JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Targeting 24 bp within Telomere Repeat Sequences with Tandem Tetramer Pyrrole–Imidazole Polyamide Probes

Yusuke Kawamoto,[†] Asuka Sasaki,[‡] Anandhakumar Chandran,[†] Kaori Hashiya,[†] Satoru Ide,[‡] Toshikazu Bando,[*,†] Kazuhiro Maeshima,[*,‡] and Hiroshi Sugiyama[*,†,§]

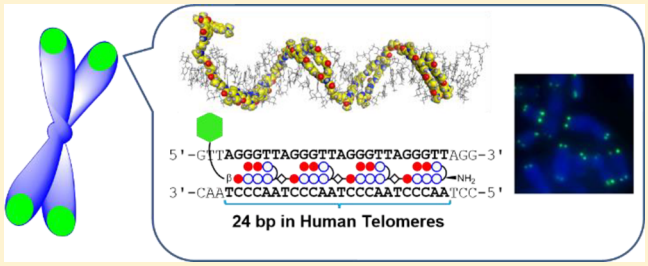
[†]Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan

[‡]Structural Biology Center, National Institute of Genetics, and Department of Genetics, School of Life Science, Graduate University for Advanced Studies (Sokendai), Mishima, Shizuoka 411-8540, Japan

[§]Institute for Integrated Cell-Material Science (WPI-iCeMS), Kyoto University, Sakyo, Kyoto 606-8501, Japan

Supporting Information

ABSTRACT: Synthetic molecules that bind sequence-specifically to DNA have been developed for varied biological applications, including anticancer activity, regulation of gene expression, and visualization of specific genomic regions. Increasing the number of base pairs targeted by synthetic molecules strengthens their sequence specificity. Our group has been working on the development of pyrrole–imidazole polyamides that bind to the minor groove of DNA in a sequence-specific manner without causing denaturation. Recently, we reported a simple synthetic method of fluorescent tandem dimer polyamide probes composed of two hairpin moieties with a linking hinge, which bound to 12 bp in human telomeric repeats (5′-(TTAGGG)$_n$-3′) and could be used to specifically visualize telomeres in chemically fixed cells under mild conditions. We also performed structural optimization and extension of the target base pairs to allow more specific staining of telomeres. In the present study, we synthesized tandem tetramer polyamides composed of four hairpin moieties, targeting 24 bp in telomeric repeats, the longest reported binding site for synthetic, non-nucleic-acid-based, sequence-specific DNA-binding molecules. The novel tandem tetramers bound with a nanomolar dissociation constant to 24 bp sequences made up of four telomeric repeats. Fluorescently labeled tandem tetramer polyamide probes could visualize human telomeres in chemically fixed cells with lower background signals than polyamide probes reported previously, suggesting that they had higher specificity for telomeres. Furthermore, high-throughput sequencing of human genomic DNA pulled down by the biotin-labeled tandem tetramer polyamide probe confirmed its effective binding to telomeric repeats in the complex chromatinized genome.



INTRODUCTION

Sequence-specific DNA-binding molecules have potential as DNA-based therapeutics and diagnostics. Various kinds of programmable, sequence-specific DNA-binding molecules have been developed, including the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated caspase 9 (Cas9) systems,[1] protein-based zinc fingers,[2] transcription activator-like effector (TALE),[3] and nucleic-acid-based triplex-forming oligonucleotides (TFO).[4] Chemically modified nucleic acid analogues such as peptide nucleic acids (PNA) can also function as TFOs.[5] Other synthetic, non-nucleic-acid-based, sequence-specific DNA-binding molecules have been reported, including amidine-benzimidazole-phenyl (ABP)-based minor groove binders,[6] DNA-binding fragments of transcription factors,[7] threading polyintercalators,[8] and minor groove-binding pyrrole–imidazole (Py–Im) polyamides.[9] Our group has been working on Py–Im polyamides first described by Dervan and co-workers. Py–Im polyamides are composed of *N*-methylpyrrole, *N*-methylimidazole, and aliphatic amino acids connected by amide bonds, and they bind sequence-specifically to the minor groove of double-stranded B-DNA by recognizing Watson–Crick base pairs. The binding rule is that antiparallel arrangement of Im/Py can bind to G·C base pairs, whereas Py/Py can bind to both A·T and T·A base pairs.[9f,g] Referred to as aliphatic amino acid residues, the C-terminal β-alanine tail and γ-aminobutyric acid or 2,4-diaminobutyric acid (Dab) turn moiety, called a γ-turn, also binds to A·T and T·A base pairs.[10] The γ-turn connects two aromatic peptides to form hairpin[11] and cyclic[12] structures, and the C-terminal β-alanine tail increases sequence specificity and binding affinity.[13a] In contrast, internal β-alanine can form β/β, β/Py, and β/Im pairs as a substitution for Py rings and can relax the curvature of the polyamides to allow their efficient binding.[13b] Py–Im polyamides can be synthesized easily by machine-assisted solid-phase peptide synthesis (SPPS)[12d,14] and have been used for many purposes; for example, visualization of specific sequences,[15] pull-down and identification of binding regions,[16] inhibition of gene expression,[17] transcriptional

