KURENAI

Kyoto University Research Information Repository

KYOTO UNIVERSITY

Title	Sequence-specific DNA alkylation and transcriptional inhibition by long-chain hairpin pyrrole-imidazole polyamide-chlorambucil conjugates targeting CAG/CTG trinucleotide repeats.
Author(s)	Asamitsu, Sefan; Kawamoto, Yusuke; Hashiya, Fumitaka; Hashiya, Kaori; Yamamoto, Makoto; Kizaki, Seiichiro; Bando, Toshikazu; Sugiyama, Hiroshi
Citation	Bioorganic & medicinal chemistry (2014), 22(17): 4646-4657
Issue Date	2014-09-01
URL	http://hdl.handle.net/2433/189821
Right	© 2014 Elsevier Ltd.
Type	Journal Article
Textversion	author

Kyoto University

Graphical Abstract

To create your abstract, type over the instructions in the template box below.
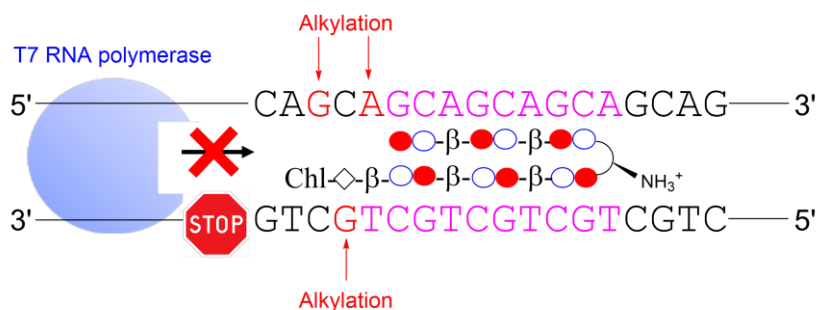Fonts or abstract dimensions should not be changed or altered.

Sequence-specific DNA alkylation and transcriptional inhibition by long-chain hairpin pyrrole–imidazole polyamide–chlorambucil conjugates targeting CAG/CTG trinucleotide repeats

Leave this area blank for abstract info.

Sefan Asamitsu[a], Yusuke Kawamoto[a], Fumitaka Hashiya[a], Kaori Hashiya[a], Makoto Yamamoto[a], Seiichiro Kizaki[a], Toshikazu Bando*[,a], and Hiroshi Sugiyama*[, a, b, c]

[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 606-8502, Japan
[c]CREST, Japan Science and Technology Corporation (JST), Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan

Sequence-specific DNA alkylation and transcriptional inhibition by long-chain hairpin pyrrole–imidazole polyamide–chlorambucil conjugates targeting CAG/CTG trinucleotide repeats

Sefan Asamitsu[a], Yusuke Kawamoto[a], Fumitaka Hashiya[a], Kaori Hashiya[a], Makoto Yamamoto[a], Seiichiro Kizaki[a], Toshikazu Bando∗,[a], and Hiroshi Sugiyama∗,[a, b, c]

[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 606-8502, Japan
[c]CREST, Japan Science and Technology Corporation (JST), Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan

ARTICLE INFO

ABSTRACT

Introducing novel building blocks to solid-phase peptide synthesis, we readily synthesized long-chain hairpin pyrrole-imidazole (PI) polyamide–chlorambucil conjugates **3** and **4** via the introduction of an amino group into a GABA (γ-turn) contained in **3**, to target CAG/CTG repeat sequences, which are associated with various hereditary disorders. A high-resolution denaturing polyacrylamide sequencing gel revealed sequence-specific alkylation both strands at the N3 of adenines or guanines in CAG/CTG repeats by conjugates **3** and **4**, with 11 bp recognition. *In vitro* transcription assays using conjugate **4** revealed that specific alkylation inhibited the progression of RNA polymerase at the alkylating sites. Chiral substitution of the γ-turn with an amino group resulted in higher binding affinity observed in SPR assays. These assays suggest that conjugates **4** with 11 bp recognition has the potential to cause specific DNA damage and transcriptional inhibition at the alkylating sites.

.

1. Introduction

Pyrrole–imidazole polyamides (PI polyamides) are ligands that bind to the DNA minor groove with high affinity and sequence specificity. These contain *N*-methylpyrrole (P) and *N*-methylimidazole (I), and recognize each of the four Watson–Crick base pairs uniquely.[1–3] The binding rule of PI polyamides, termed Dervan's rule, is the recognition of a G/C base pair by antiparallel pairing of an I/P, whereas a P/P pair recognizes an A/T or a T/A base pair. A β-alanine/β-alanine (β/β) pair reads an A/T or a T/A base pair in the same way as a P/P pair does. Recently, the introduction of PI polyamides into DNA-alkylating agents has attracted attention because of their sequence-specific alkylation at target sites.[4]

