with dichloromethane−diethyl ether, and then dried in vacuo. The dried crude residue was purified with HPLC (0.1% TFA/acetonitrile 0-50% linear gradient over 20 min, at 254 nm) to obtain **1** and **2** as a yellow powder.

1: ¹H NMR (600 MHz, DMSO-d₆) δ=10.61 (s, 1H), 10.32 (s, 1H), 10.28 (s, 1H), 9.971 (s, 1H), 9.923 (s, 1H), 9.72 (s, 2H), 8.30 (d, *J*=1.5 Hz, 1H), 8.24 (t, *J*=6.0 Hz, 1H), 8.08(t, *J*=6.1 Hz, 1H), 8.07 (t, *J*=6.0 Hz, 1H), 7.57 (s, 2H), 7.56 (s, 1H), 7.46 (s, 1H), 7.50 (s, 1H), 7.28 (s.1H), 7.26 (s, 1H), 7.25 (s, 1H), 7.22 (s, 1H), 7.14 (s, 1H), 7.09 (s, 1H), 7.07 (s, 1H), 7.05 (d, *J*=1.5 Hz, 1H) 6.95 (s, 1H), 4.01 (s, 3H), 4.00 (s, 3H), 3.94 (s, 3H), 3.87 (s, 3H), 3.86 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.39 (dd, *J*=1.5 Hz, *J*=7.8 Hz, 2H), 3.32-2.29 (m, 4H), 3.13 (t, *J*=5.4 Hz, 1H), 3.07-3.03 (m, 4H), 2.89 (s, 3H), 2.39 (t, *J*=7.2 Hz, 4H), 1.99 (m, 2H), 1.91 (m, 2H), 1.78 (m, 2H), 1.24 ppm (m, 2H) ESI-TOF-MS m/z calcd for $C_{53}H_{70}N_{22}O_9$ [M + H]⁺ 1159.57; found 1159.56; analytical HPLC : t_R=12.9 min (0.1% TFA/CH₃CN 0–50% linear gradient, 0–20 min).

2: ¹H NMR (600 MHz, DMSO-d₆) δ=10.56 (s, 1H), 10.32 (s, 1H), 9.97 (s, 1H), 9.92 (s, 1H), 9.87 (s, 1H), 9.73 (s, 2H), 8.31 (d, *J*=1.5 Hz, 1H), 8.24 (t, *J*=6.0 Hz, 1H), 8.08(t, *J*=6.1 Hz, 1H), 8.03 (t, *J*=6.0Hz, 1H), 7.57 (s, 1H), 7.56 (s, 1H), 7.46 (s, 1H), 7.26 (s, 1H), 7.25 (s.1H), 7.22 (s, 1H), 7.20(s, 1H), 7.15 (s, 1H), 7.09(s, 1H), 7.08 (s, 1H), 7.05 (d, *J*=1.5Hz, 1H), 6.93 (s, 1H) 6.89 (s, 1H), 4.01 (s, 3H), 4.00 (s, 3H), 3.88 (s, 3H), 3.87 (s, 3H), 3.86 (s, 3H), 3.83 (s, 3H), 3.81 (s, 3H), 3.39 (dd, *J*=1.5 Hz, *J*=7.8 Hz, 2H), 3.32-2.28 (m, 4H), 3.12 (t, *J*=5.4 Hz,1H), 3.08-3.03 (m, 4H), 2.87 (s, 3H), 2.36 (t, *J*=7.2 Hz, 4H), 2.00 (m, 2H), 1.91 (m, 2H), 1.78 (m , 2H), 1.24 ppm (m 2H) ESI-TOF-MS m/z calcd for $C_{54}H_{71}N_{21}O_9$ [M+H]⁺ 1158.57; found 1578.58; analytical HPLC: t_R=12.8 min (0.1 % TFA/CH₃CN 0–50 % linear gradient, 0–20 min).

(*R*)-Dab(ImImImPy)-PyPyPyPy-β-Dp-NH₂ (3). Boc-D-Dab(ImImImPy)-PyPyPyPy-β-Wang resin was synthesized in a stepwise reaction by Fmoc solid phase protocol. A sample of resin was cleaved with DMDPA (3 h, 55 °C) and workup procedure was done as described above. The crude residue was treated with TFA-DCM for deprotection and concentrated in vacuo and the residue was dissolved in the minimum amount of DCM, washed with diethyl ether to afford the residue, followed by purification with HPLC (0.1% TFA/acetonitrile 0 - 50% linear gradient over 20 min, at 254 nm) to obtain **3** as a yellow powder.

3: ¹H NMR (600 MHz, DMSO-d₆) δ=10.56 (s, 1H), 10.36 (s, 1H), 10.04 (s, 1H), 9.96 (s, 1H), 9.91 (s, 1H), 9.87 (s, 1H), 9.57 (s, 2H), 8.27 (d, *J*=1.5 Hz, 1H), 8.26 (t, *J*=6.0 Hz, 1H), 8.09 (t, *J*=6.0 Hz, 1H), 8.03 (t, *J*=6.1 Hz, 1H), 7.65 (s, 1H), 7.59 (s, 1H), 7.46 (s, 1H), 7.26 (s, 1H), 7.25 (s.1H), 7.22 (s, 1H), 7.19(s, 1H), 7.14 (s, 1H), 7.10 (s, 1H) 7.09(s, 1H), 7.08 (s, 1H), 7.07 (d, *J*=1.5 Hz, 1H), 6.95 (s, 1H) 6.91 (s, 1H), 4.03 (s, 3H), 4.01 (s, 3H), 4.00 (s, 3H), 3.88 (s, 3H), 3.86 (s, 3H), 3.85 (s, 3H) 3.83 (s, 3H), 3.81 (s, 3H), 3.39 (dd, *J*=1.5 Hz, *J*=7.8 Hz, 2H), 3.32-2.28 (m, 4H), 3.12 (t, 1H), 3.08-3.03 (m, 4H), 2.87 (s, 3H), 2.36 (t, *J*=7.2 Hz,

4H), 2.00 (m, 2H), 1.91 (m, 2H), 1.78 (m, 2H), 1.24 ppm (m, 2H) ESI-TOF-MS m/z calcd for $C_{59}H_{76}N_{24}O_{10}$ [M+H]⁺ 1281.62; found 1281.63; analytical HPLC: t_R=13.4 min (0.1 % TFA/CH₃CN 0–50 % linear gradient, 0–20 min).