Received: August 29, 2016
Published: October 3, 2016

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100–14107

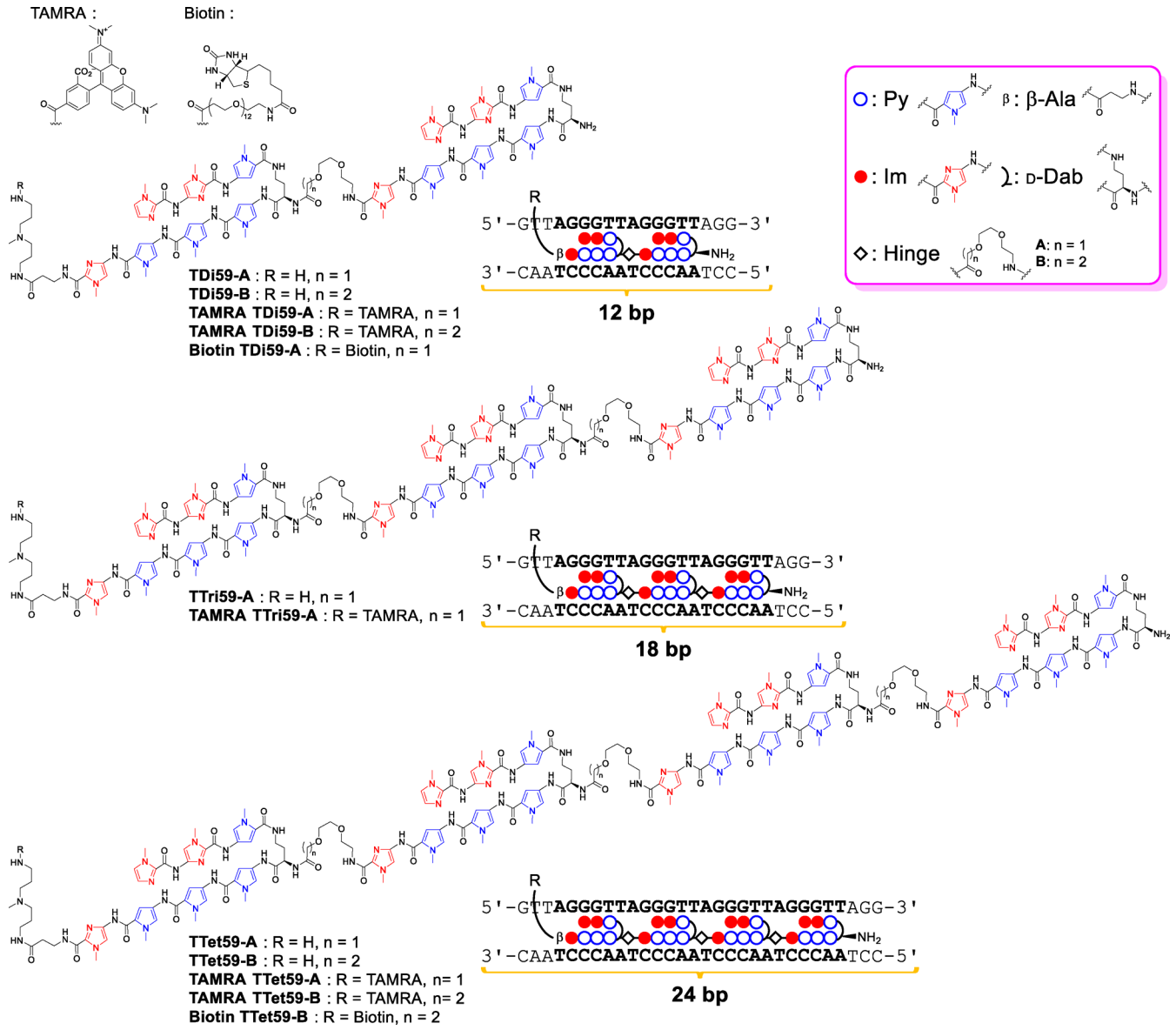


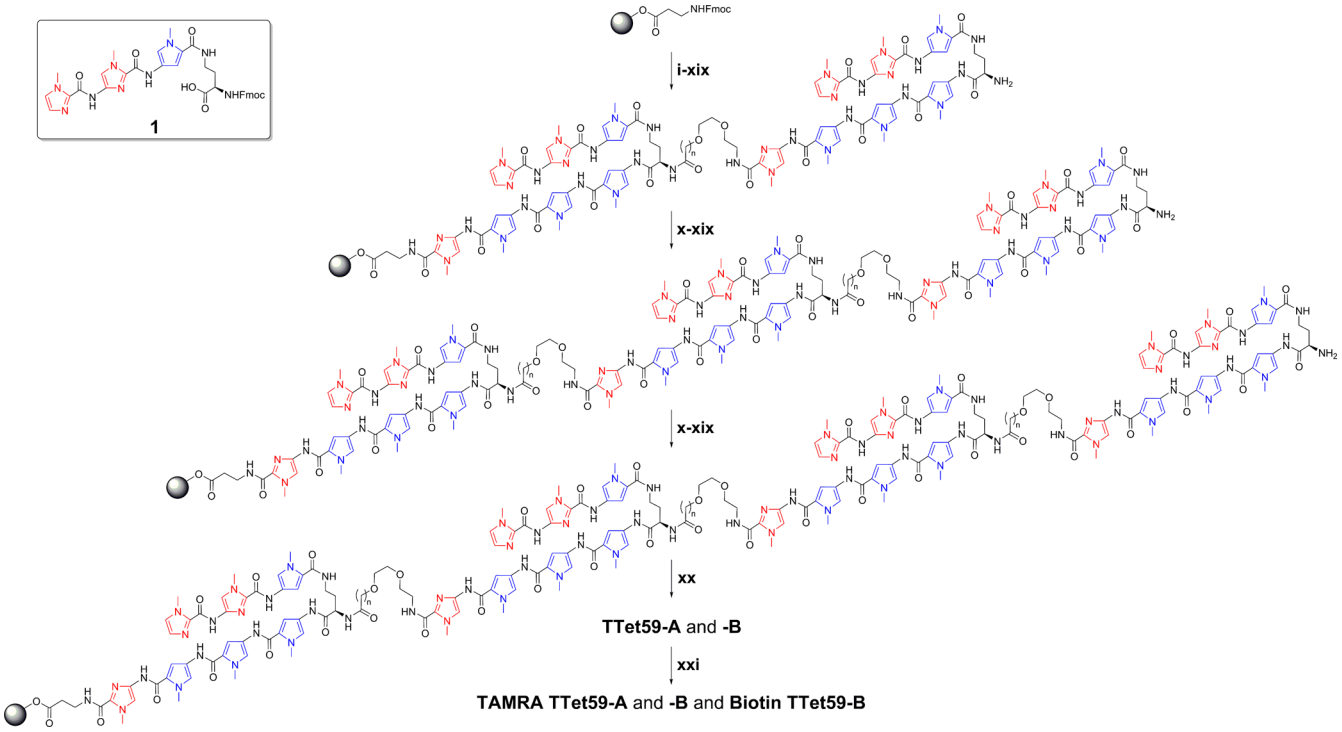
Figure 1. Chemical structures and schematic representations of tandem dimer **TDi59-A** and **-B**, **TAMRA TDi59-A** and **-B** and **Biotin TDi59-A**, tandem trimer **TTri59-A** and **TAMRA TTri59-A**, and tandem tetramer **TTet59-A** and **-B** and **TAMRA TTet59-A** and **-B** and **Biotin TTet59-B** targeting for 12, 18, and 24 bp in human telomere sequence, respectively.

activation,[18,19] and detection of specific sequences on a solid substrate.[20] To target a single site in the whole human genome, which comprises about 3 billion base pairs, a sequence of ≥16 bp must be recognized by DNA-binding molecules; therefore, many efforts have been made to extend the target base pair sequence of such DNA ligands. Wilson and co-workers succeeded in selectively targeting a 10 bp sequence with ABP motifs.[6a] Mascareñas and co-workers have recently synthesized a DNA-binding peptide targeting 13 bp, which is composed of the DNA-binding fragments of two transcription factors connected by a hook.[7] Iverson and co-workers have developed threading polyintercalators, composed of naphthalene diimide intercalators and minor groove and major groove binding aliphatic linkers, which target 14 and 22 bp.[8] This 22 bp sequence was at that time the longest site recognized by a synthetic, non-nucleic-acid-based DNA-binding molecule. Regarding Py−Im polyamides, a linear Py−Im polyamide monomer forming a homodimer with two molecules, and a cysteine-derived Py−Im polyamide dimer were both reported to bind to 16 bp sequences.[21] Furthermore, tandem hairpin motifs composed of several hairpin moieties with connecting hinge segments have also been demonstrated.[22,23]