Trinucleotide repeat sequences exist in some regions of genomic DNA, and the expansion of the repeat sequences often causes neurological disorders.[5,6] For example, the expansion of CAG trinucleotide repeats within the first exon of the Huntingtin (HTT) gene (from 36 to > 100 repeats) yields a full-length HTT protein harboring expanded polyglutamine tracts, which gain some

functions, lead to the selective disruption of neuronal protein function, and ultimately, cause Huntington's disease.[7,8,9] In myotonic dystrophy type 1 (DM1) and Huntington's disease-like 2, which are associated with CTG repeat sequences located in the 3′ UTR of genes, transcriptional products containing excess CUG repeats accumulate within the nucleus and form a 1 bp mismatch-containing hairpin structure, bind to proteins such as splicing regulators, cause inactivation of their alternative pre-mRNA splicing, and lead to the pathogenesis of each of these diseases.[10–13] Surprisingly, these trinucleotide repeat sequences have a tendency to expand as a result of the formation of a hairpin-like structure containing a d(CAG/CAG) or d(CTG/CTG) 1 bp mismatch internal loop motif, by reannealing during DNA replication, repair, and recombination. The expansion leads to a decreased age of onset and increased severity of the diseases in successive generations.[5] Several mechanisms to explain this characteristic inheritance have been proposed; however, a genuine one remains unknown. Recently, studies on antisense oligonucleotides that target the internal loop motifs located in the hairpin structure have been reported.[13] Moreover, Disney and coworkers reported previously that covalent adduct formation
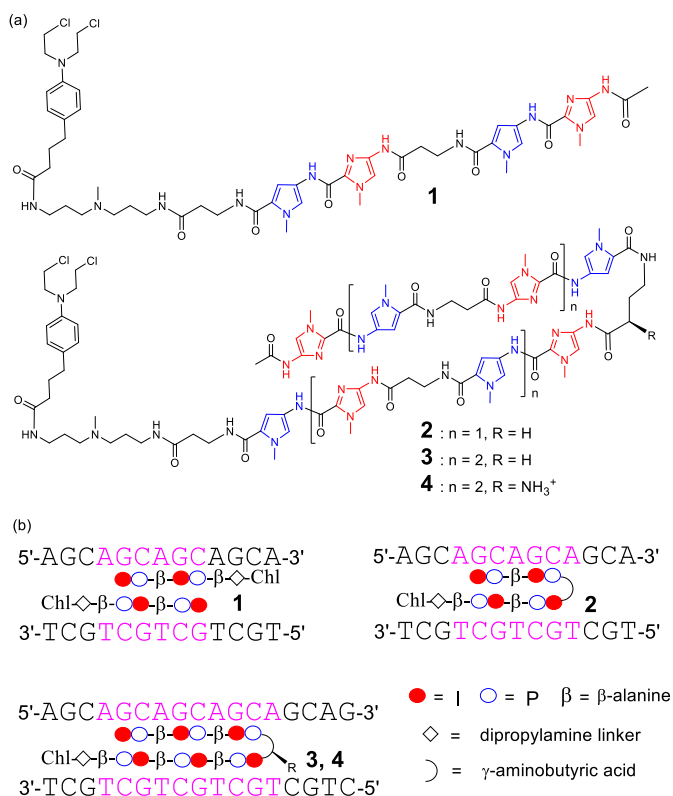
Figure 1. A (a) Chemical structure of PI polyamide–chlorambucil (Chl) conjugates (**1–4**) targeting the CAG/CTG repeat sequences (b) Schematic representation of the binding models of conjugates **1–4** to the CAG/CTG repeat region via a dimer (conjugate **1**) or hairpin (conjugate **2–4**) conformation.

between expanded rCUG transcripts in the hairpin structure and moieties that target the r(CUG/CUG) mismatch internal loop motif improved DM1-associated pre-mRNA splicing defects.[14] In addition, Nakatani *et al.* developed a small molecule that binds a hairpin d(CAG/CAG) region and flips cytidine nucleotides out.[15] Despite those studies, to our knowledge, there are few reports of the targeting of Watson–Crick base pairs in CAG/CTG repeat regions, which are present even in disease-associated expanded CAG/CTG repeats. Our laboratory has developed polyamides with fluorescence dyes that detect CAG repeated sequences as a diagnostic tool.[16,17]