4.2. Thermal Denaturation Analysis

Thermal stabilization of polyamide−DNA complex can be analyzed by thermal melting temperature (T_m) analysis, and this method has been used for the measurement of the relative binding affinity and the ability to discriminate mismatch sequences. Polyamide concentrations were calculated with a Nano drop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc.) using an extinction coefficient of 9900 M⁻¹cm⁻¹ per one pyrrole or imidazole moietry at $λ_{max}$ near 310 nm. UV−vis spectra were measured on a spectophotometer V-650 (JASCO). Melting temperature analyses were performed on a spectophotometer V-650 (JASCO) equipped with a thermocontrolled PAC-743R cell changer (JASCO) and a refrigerated and heating circulator F25-ED (Julabo). In order to compare analysis between Py-Im polyamides **1-3**, the relative binding affinity and specificities can be calculated by using match and mismatched sequences. The sequences of dsDNA were 5′-GGTTAGGGTTAGG-3′ (ODN-1) and 3′-CCAATCCCAATCC-5′ (ODN-2), 5′-GGTTAAGGTTAGG-3′ (ODN-3) and 3′-CCAATTCCAATCC-5′(ODN-4), 5′-GGTTAGAGTTAGG-3′ (ODN-5) and 3′-CCAATCTCAATCC-5′ (ODN-6). The underlined bases were the binding sites of polyamides, and the bold bases showed mismatched parts. The analysis buffer was the aqueous solution of 10 mM sodium chloride and 10 mM sodium cacodylate at pH 7.0 containing 2.5% v/v DMF. The concentration of polyamides and dsDNA was 2.5 μM, respectively. Before analyses, samples were annealed from 95 °C to 20 °C at a rate of 1.0 °C/min. Denaturation profiles were recorded at λ = 260 nm from 20 to 95 °C at a rate of 1.0 °C/min.

4.3. SPR Analysis

SPR experiments were performed on a Biacore X instrument. A biotinylated hairpin DNA was purchased from Sigma and the hairpin DNA was shown in Figures 3-5. A streptavidin-functionalized SA sensor chip was purchased from Biacore. The biotinylated DNA is immobilized to the chip to obtain the desired immobilisation level. SPR assays were carried out using HBS-EP buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005% Surfactant P20) with 0.1% DMSO at 25 °C. A series of sample solutions with various concentrations were prepared in the buffer with 0.1% DMSO and injected at a flow rate of 20 μl/min. To measure the values of binding affinity and kinetics parameter, data processing was performed with an appropriate fitting model using BIAevaluation 4.1 program. The predefined models (1:1 binding model with mass transfer) were used for fitting the sensorgrams of polyamides **1**, **2**, and **3** to give better fitting.

Acknowledgments

.

References and notes

1. Blackburn, E. H. *Angew. Chem., Int. Ed.* **2010**, *49*, 7405−7421.
2. Zakian, V. A. *Exp. Cell Res.* **2012**, 318, 1456−1460.

3. Smogorzewska, A.; de Lange, T. *Annu. Rev. Biochem.* **2004**, *73*, 177−208.

4. (a) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235. (b) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Str. Biol.* **2003**, *13*, 284-299. (c) Dervan, P. B.; Doss, R. M.; Marques, M. A. *Curr. Med. Chem.: Anti-Cancer Agents* **2005**, *5*, 373−387. (d) Bando, T.; Sugiyama, H. *Acc. Chem. Res.* **2006**, *39*, 935-944. (e) Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem.* **2013**, *21*, 6101−6114.

5. (a) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468−471. (b) Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 6219−6226. (c) Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382-1391. (d) Swalley, S. E.; Eldon E. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 1113–1120.

6. Maeshima, K.; Janssen, S.; Laemmli, U. K. *EMBO J.* **2001**, *20*, 3218−3228.

7. (a) Kawamoto, Y.; Bando, T.; Kamada, F.; Li, Y.; Hashiya, K.; Maeshima, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2013**, *135*, 16468-16477. (b) Hirata, A.; Nokihara, K.; Kawamoto, Y.; Bando, T.; Sasaki, A.; Ide, S.; Maeshima, K.; Kasama, T.; Sugiyama, H. *J. Am. Chem. Soc.* **2014**, *136*, 11546-11554. (c) Yamamoto, M.; Bando, T.; Kawamoto, Y.; Taylor, R.; Hashiya, K.; Sugiyama, H. *Bioconjugate Chem.* **2014**, *25*, 552-559. (d) Taylor, R, D.; Kawamoto, Y.; Hashiya, K.; Bando, T.; Sugiyama, H. *Chem. Asian J.* **2014**, *9*, 2527-2533.

8. Pilch, D. S.; Poklar, N.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306−8311.

9. Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J .Am. Chem. Soc.* **2002**, *124*, 2153-2163.

10. Wurtz, N. R.; Turner. J. M.; Baird, E. E.; Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201-1203.

11. Li, B. C.; Montgomery, D. C.; Puckett, J. W.; Dervan, P. B. *J. Org. Chem.* **2013**, *78*, 124-133.