Recently, our group designed tandem hairpin Py−Im polyamides specific for the human telomere repetitive sequences 5′-$(TTAGGG)_n$-3′.[23] The human telomere is composed of tandem repeats of 5′-TTAGGG-3′ and the protein complex called shelterin including TRF1, TRF2, Rap1, TIN2, TPP1, and POT1. The telomere protects the ends of the chromosomes from nucleolytic degradation and DNA recombination and thus is important in chromosome stability.[24] Human telomere DNA has a duplex region and a single-stranded 3′ overhang, which is suggested to form G-quadruplexes or t-loops.[25] In normal mammalian cells, the number of telomere repeats decreases with cell division and is related to the aging process and cancer.[24a,c,f] Therefore, the telomere length is an important biomarker in studies of these processes, and many techniques to visualize telomeres and to measure telomere length have been reported.[15b] Fluorescence

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100−14107

Scheme 1. Solid-Phase Synthesis of Tandem Tetramer Py−Im Polyamides and the Structure of Building Block 1[a]



[a]Reagents and conditions: (i) 20% piperidine, DMF; (ii) Fmoc-PyIm-CO_2H, HCTU, DIEA, NMP; (iii) 20% piperidine, DMF; (iv) Fmoc-Py-CO_2H, HCTU, DIEA, NMP; (v) 20% piperidine, DMF; (vi) Fmoc-Py-CO_2H, HCTU, DIEA, NMP; (vii) 20% piperidine, DMF; (viii) **1**, HCTU, DIEA, NMP; (ix) 20% piperidine, DMF; (x) Fmoc-mini-PEG or *N*-Fmoc-amido-dPEG$_2$ acid, HCTU, DIEA, NMP; (xi) 20% piperidine, DMF; (xii) Fmoc-PyIm-CO_2H, HCTU, DIEA, NMP; (xiii) 20% piperidine, DMF; (xiv) Fmoc-Py-CO_2H, HCTU, DIEA, NMP; (xv) 20% piperidine, DMF; (xvi) Fmoc-Py-CO_2H, HCTU, DIEA, NMP; (xvii) 20% piperidine, DMF; (xviii) **1**, HCTU, DIEA, NMP; (xix) 20% piperidine, DMF; (xx) 3,3′-diamino-*N*-methyldipropylamine, 45 °C; (xxi) 5-carboxytetramethylrhodamine succinimidyl ester or EZ-Link NHS-PEG$_{12}$-biotin, DIEA, DMF.

in situ hybridization using fluorescent PNA probes has been widely used, but this method normally denatures the DNA under the harsh conditions required for efficient probe binding to telomeres and could lead to destruction of telomere structures.[26] In contrast, Py−Im polyamides can bind to target sequences under mild conditions and thus can visualize target regions without destroying their structures.[23e] Based on a previous report of visualization of telomeres and estimation of their length in chemically fixed cells with Texas Red-labeled Py−Im polyamide probes,[23a] our group developed a simple method for the synthesis of Py−Im polyamide probes labeled with 5-carboxytetramethylrhodamine (TAMRA) that bound to 12 bp of the human telomere sequences. The probes were named tandem dimer **TDi59-A** and **TAMRA TDi59-A** (Figure 1) and were synthesized from a fluorenylmethyloxycarbonyl (Fmoc) building block comprising the sequence from (R)-Dab to the N-terminus of a hairpin moiety.[23b−d] Costaining with the fluorescent Py−Im polyamide probe and anti-TRF1 antibody under mild conditions confirmed that fluorescently labeled **TDi59-A** could specifically visualize human telomeres in chemically fixed cells and that telomere length is related to the amount of shelterin.[23a,b] To decrease the background signals from nonspecific binding in telomere staining, we also developed **TAMRA TDi59-B**, containing hinge B lengthened by one methylene unit,[23c] and tandem trimer **TAMRA TTri59-A**, whose target sequence was extended to 18 bp, the longest specific sequence reported for Py−Im polyamides.[23d] Furthermore, the Py−Im polyamide probes can also visualize telomeres in tissue sections.[23e]

In this study, we have synthesized, using an automatic solid-phase peptide synthesizer, novel tandem tetramer Py−Im polyamides **TTet59-A** and **-B** (Figure 1) targeting 24 bp in the human telomere 5′-(TTAGGG)$_n$-3′ repeats. These Py−Im polyamides are composed of four hairpin units and three hinges, either A or B. To our knowledge, 24 bp is the new record for the longest binding site of synthetic, non-nucleic-acid-based, sequence-specific DNA-binding molecules. Surface plasmon resonance (SPR) analysis was performed to assess the binding affinity to four telomere repeats of the two tandem tetramer Py−Im polyamides and tandem trimer **TTri59-A**. After conjugation of **TTet59-A** and **-B** with TAMRA to obtain **TAMRA TTet59-A** and **-B**, we stained telomeres in chemically fixed cells to compare the new Py−Im polyamide probes with those previously reported and found that the background signals derived from nonspecific binding were much lower for the tetramers than for previously reported Py−Im polyamide probes. Finally, **TDi59-A** and **TTet59-B** were conjugated to biotin, allowing us to assess their binding to telomeric repeats in the whole genome using high-throughput sequencing technology.[16f,g,27]

RESULTS AND DISCUSSION

Synthesis of Py−Im Polyamides. As shown in Scheme 1, tandem tetramer Py−Im polyamides **TTet59-A** and **-B** were synthesized based on the reported methodology, using building block **1** corresponding to Dab and three-ring Py and Im beside the N-terminal.[23b−d] In detail, machine-assisted Fmoc SPPS from Fmoc-*β*-Ala-Wang resin was performed. Fmoc-protected resin was treated with 20% piperidine/*N*,*N*-dimethylformamide

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100−14107

(DMF) for deprotection, followed by coupling with the next Fmoc block (Fmoc-Py-OH, Fmoc-PyIm-OH,[28] **1**,[23b−d] Fmoc-mini-PEG or *N*-Fmoc-amido-dPEG$_2$ acid) activated with *N*,*N*-diisopropylethylamine (DIEA) and *N*,*N*,*N*,*N*-tetramethyl-*O*-(6-chloro-1*H*-benzotriazol-1-yl)uronium hexafluorophosphate (HCTU) in *N*-methyl-2-pyrrolidone (NMP). After these two procedures were repeated 19 times, the N-terminal Fmoc-protected amino group was deprotected with 20% piperidine and then cleaved from the resin with 3,3′-diamino-*N*-methyldipropylamine at 45 °C to produce **TTet59-A** and **-B** with 2.7 and 7.7% yield, respectively. The resulting Py−Im polyamides were coupled with 5-carboxytetramethylrhodamine or biotin succinimidyl ester in DIEA and DMF to produce the fluorescent probes **TAMRA TTet59-A** and **-B** and the biotinylated Py−Im polyamides **Biotin TDi59-A** and **Biotin TTet59-B**. Analytical high-performance liquid chromatography (HPLC) profiles and electrospray ionization time-of-flight mass (ESI-TOF-MS) spectrometry spectra are shown in Figures S1 and S2, respectively.