Based on these reports, we hypothesized that the inhibition of transcripts at CAG/CTG repeat sequences that form Watson–Crick base pairs, via specific covalent adduction of an alkylating agent possessing polyamides that recognizes CAG/CTG repeat sequences may prevent the subsequent occurrence of expanded polyglutamine tracts and the sequestration of expanded rCUG transcripts.[18] Moreover, Sinden and coworkers reported that DNA damaging agents, such as alkylating agents, UV irradiation, and oxidation, increase the rate of triplet repeat deletion.[19,20] These are effective approaches for genetic disorders, but have a very limited DNA sequence specificity, which must be related to side effects. To suppress it, we also hypothesized that sequence-specific alkylation damage may lead to an increase in the rate of d(CAG/CTG) repeat deletion without damaging other sequences. Our laboratory reported previously that polyamides covalently linked to DNA-alkylating agents predominantly alkylated their targeting sequences that are associated with cancer (5′–(GGGTTA)n–3′, human telomere repeat sequences,[21–23] or 5′–ACGTCACCA–3′ in KRAS codon 13 mutation[24]). In the present study, we synthesized PI polyamides conjugated to a DNA-alkylating agent, chlorambucil (Chl), to target CAG/CTG repeat

sequences, and their length of sequence recognition was elongated by introducing novel building blocks used in Fmoc solid-phase peptide synthesis (SPPS).[25] The activities of the compounds obtained were investigated in terms of DNA sequence-specific alkylation, transcriptional inhibition, and dynamics of the interaction with the target sequence.

2. Results and discussion

2.1. Molecular design and synthesis

We rationally designed three types of PI polyamide–Chl conjugates (**1–4**) to target CAG/CTG repeat sequences with different lengths or structures (Figure 1).[26,27] Chl is a nitrogen mustard alkylating agent that belongs to the group of anticancer drugs and has been used widely as a first line of treatment for chronic lymphocytic leukemia. Its efficacy has been confirmed at the clinical level.[28,29] The potency of hairpin PI polyamide–Chl conjugates has been investigated in terms of chemical and biological activity.[30,31] Hartley and coworkers previously demonstrated preferential alkylation at the N3 of guanine or adenine on 5′–TTTTGG–3′ or 5′–TTTTGA–3′ sequences, respectively, by tripyrrole conjugated with BAM,[32] which is one of the nitrogen mustards. In addition, it has been reported that Chl is more cytotoxic than BAM in the human colonic adenocarcinoma LS174T and leukemic K562 cell lines.[33]

From these reports, we inferred that hairpin PI polyamide–Chl conjugates, the specific alkylation of which at the N3 of guanines has never been investigated, were appropriate for targeting AG-rich sequences as alkylating agents. Moreover, the flexibility of Chl within the structure may allow reaching the opposite strand.[34] To promote this function and the alkylation at guanines or adenines of both strands by Chl moieties, β-alanine and subsequently 3,3′-diamino-*N*-methyldipropylamine (DMDPA; as a relatively long linker, to obtain additional flexibility) were added to the C terminus. We also designed PI polyamides to allow them to have symmetry and therefore the polyamides binds the match sequence in two orientations.[30] This leads to an increase in the number of alkylating sites, which are placed on either side of the binding site. The appropriate positions of PI polyamide moieties were linked with β-alanine, which allows for optimal H-bond donor/acceptor interactions while retaining flexibility, resulting in an increase in overall affinity and specificity.[35–38]