Comparative Analysis of Binding To Match Sequence. Binding affinities of tetramers **TTet59-A** and **-B** and tandem trimer **TTri59-A** to telomere repeats were assessed with SPR.[23c,d,29] 5′-Biotinylated ODN-1 (5′-biotin-GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTTTCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACC-3′) containing four telomeric repeats was immobilized to a sensor chip through a biotin−streptavidin interaction, and then the trimer and tetramers were passed over the DNA on the sensor chip, and sensorgrams (shown in Figure S3) and values (summarized in Table 1) were

Table 1. Binding Affinities of Polyamides TTet59-A, TTet59-B, and TTri59-A against Match Sequence (ODN-1) Calculated with SPR

	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (M)
5′-Biotin-GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG 3′-CCAATCCCAATCCCAATCCCAATCCCAATCC TTet59-A	5.7 x 10^4	4.3 x 10^{-4}	7.5 x 10^{-9}
5′-Biotin-GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG 3′-CCAATCCCAATCCCAATCCCAATCCCAATCC TTet59-B	4.1 x 10^4	1.6 x 10^{-4}	4.0 x 10^{-9}
5′-Biotin-GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG 3′-CCAATCCCAATCCCAATCCCAATCCCAATCC TTri59-A	7.9 x 10^4	4.8 x 10^{-4}	6.1 x 10^{-9}

obtained. The underlined bases are the binding site of these three types of Py−Im polyamides. Tetramer **TTet59-A** showed strong binding affinity to ODN-1 (K_D = 7.5 nM), comparable to other types of Py−Im polyamides.[9] Trimer **TTri59-A** showed a slightly higher association constant and stronger binding affinity than **TTet59-A** (K_D = 6.1 nM). As suggested previously,[23d] a major reason for this result is that the tetramer's higher steric hindrance influenced its accessibility to the match site. Of the three Py−Im polyamides, **TTet59-B** showed the strongest binding affinity (K_D = 4.0 nM), attributed to a much slower dissociation, suggesting that the hinge region prevented the polyamide's dissociation.

Visualization of Human Telomeres To Compare Py−Im Polyamide Probes. To compare the abilities of Py−Im polyamide probes to stain telomeres specifically, we doubly stained human HeLa 1.3 cell spreads with 4′,6-diamidino-2-phenylindole (DAPI) and five Py−Im polyamide probes: **TAMRA TDi59-A** and **-B**, **TAMRA TTri59-A**, and **TAMRA TTet59-A** and **-B**.[23] Cell spreads are used for clinical karyotyping tests and are prepared from colcemid-treated mitotic cells fixed with MeOH/AcOH. The resulting images of chromosomes stained with

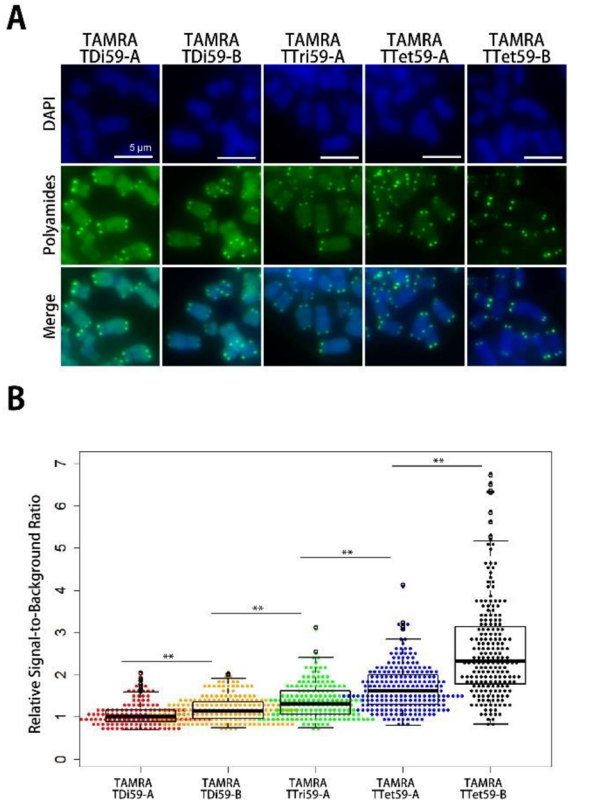


Figure 2. Telomere staining of HeLa 1.3 cell spreads with 75 nM fluorescent Py−Im polyamide probes. (A) Images of HeLa 1.3 cell spreads stained with DAPI (first row) and the fluorescent polyamide probes (second row). The merged images are shown in the third row. Original images are shown in Figure S4A. (B) Signal-to-background ratios of telomeric foci in images. Dot plots and box plots are based on the images shown in Figure S4A (for each group, number of dots = 167−231). All the values were expressed relative to the median value of **TAMRA TDi59-A**. $^{**}P < 0.01$, Wilcoxon rank sum test.

75 nM Py−Im polyamide probes are shown in Figures 2, S4A, and S5A. Chromosomal regions and nuclei were visualized with DAPI. Each Py−Im polyamide probe showed two foci at the ends of every chromosomes, suggesting that all probes, including the new tandem tetramer probes **TAMRA TTet59-A** and **-B**, could stain telomeres specifically. As the number of targeted base pairs increased, the fluorescence of the polyamide probes became lower, presumably because increasing the size of the polyamide probes decreased the number of molecules binding to telomeres, but the background signals derived from nonspecific binding of polyamides decreased. When treating cells with **TAMRA TDi59-A** or **-B** or **TAMRA TTri59-A**, we could easily observe the shape of chromosomes from the background signals of polyamides, and stronger background signals for **TAMRA TDi59-A**, in particular, were detected from the whole chromosomes. However, for treatment with **TAMRA TTet59-A** or **-B**, the intensity of background signals was less, showing that extending the target base pair sequence significantly improves telomere specificity in cells. In addition, **TAMRA TTet59-B** had less background signal along the chromosomes than **TAMRA TTet59-A**, suggesting that the change of hinge from A to B is also effective in improving specificity; this was also shown by the comparison between **TAMRA TDi59-A** and **-B**.[23c] For quantitative evaluation of the telomere staining, we compared relative signal-to-background (S/B) ratios between polyamide probes (Figure 2B). Each polyamide signal at chromosome ends was extracted and

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100−14107

divided by background signal at its neighboring chromosomal region. **TAMRA TTet59-B** had the highest relative S/B ratio. These results demonstrated that increasing the number of target base pairs to 24 and changing the hinge from A to B suppressed the nonspecific binding and allowed us to observe telomeres in HeLa 1.3 cell spreads with high specificity.

Images of telomere staining of HeLa 1.3 cells fixed in formaldehyde (Figures 3, S4B, and S5B), a commonly used fixation method in cell biology, and the quantitative analysis of S/B ratios (Figure 3B) suggested that the results were consistent with those of cell spreads staining. Many sharp foci from the telomeric regions were observed in the nuclei of cells.[23a,b] As reported previously, **TAMRA TDi59-B** and **TAMRA TTri59-A** showed lower background signals than those of **TAMRA TDi59-A**.[23c,d] The background signals of the two tandem tetramer probes **TAMRA TTet59-A** and **-B** were lower and their S/B ratios were higher than those of the other three probes. As can be seen from the polyamide signals of **TAMRA TTet59-A** and **-B** shown in Figure 3, the shapes of the whole nuclei were less distinct. In the images of formaldehyde-fixed cells, **TAMRA TTet59-B** also showed the highest S/B ratio. These results also demonstrated that **TAMRA TTet59-A** and **-B** could visualize telomeres with higher specificity in the fixed cells or formaldehyde-fixed cells and that **TAMRA TTet59-B** targeting 24 bp and composed of four hairpins and three B hinges had the highest telomere specificity of the fluorescent Py–Im polyamide-based telomere probes.

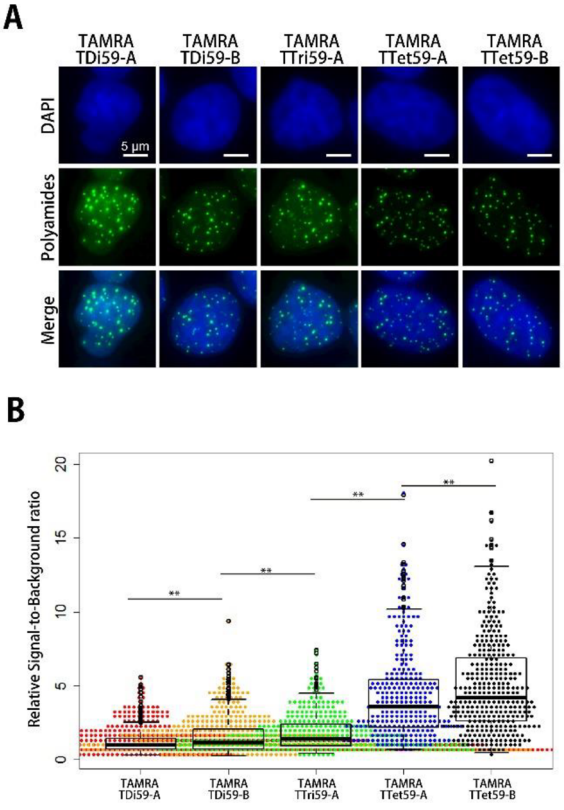


Figure 3. Telomere staining of HeLa 1.3 cells with 15 nM fluorescent Py–Im polyamide probes. (A) Images of HeLa 1.3 cells stained with DAPI (first row) and fluorescent polyamide probes (second row). The merged images are shown in the third row. Original images are shown in Figure S4B. (B) Signal-to-background ratios of telomeric foci in images. Dot plots and box plots are based on the images shown in Figure S4B (for each group, number of dots = 295–606). All the values were expressed relative to the median value of **TAMRA TDi59-A**. $^{**}P < 0.01$, Wilcoxon rank sum test.

High-Throughput Sequencing of the Binding Sites of Py–Im Polyamide Probes for Telomeric Repeats. Images of chromosome spreads treated with the Py–Im polyamide probes (Figure 2) suggested that they specifically stained telomeres. To assess their efficient binding to telomeric repeats in a biologically active, histone-packed, chromatinized genome, we performed high-throughput sequencing of DNA bound to **Biotin TDi59-A** and **Biotin TTet59-B**, based on the reported procedure.[16f,g] The workflow is shown in Scheme 2. In brief, nuclei were extracted from BJ fibroblast cells and then incubated with 400 nM **Biotin TDi59-A** or **Biotin TTet59-B** for 16 h at 4 °C. After this, the extracted nuclei were treated with micrococcal nuclease to produce mononucleosomes to generate the optimally sized DNA for construction of a sequencing library. Nuclear proteins were digested with proteinase K, and then DNA bound to the biotinylated Py–Im polyamide probes was extracted through a biotin–streptavidin interaction on streptavidin-coated magnetic beads (MyOne Streptavidin C1 beads). The extracted and purified DNA was ligated with sequencing platform-specific adapters and amplified with a polymerase chain reaction to obtain sufficient amounts for a sequencing library. The qualified libraries were sequenced using the ion proton sequencing system. Sequenced and filtered reads were aligned to human reference genome hg38. MACS peaks were called to identify the differentially enriched regions compared with control data.[30a]

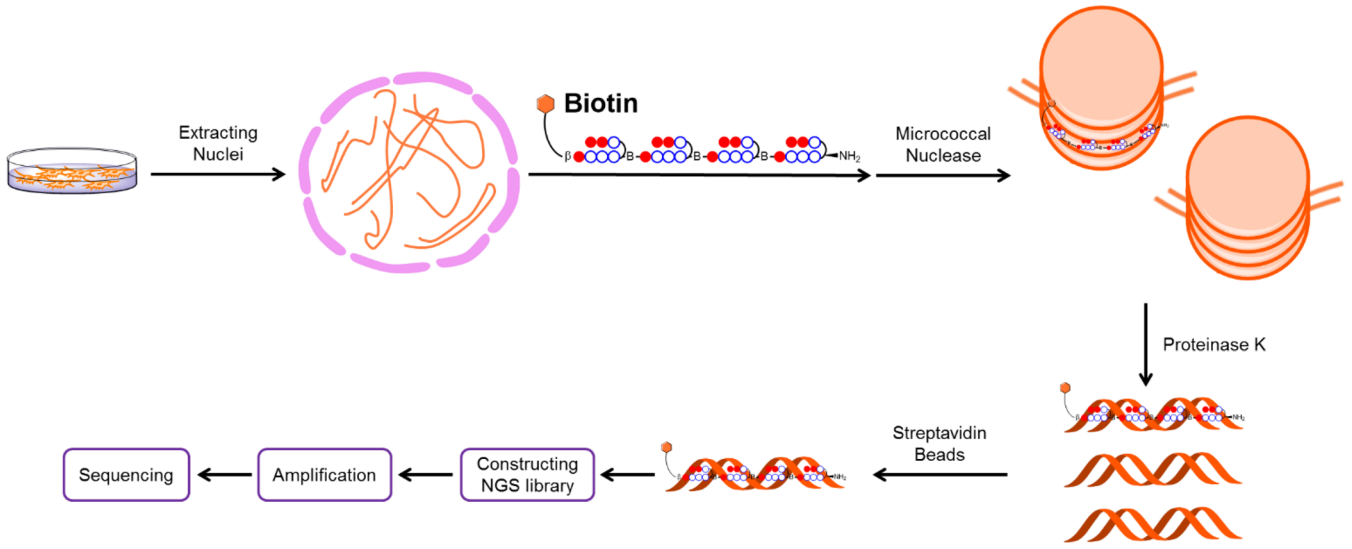
Figure S6 shows a representative result indicating the enriched regions in chromosomes 1 and 18, and a close-up of the enriched telomeric repeats from a nontreated control sample and after **Biotin TDi59-A** and **Biotin TTet59-B** treatment is shown in Figure 4. In both left and right arm termini of chromosomes 1 and 18, we observed significant enrichments with both polyamides compared with the control. The enrichment was consistent with telomeric repeat sequences (TTAGGG/CCCTAA) of the reference genome, suggesting that we could successfully validate the binding of **Biotin TDi59-A** and **Biotin TTet59-B** to telomeric repeats with high-throughput sequencing. Other mild enrichments appeared to be the results of partial or nonspecific binding.