Regarding the synthesis of long-chain hairpin PI polyamides, only monomeric or/and dimeric units have been used in Fmoc SPPS to produce relatively long hairpin polyamides;[39] therefore, their yields were worse because of the increase in the number of reaction steps. In particular, Fmoc SPPS had difficulty in coupling with a P after the introduction of a GABA (γ-turn) , which functions to transform polyamides into hairpin structures that allow high affinity and specificity.[24,40] To overcome these difficulties, we plan to synthesize FmocHN-P-γ-CO₂H (**5**), Boc-D-Dab (FmocHN-P)-CO₂H (**6**), and FmocHN-P-β-I-CO₂H (**10**) as building blocks of Fmoc SPPS in solution phase,[39,41,42] in which time-consuming purifications are never required in any of the reaction steps (Scheme 1). All of them are washed and dried to produce powder states that can then be subjected to Fmoc SPPS without further purification. Compound **5** and **6**, which were used in Fmoc SPPS of polyamide **13** and **14**, respectively (Scheme 2), played an important role in preventing the P-lacking products and could be prepared using only one step. As compound **10** was used to construct a core moiety of the parent polyamide in this study, its introduction into Fmoc SPPS, in addition to that of compound **5** or **6**, enabled a great reduction in reaction steps (from 17 to eight steps)(Scheme 2). Consequently, we succeeded in reducing failure sequence products and enhanced the yield of the reaction. As we
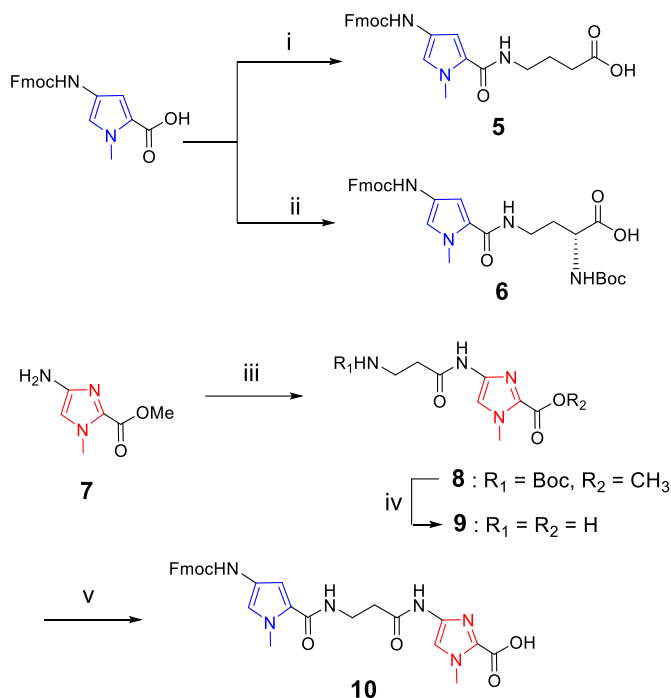
Scheme 1. (i) HCTU, DIEA, DMF, and then a solution of GABA in DIEA and DMF, rt, 3 h, 94%; (ii) PyBOP, DIEA, DMF, and then a solution of Boc-D-Dab-OH in DMF, rt, 3 h, quant; (iii) BocHN-β-alanine, HCTU, DIEA, DMF, rt, 1.5 h, 84%; (iv) TFA, H_2O, rt, then NaOH, MeOH, H_2O, 40°C, 2 h; (v) a solution of FmocHN-P-CO_2H and HCTU in DIEA and DMF, rt, 6 h, 73%.
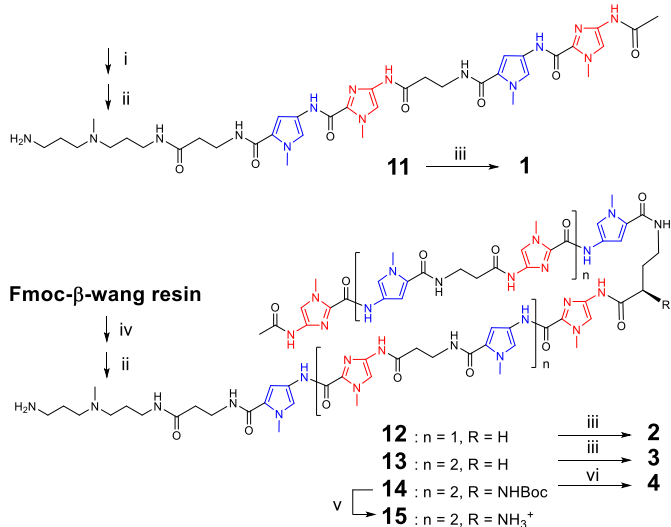
Scheme 2. (i) Fmoc solid-phase synthesis of **11**: FmocHN-P-CO_2H, **10**, FmocHN-I-CO_2H (in this order, total three steps), 20% piperidine, HCTU, DIEA, DMF, NMP in each step; (ii) DMDPA; (iii) Chl, PyBOP, DIEA, DMF, rt; (iv) Fmoc solid-phase synthesis of **12**: FmocHN-P-CO_2H, **10**, FmocHN-I-CO_2H, **5**, **10**, FmocHN-I-CO_2H (in this order, total six steps), 20% piperidine, HCTU, DIEA, DMF, NMP in each step; Fmoc solid-phase synthesis of **13** (**14**): FmocHN-P-CO_2H, **10**, FmocHN-I-CO_2H, **5** (**6**), **10**, FmocHN-I-CO_2H (in this order, total eight steps), 20% piperidine, HCTU, DIEA, DMF, NMP in each step; (v) TFA-DCM, rt; (vi) Chl, HCTU, DIEA, DMF, then TFA-DCM, 0 °C.

anticipated that it was less difficult for such a long-chain polyamide in a hairpin structure to be formed into the open state, we performed a chiral substitution of the γ-turn by an amino group (to produce polyamide **15** and conjugate **4**).[2,43]

The products obtained by programmable Fmoc SPPS introducing the building blocks were cleaved with DMDPA, to produce the amino PI polyamides **11–14** (Scheme 2). Conjugates **1–3** were produced via the PyBOP coupling with Chl. By conversion from **14**, the PI polyamide–Chl conjugate **4** was synthesized via the deprotection of the Boc group by TFA-DCM in the last step. The diamino PI polyamide **15** was also prepared by acid deprotection for SPR assays. Conjugates **1–4** were purified by reversed-phase HPLC, and confirmed by ESI-TOF-MS.