The enrichments in the termini of each chromosome are shown in Figure S7. Duplicate experiments ensured the consistency of enrichments in chromosomal termini. However, we could not quantify the enrichment level on Chr 3R, 5R, 6L, 6R, 8L, 8R, 9R, 11L, 13L, 14L, 14R, 15L, 16R, 17L, 19L, 20L, 20R, 21L, and 22L (L, left arm; R, right arm) chromosomal termini. This may be because of lacking telomeric repeat sequence information on the above-mentioned chromosomes in the human reference genome hg38 data. Even though chromosomes 13, 14, 15, 21, and 22 lack the telomeric repeats information at the left termini presumably because of rDNA whose sequences are not included in the human reference genome hg38, we could observe enrichments from ~16000000 bp, ~16000000 bp, ~17000000 bp, ~5010000 bp, and ~10510000 bp, respectively. These results show that **Biotin TDi59-A** and **Biotin TTet59-B** could effectively recognize and bind to telomeric repeats.

To identify high affinity binding motif of **Biotin TDi59-A** and **Biotin TTet59-B** from the enrichments, we used the Homer motif analysis program with the enriched peaks.[30b] One of the highly enriched motifs corresponded to telomeric 5′-$(TTAGGG)_n$-3′ repeats with a p value of e^{-8595} and e^{-100} for **Biotin TDi59-A** and **Biotin TTet59-B**, respectively (Figure 4B).

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100–14107

Scheme 2. Extraction and High-Throughput Sequencing of Genomic DNA Bound by Biotinylated Py−Im Polyamide Probes



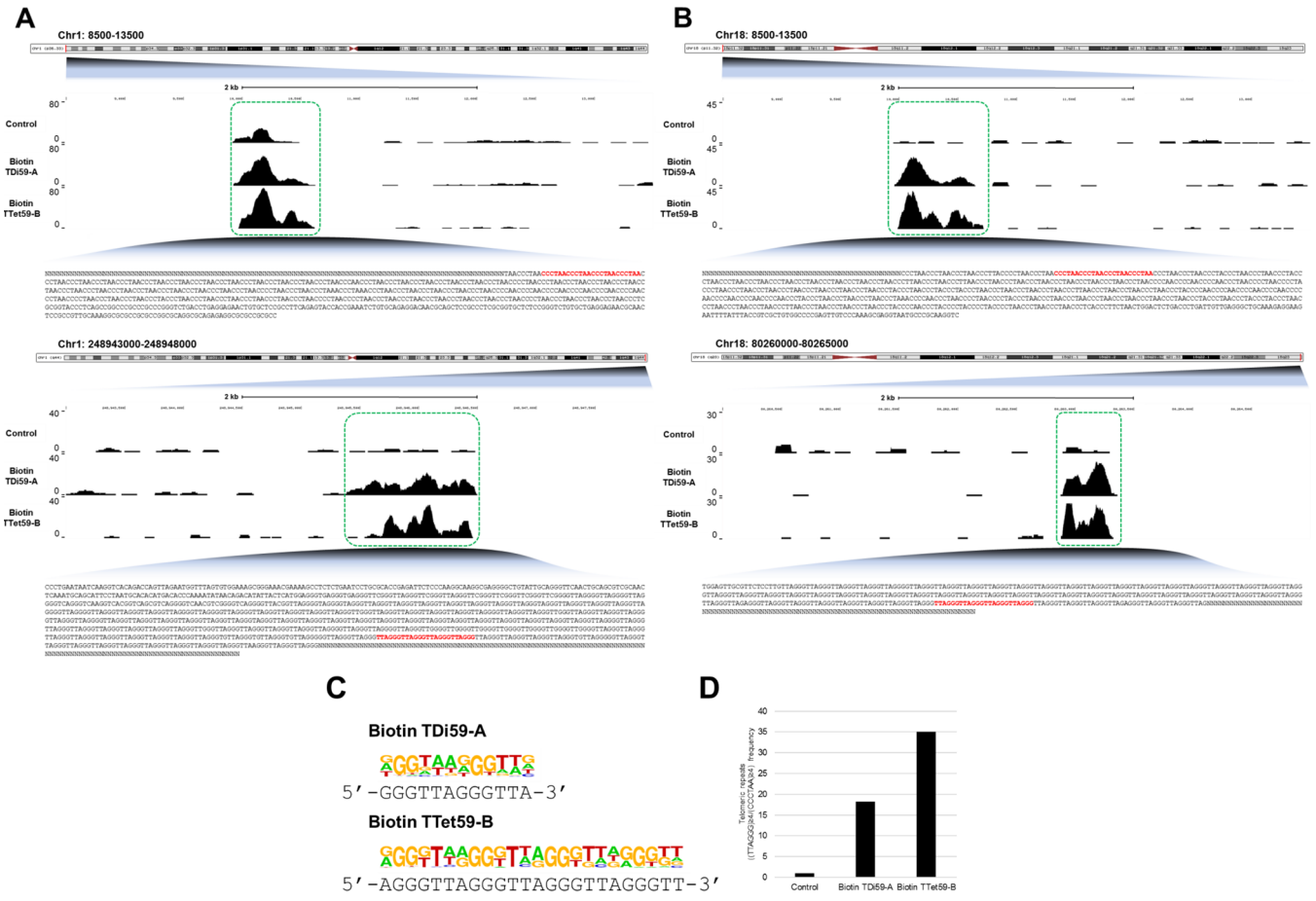


Figure 4. Assessment of efficient binding of **Biotin TDi59-A** and **Biotin TTet59-B** to telomeric repeats with high-throughput sequencing. Mapping of the biotinylated Py−Im polyamide probe binding and enriched sites in both termini of chromosomes 1 and 18 are shown in (A) and (B), respectively. (C) Binding motifs of the probes corresponding to telomeric repeats identified in human genomic enriched sequences. (D) Ratio of sequenced reads containing ≥4 telomeric repeats to whole reads as normalized to that from the control.

This result also suggests efficient binding of **Biotin TDi59-A** and **Biotin TTet59-B** to telomeres.

Finally, to compare the recognition of telomeric repeats by **Biotin TDi59-A** with that by **Biotin TTet59-B**, the ratio of reads containing ≥4 TTAGGG/CCCTAA repeats to whole sequenced reads was calculated. As shown in Figure 4C, the ratio for **Biotin TDi59-A** was 18.2 times higher than that of the nontreated control, showing that **Biotin TDi59-A** could enrich

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100−14107

telomeric DNA. The ratio for **Biotin TTet59-B** was 1.9 times higher than that for **Biotin TDi59-A**, suggesting that **TTet59-B** exhibited better recognition ability toward telomeric repeats. This result is also consistent with the higher telomere-specific staining of **TAMRA TTet59-B**.

High-throughput sequencing confirmed the binding of **Biotin TDi59-A** and **Biotin TTet59-B** to telomeric repeats in the whole genome. In contrast, foci of fluorescently labeled Py−Im polyamide probes were specifically detected at the ends of the chromosomes (Figure 2). The principal reason for this is that accumulation of the Py−Im polyamides on long telomeric repeats in telomere regions was higher than that in nonspecific regions, and therefore, much stronger fluorescence was detected at the termini of chromosomes than from other chromosomal regions.