2.2. DNA-alkylating activity of conjugates 1-4

First, The DNA-alkylating activities of conjugates **1–3** were evaluated using polyacrylamide gel electrophoresis (PAGE), to investigate the effects of the structure or recognition length of PI polyamides on specific alkylation within CAG/CTG repeat sequences.[44,45] 5′-Texas Red-labeled 214 bp DNA fragments including a $(CAG/CTG)_{12}$ repeat sequence were prepared by transformation into pGEM-T Easy vectors, and subsequently by PCR amplification and purification (Figure S1).[46] Alkylation was carried out at 23°C for 18 h, followed by quenching by the addition of calf thymus DNA. The samples were heated at 95°C under neutral conditions for 10 min. Under these conditions, all the N3 of purines at the alkylated sites in the DNA fragment produced cleavage bands quantitatively on the gel.[47,48]

The results of the sequencing gel analysis of the DNA fragment treated with conjugates **1–3** are shown in Figure 2. Regarding the alkylation of the $(CAG)_{12}$ repeat sequence (Figure 2a), no alkylation by conjugates **1** and **2** occurred at the sites of the CAG repeat, and alkylation by conjugate **1** occurred at *site a*, which can be explained by recognition of expanded state with 1 bp mismatch, instead of dimer recognition.[49] Even at 10 times concentrations, conjugate **1** did not specifically alkylate the CAG repeat sequence (data not shown). At around 150 nanomolar concentrations, alkylation by conjugate **2** occurred within the CAG repeat, albeit with a specificity that was much lower than that of conjugate **3** (data not shown). However, we observed specific alkylation by conjugate **3** at sites of the target region with higher alkylating activities and sequence specificity at nanomolar concentrations (lanes 2–5), except alkylation at *site b* with 3 bp mismatch. Conversely, within the $(CTG)_{12}$ repeat sequence (Figure 2b), we observed specific alkylation by both conjugates **2** and **3**, except alkylation at *site c* and *d* with 3 bp mismatch by conjugate **2** and **3**, respectively, and no alkylation by conjugate **1**. These results indicate that, although 8 bp recognition through dimer (conjugate **1**) or hairpin (conjugate **2**) conformation is not sufficient, 11 bp recognition through hairpin conformation (conjugate **3**) is required for sequence-specific alkylation in both strands including CAG/CTG repeat sequences.

Next, having established that, among conjugates **1–3**, conjugate **3** is the most appropriate for targeting CAG/CTG repeat sequences, we used conjugates **3** and **4** to investigate DNA-alkylating activities for CAG/CTG repeat sequences in detail and the effect of chiral substitution at the α position of the γ-turn by an amino group on DNA-alkylating activities. The results of the sequencing gel analysis using a similar treatment to that mentioned above are shown in Figure 3. The $(CAG/CTG)_{12}$-containing DNA fragment prepared in this study comprised the target region (5′−$(CAG)_{12}$CAA–3′/5′−TTG$(CTG)_{12}$–3′) for these conjugates (Figure 3c). In fact, specific alkylation by conjugates **3** and **4** was observed within the target region between *sites 1* and *13*, or *sites 1'* and *15'*, except non-specific alkylation (*site b* and *site d*) as also seen in Figure 2 (Figure 3). For the CTG repeat sequence (Figure 3b), both conjugates alkylated at the N3 position of the guanines in the CTG trinucleotides, particularly from the third CTG (*site 6'*) to the fifth
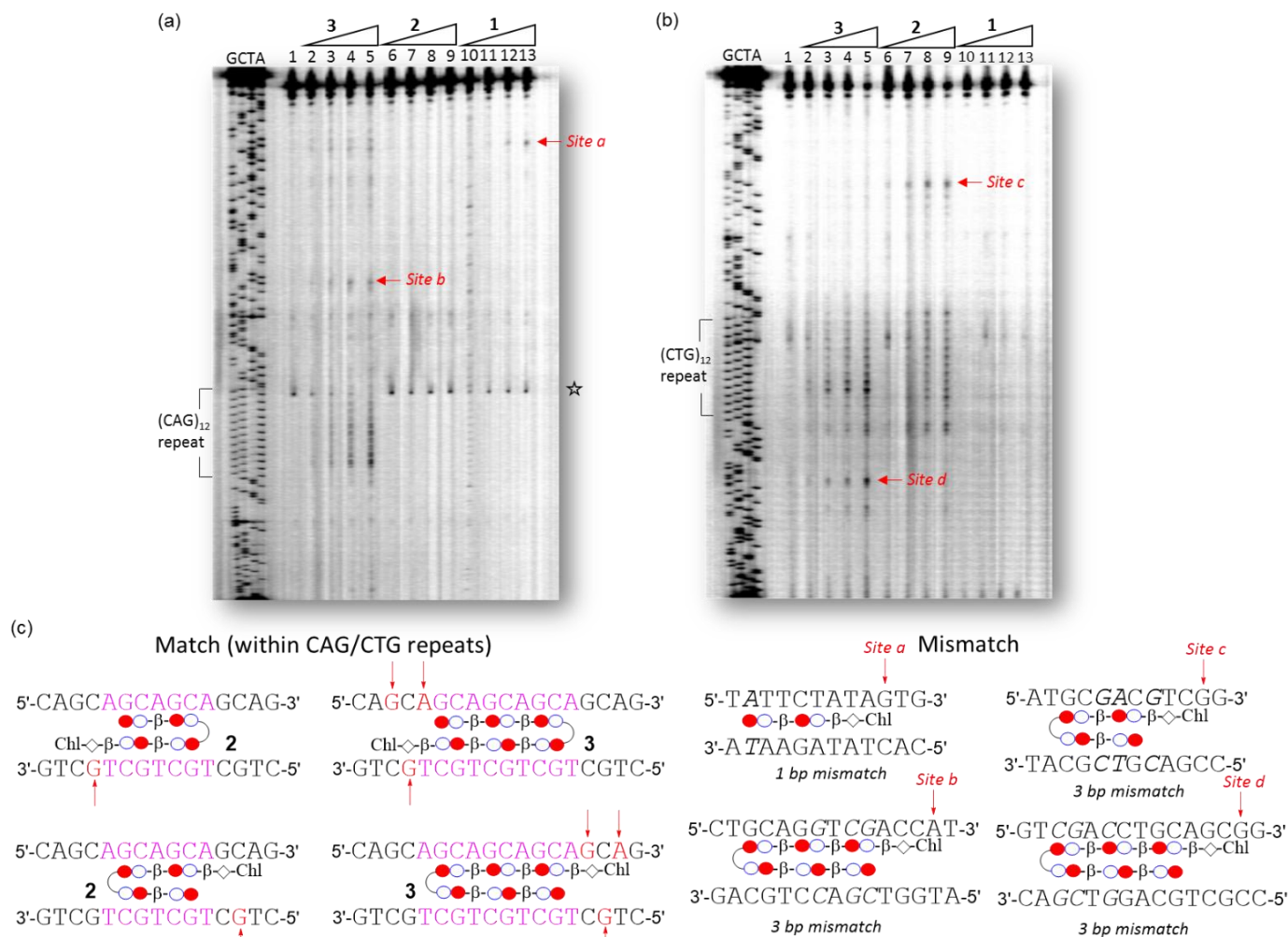
Figure 2. Thermally induced strand cleavage of 214 bp DNA fragments derived from 5′-Texas Red-labeled strands including (a) a $(CAG)_{12}$ or (b) $(CTG)_{12}$ repeat (8 nM) by conjugates **1**, **2**, and **3** at 23°C for 18 h. These two DNA strands are complementary. Lane 1: DNA control; lanes 2–5: 2.5, 5, 10, and 20 nM of **3**; lanes 6–9: 2.5, 5, 10, and 20 nM of **2**; lanes 10–13: 2.5, 5, 10, and 20 nM of **1**. (c) Schematic representation of the recognition and alkylation models of these conjugates. The arrows indicate the alkylating sites that were inferred from the sequencing gel analysis. Note that alkylation does not occur at the site indicated by ☆, as this band is also observed in lane 1 (control).[45]