CONCLUSION

In this study, we have synthesized novel tandem tetramer Py−Im polyamides composed of four hairpins and three hinges that targeted 24 bp of the human telomere sequences. This is a new record for the longest binding site of synthetic, non-nucleic-acid-based, sequence-specific DNA-binding molecules. SPR analysis revealed that tandem tetramers could bind to four telomeric repeats with nanomolar K_D values. TAMRA-labeled tandem tetramer Py−Im polyamide probes could stain clearly the telomere foci in chemically fixed cells. Compared with previously reported tandem dimer and trimer Py−Im polyamides, the tandem tetramers had higher specificity for telomeres in cells, and **TAMRA TTet59-B** containing hinge B had the highest specificity. Furthermore, high-throughput sequencing of chromatin pulled down by biotin-labeled Py−Im polyamide probes confirmed their high recognition and binding toward telomeric DNA. To the best of our knowledge, this is the first report of high-throughput sequencing of DNA pulled down by a synthetic DNA binder targeting human telomeres. To allow the application of these tetramers, our group continues to optimize further their chemical structures and to develop new functions.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b09023.

Materials and methods, mass spectra and HPLC profiles of tandem tetramers, SPR sensorgrams, cell images, and some enrichment data obtained from high-throughput sequencing (PDF)

AUTHOR INFORMATION

Corresponding Authors

*bando@kuchem.kyoto-u.ac.jp
*kmaeshim@nig.ac.jp
*hs@kuchem.kyoto-u.ac.jp

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. T. de Lange (Rockefeller University) for the gift of HeLa 1.3 cells, Dr. S. Sugano (Kyoto University) for helpful discussion about high-throughput sequencing, S. Asamitsu (Kyoto University) for assistance in SPR data analysis, and S. Sato (Kyoto University) for culturing BJ cells. Y.K. is thankful for a JSPS research fellowship for young scientists (DC1). This work was supported by the Basic Science and Platform Technology Program for Innovative Biological Medicine to H.S., JSPS-NSF International Collaborations in Chemistry (ICC) to H.S., NIG Collaborative Research Program (2015-B6) to K.M. and H.S., JST-CREST to K.M., and JSPS KAKENHI Grant Nos. 24225005 to H.S., 24310155 to T.B., and 15J00928 to Y.K.

REFERENCES

(1) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. *Science* **2012**, *337*, 816−821.

(2) Klug, A. *Annu. Rev. Biochem.* **2010**, *79*, 213−231.

(3) Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. *Science* **2009**, *326*, 1509−1512.

(4) (a) Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P. B. *Nucleic Acids Res.* **2008**, *36*, 5123−5138. (b) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645−650.

(5) (a) Nielsen, P. E. *Chem. Biodiversity* **2010**, *7*, 786−804. (b) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497−1500.

(6) (a) Liu, Y.; Chai, Y.; Kumar, A.; Tidwell, R. R.; Boykin, D. W.; Wilson, W. D. *J. Am. Chem. Soc.* **2012**, *134*, 5290−5299. (b) Paul, A.; Nanjunda, R.; Kumar, A.; Laughlin, S.; Nhili, R.; Depauw, S.; Deuser, S. S.; Chai, Y.; Chaudhary, A. S.; David-Cordonnier, M. H.; Boykin, D. W.; Wilson, W. D. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4927−4932.

(7) Rodríguez, J.; Mosquera, J.; García-Fandiño, R.; Vázquez, M. E.; Mascareñas, J. L. *Chem. Sci.* **2016**, *7*, 3298−3303.

(8) (a) Holman, G. G.; Zewail-Foote, M.; Smith, A. R.; Johnson, K. A.; Iverson, B. L. *Nat. Chem.* **2011**, *3*, 875−881. (b) Rhoden Smith, A.; Iverson, B. L. *J. Am. Chem. Soc.* **2013**, *135*, 12783−12789.

(9) (a) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215−2235. (b) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. 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. (f) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559−561. (g) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468−471.

(10) (a) Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382−1391. (b) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 1113−1120.

(11) (a) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983−7988. (b) de Clairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909−7916.

(12) (a) Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 1121−1129. (b) Chenoweth, D. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13175−13179. (c) Chenoweth, D. M.; Dervan, P. B. *J. Am. Chem. Soc.* **2010**, *132*, 14521−14529. (d) Morinaga, H.; Bando, T.; Takagaki, T.; Yamamoto, H.; Hashiya, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2011**, *133*, 18924−18930. (e) Li, B. C.; Montgomery, D. C.; Puckett, J. W.; Dervan, P. B. *J. Org. Chem.* **2013**, *78*, 124−133.

(13) (a) Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147−6152. (b) Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 6219−6226.

(14) (a) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141−6146. (b) Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201−1203. (c) Fang, L.; Yao, G.; Pan, Z.; Wu, C.; Wang, H. S.; Burley, G. A.; Su, W. *Org. Lett.* **2015**, *17*, 158−161.

(15) (a) Vaijayanthi, T.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2012**, *13*, 2170−2185. (b) Boutorine, A. S.; Novopashina, D. S.; Krasheninina, O. A.; Nozeret, K.; Venyaminova, A. G. *Molecules* **2013**, *18*, 15357−15397. (c) Nozeret, K.; Loll, F.; Escudé, C.; Boutorine, A. S. *ChemBioChem* **2015**, *16*, 549−554.

(16) (a) Anandhakumar, C.; Kizaki, S.; Bando, T.; Pandian, G. N.; Sugiyama, H. *ChemBioChem* **2015**, *16*, 20−38. (b) Meier, J. L.; Yu, A. S.; Korf, I.; Segal, D. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2012**, *134*, 17814−17822. (c) Kang, J. S.; Meier, J. L.; Dervan, P. B. *J. Am. Chem. Soc.* **2014**, *136*, 3687−3694. (d) Erwin, G. S.; Bhimsaria, D.; Eguchi, A.; Ansari, A.

Z. *Angew. Chem., Int. Ed.* **2014**, *53*, 10124−10128. (e) Anandhakumar, C.; Li, Y.; Kizaki, S.; Pandian, G. N.; Hashiya, K.; Bando, T.; Sugiyama, H. *ChemBioChem* **2014**, *15*, 2647−2651. (f) Chandran, A.; Syed, J.; Taylor, R. D.; Kashiwazaki, G.; Sato, S.; Hashiya, K.; Bando, T.; Sugiyama, H. *Nucleic Acids Res.* **2016**, *44*, 4014−4024. (g) Chandran, A.; Syed, J.; Li, Y.; Sato, S.; Bando, T.; Sugiyama, H. *ChemBioChem* **2016**, DOI: 10.1002/cbic.201600274.