CTG trinucleotide (*site 8'*) numbered from the 5′ side. These sites are included in ones which can be alkylated by the conjugates through two binding modes (*site 6' to 11'*), thus showing the relative higher cleavage intensities (Figure 3c). In total, we observed 15 cleavage bands as a result of the N3 adenine or guanine alkylation, which corresponded to prospective alkylating sites within the target region (Figure 3b and c). Conversely, specific alkylation by conjugates **3** and **4** within the CAG repeat occurred similarly and produced 14 cleavage bands. As with the case for the CTG repeats, the cleavage bands with the relatively higher intensities (*site 4 to 9*) mostly correspond to the sites which can be alkylated by the conjugates through two binding modes (*site 4 to 10*). In addition, we observed that these cleavage bands on the gel were a little broader than those obtained for the CTG repeat, which suggests that both N3 adenines and their adjacent guanines in CAG trinucleotides were specifically alkylated by conjugates **3** and **4** (Figure 3a and c). We found that, although the effect of a chiral substitution on DNA-alkylating activities was not observed in this experimental system, rationally designed long-chain PI polyamide–Chl conjugates with a relatively long linker had the ability to alkylate both of the strands and increase the number of the alkylating sites by binding in the two orientations.

2.3. The effect of conjugate 4 on transcription

To examine the influence of alkylation by a PI polyamide–Chl conjugate on transcription, non-labelled DNA fragments including a $(CAG/CTG)_{12}$ repeat (sequences **A** and **B**, as shown in Figure 4a) were treated with conjugate **4** and its parent polyamide **15** as a control, and then transcribed by T7 RNA polymerase. The DNA fragments were prepared by a similar method for the sequence gel analysis. Sequence **B** was prepared by inserting a $(CAG/CTG)_{12}$ repeat and its sequence is the same as that used in the sequencing gel analysis. Sequence **A** was prepared by inserting a inversed $(CAG/CTG)_{12}$ repeat (Figure 4a and S2). Its transcripts were analyzed by 10% denaturing PAGE,[50] and the results of the gel stained with SYBR Gold are shown in Figure 4b and c. In only lanes 2–4 of figure 4b and c, in addition to full-length RNAs (predicted length: 197 nt), truncated RNAs (predicted length : Ca. 60–101 nt) appeared as ladder-like bands in a dose-dependent manner, suggesting that conjugate **4** alkylated the CAG/CTG repeat region and arrest the progression of RNA polymerase at the alkylating sites. Furthermore, we observed that within the target region, the number of the ladder-like bands obtained from transcription of sequence **A** was greater than that obtained from sequence **B**. this shows the alkylation by conjugate **4** at both adenines and guanines in the $(CAG)_{12}$ repeat, which corresponds to the results of the sequence gel analysis (Figure 3a). Moreover, this observation demonstrates that alkylation of the template
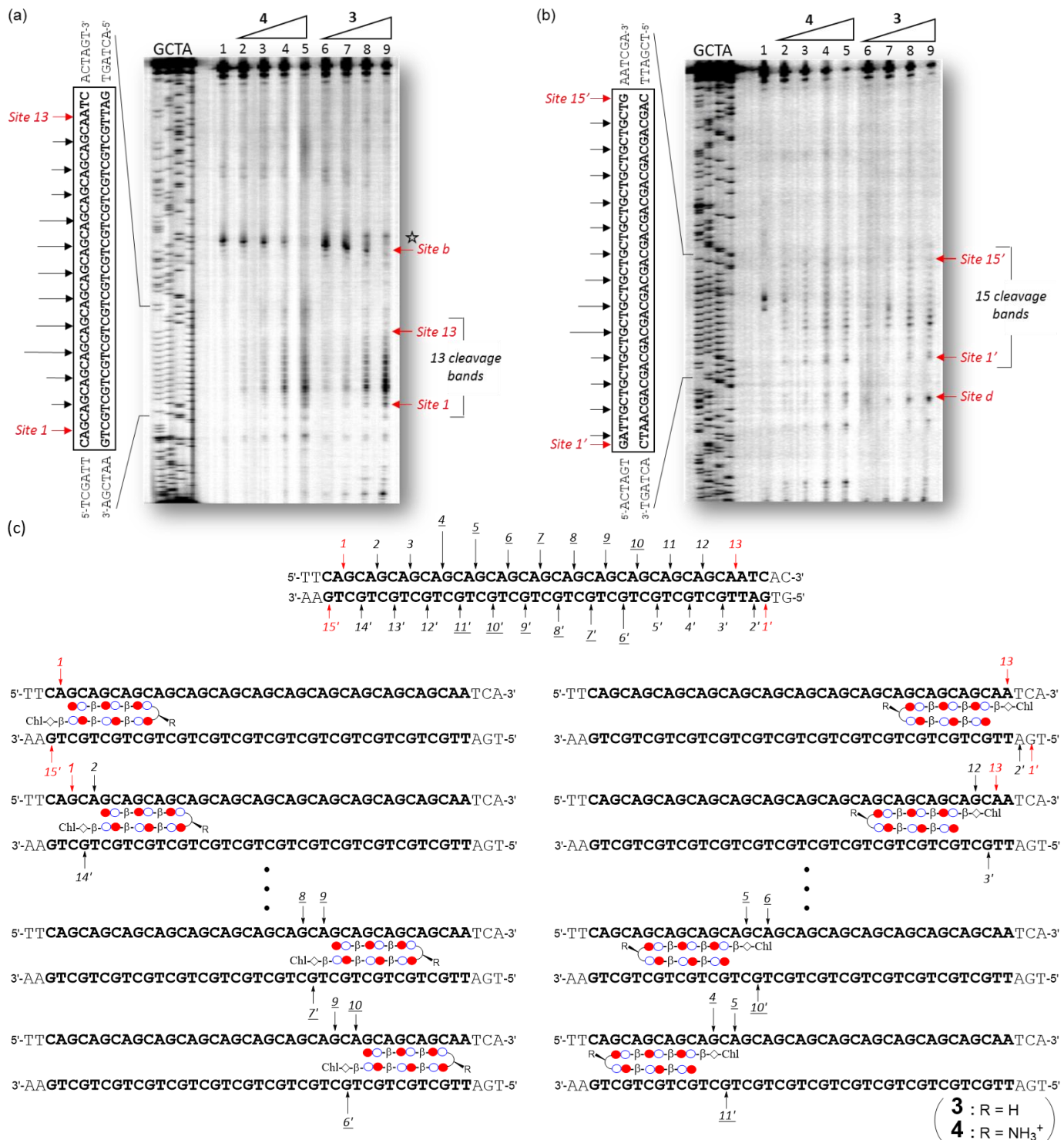
Figure 3. Thermally induced strand cleavage of 214 bp DNA fragments derived from 5′-Texas Red-labeled strands including (a) a $(CAG)_{12}$ or (b) $(CTG)_{12}$ repeat (8 nM) by conjugates **3** and **4** at 23°C for 18 h. These two DNA strands are complementary. Lane 1: DNA control; lanes 2–5: 2.5, 5, 10, and 20 nM of **4**; lanes 6–9: 2.5, 5, 10, and 20 nM of **3**. (c) Schematic representation of the recognition and alkylation models of these conjugates with two binding modes. The arrows indicate the alkylating sites within the target region that were inferred from the sequencing gel analysis. The underlined sites show the ones at which the conjugates can alkylate with two binding modes. Note that alkylation does not occur at the site indicated by ☆, as this band is also observed in lane 1 (control).[45]