(17) (a) Murty, M. S. R. C.; Sugiyama, H. *Biol. Pharm. Bull.* **2004**, *27*, 468−474. (b) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202−205. (c) Lai, Y.−M.; Fukuda, N.; Ueno, T.; Kishioka, H.; Matsuda, H.; Saito, S.; Matsumoto, K.; Ayame, H.; Bando, T.; Sugiyama, H.; Mugishima, H.; Serie, K. *J. Pharmacol. Exp. Ther.* **2005**, *315*, 571−575. (d) Hiraoka, K.; Inoue, T.; Taylor, R. D.; Watanabe, T.; Koshikawa, N.; Yoda, H.; Shinohara, K.; Takatori, A.; Sugimoto, K.; Maru, Y.; Denda, T.; Fujiwara, K.; Balmain, A.; Ozaki, T.; Bando, T.; Sugiyama, H.; Nagase, H. *Nat. Commun.* **2015**, *6*, 6706.

(18) (a) Xiao, X.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew. Chem., Int. Ed.* **2007**, *46*, 2865−2868. (b) Patel, S.; Jung, D.; Yin, P.; Carlton, P.; Yamamoto, M.; Bando, T.; Sugiyama, H.; Lee, K. B. *ACS Nano* **2014**, *8*, 8959−8967.

(19) (a) Han, L.; Pandian, G. N.; Junetha, S.; Sato, S.; Chandran, A.; Taniguchi, J.; Saha, A.; Bando, T.; Nagase, H.; Sugiyama, H. *Angew. Chem., Int. Ed.* **2013**, *52*, 13410−13413. (b) Pandian, G. N.; Taniguchi, J.; Junetha, S.; Sato, S.; Han, L.; Saha, A.; Chandran, A.; Bando, T.; Nagase, H.; Vaijayanthi, T.; Taylor, R. D.; Sugiyama, H. *Sci. Rep.* **2014**, *4*, 3843. (c) Han, L.; Pandian, G. N.; Chandran, A.; Sato, S.; Taniguchi, J.; Kashiwazaki, G.; Sawatani, Y.; Hashiya, K.; Bando, T.; Xu, Y.; Qian, X.; Sugiyama, H. *Angew. Chem., Int. Ed.* **2015**, *54*, 8700−8703.

(20) Singh, I.; Wendeln, C.; Clark, A. W.; Cooper, J. M.; Ravoo, B. J.; Burley, G. A. *J. Am. Chem. Soc.* **2013**, *135*, 3449−3457.

(21) (a) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 3534−3535. (b) Yamamoto, M.; Bando, T.; Morinaga, N.; Kawamoto, Y.; Hashiya, K.; Sugiyama, H. *Chem. - Eur. J.* **2014**, *20*, 752−759.

(22) (a) Herman, D. M.; Baird, E. E.; Dervan, P. B. *Chem. - Eur. J.* **1999**, *5*, 975−983. (b) Kers, I.; Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3339−3349. (c) Schaal, T. D.; Mallet, W. G.; McMinn, D. L.; Nguyen, N. V.; Sopko, M. M.; John, S.; Parekh, B. S. *Nucleic Acids Res.* **2003**, *31*, 1282−1291. (d) Sasaki, S.; Bando, T.; Minoshima, M.; Shinohara, K.; Sugiyama, H. *Chem. - Eur. J.* **2008**, *14*, 864−870. (e) Yamamoto, M.; Bando, T.; Kawamoto, Y.; Taylor, R.; Hashiya, K.; Sugiyama, H. *Bioconjugate Chem.* **2014**, *25*, 552−559. (f) Taylor, R. D.; Kawamoto, Y.; Hashiya, K.; Bando, T.; Sugiyama, H. *Chem. - Asian J.* **2014**, *9*, 2527−2533.

(23) (a) Maeshima, K.; Janssen, S.; Laemmli, U. K. *EMBO J.* **2001**, *20*, 3218−3228. (b) Kawamoto, Y.; Bando, T.; Kamada, F.; Li, Y.; Hashiya, K.; Maeshima, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2013**, *135*, 16468−16477. (c) 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. (d) Kawamoto, Y.; Sasaki, A.; Hashiya, K.; Ide, S.; Bando, T.; Maeshima, K.; Sugiyama, H. *Chem. Sci.* **2015**, *6*, 2307−2312. (e) Sasaki, A.; Ide, S.; Kawamoto, Y.; Bando, T.; Murata, Y.; Shimura, M.; Yamada, K.; Hirata, A.; Nokihara, K.; Hirata, T.; Sugiyama, H.; Maeshima, K. *Sci. Rep.* **2016**, *6*, 29261.

(24) (a) Smogorzewska, A.; de Lange, T. *Annu. Rev. Biochem.* **2004**, *73*, 177−208. (b) Palm, W.; de Lange, T. *Annu. Rev. Genet.* **2008**, *42*, 301−334. (c) Blackburn, E. H. *Angew. Chem., Int. Ed.* **2010**, *49*, 7405−7421. (d) Xu, Y. *Chem. Soc. Rev.* **2011**, *40*, 2719−2740. (e) Nandakumar, J.; Cech, T. R. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 69−82. (f) Zakian, V. A. *Exp. Cell Res.* **2012**, *318*, 1456−1460. (g) Smogorzewska, A.; van Steensel, B.; Bianchi, A.; Oelmann, S.; Schaefer, M. R.; Schnapp, G.; de Lange, T. *Mol. Cell. Biol.* **2000**, *20*, 1659−1668.

(25) (a) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. *Nat. Chem.* **2013**, *5*, 182−186. (b) Doksani, Y.; Wu, J. Y.; de Lange, T.; Zhuang, X. *Cell* **2013**, *155*, 345−356.

(26) Lansdorp, P. M.; Verwoerd, N. P.; van de Rijke, F. M.; Dragowska, V.; Little, M. T.; Dirks, R. W.; Raap, A. K.; Tanke, H. J. *Hum. Mol. Genet.* **1996**, *5*, 685−691.

(27) Anders, L.; Guenther, M. G.; Qi, J.; Fan, Z. P.; Marineau, J. J.; Rahl, P. B.; Lovén, J.; Sigova, A. A.; Smith, W. B.; Lee, T. I.; Bradner, J. E.; Young, R. A. *Nat. Biotechnol.* **2014**, *32*, 92−96.

(28) (a) Minoshima, M.; Bando, T.; Sasaki, S.; Fujimoto, J.; Sugiyama, H. *Nucleic Acids Res.* **2008**, *36*, 2889−2894. (b) Wetzler, M.; Wemmer, D. E. *Org. Lett.* **2010**, *12*, 3488−3490.

(29) (a) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153−2163. (b) Wang, S.; Aston, K.; Koeller, K. J.; Harris, G. D.; Rath, N. P.; Bashkin, J. K.; Wilson, W. D. *Org. Biomol. Chem.* **2014**, *12*, 7523−7536.

(30) (a) Feng, J.; Liu, T.; Qin, B.; Zhang, Y.; Liu, X. S. *Nat. Protoc.* **2012**, *7*, 1728−1740. (b) Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y. C.; Laslo, P.; Cheng, J. X.; Murre, C.; Singh, H.; Glass, C. K. *Mol. Cell* **2010**, *38*, 576−589.

DOI: 10.1021/jacs.6b09023
J. Am. Chem. Soc. 2016, 138, 14100−14107