strands by conjugate **4** terminates transcriptional elongation, which is also supported by the fact that other retarded bands on the gels of Figure 4b and c amount to ones indicating non-specific alkylation on the gel of Figure 3b. Conversely, little appearance of retarded bands in lanes 6–8 shows that the parent polyamide **15** dose not inhibit the transcription. This result is consistent with the report that polyamides not conjugated with alkylating reagents dose not inhibit transcriptions by T7 RNA polymerase in the absence of histones.[51]

2.4. Binding properties of the parent polyamides to target hairpin DNA

To obtain additional information about the interaction of polyamides with the CAG/CTG repeat region, using SPR assays we evaluated and compared the binding affinities and kinetics parameters of the four parent polyamides **11–13** and **15** (which were purified on a reversed-phase flash column and confirmed by ^{1}H-NMR and ESI-TOF-MS) for a biotinylated hairpin DNA that included the target region (5′–AGCAGCAGCA–3′/5′–

TGCTGCTGCT–3′) (Figure 5a).[52,53] The solution of each polyamide in HBS buffer with 0.1% DMSO flowed on a sensor chip onto which the designed hairpin DNA was immobilized, and the interaction was visualized as a sensorgram (Figure 5b–e). The values of k_a, k_d, and K_D obtained by curve fitting are shown in Table 1. In terms of the influence of the recognition length over binding affinity, the results of this experiment showed that **12** had about 4 times higher binding ability, which means the lower K_D values, than that of **13**, mainly owing to the difference in the rates of association (k_a). One possible explanation for this observation is that it is more difficult for **13** to penetrate into the DNA minor groove of the target sequence with the closed state (hairpin conformation), because of its flexibility resulting from the long-chain. Another possibility of the stronger binding by **12** is that although this polyamide interacts with the DNA with the 1:1 binding mode, it has two matching positions in the target region. This facilitates the penetration of the polyamide, which is reflected in the k_a value. Concerning the rates of dissociation (k_d), we observed that the extension of the recognition length results in a small decrease in the k_d values. This indicates that the increase in the hydrogen bonds from 7 bp recognition does not affect the tendency to localize at the target region in the DNA minor groove, even with the introduction of the cationic amino substituent (polyamide **15**). Conversely, comparing the binding affinity of **13** and **15** allowed us to investigate the effect of a chiral substitution on the γ-turn with an amino group, and revealed that **15** had about 5 times higher

Figure 4. Transcribed DNA sequences and transcription products. (a) Sequence **A** has a template strand that contains a (CAG)$_{12}$ repeat and Sequence **B** has one that contains a (CTG)$_{12}$ repeat. The numbers indicate position of bases in the sequence from the beginning of transcription. Bold domains indicate the inserted (CAG/CTG)$_{12}$ repeat sequence. The arrows show probably alkylated bases. (b and c) Effects on transcription of conjugate 4 and its parent polyamide 15; (b) Sequence **A** and (c) Sequence **B**. Lane 1: without any polyamides; lanes 2–4: 10, 20, and 30 nM of 4; lane 5: ssRNA markers, lanes 6–8: 10, 20, and 30 nM of 15.

binding ability than that of **13**. Particularly, we observed the difference in the k_a values, suggesting that **15** is more likely to access the DNA with the cationic amino group positioned in the DNA minor groove to give better electrostatic attraction, retaining the forward hairpin folding of **15**.[43] Since Polyamide **11** was designed to bind to the target region with a dimer conformation, the sensorgram obtained for **11** was fitted to 2:1 stoichiometry binding model that represents dynamics in forming the complexes with two polyamides bound side by side in the same region (see experimental section).[26,27,52] As seen in Table 1, the value of K_{D2} (1.3×10^{-8} M) is lower than that of K_{D1} (6.1×10^{-7} M), indicating the cooperative binding for the target region. In order to compare the other K_D values that were obtained by fitting to the 1:1 binding model, a $(K_{D1} \cdot K_{D2})^{1/2}$ value was used, which is the value adjusted to a per bound molecule.[52] As a result, the value for **11** (8.9×10^{-8} M) was the highest of those in this experiment, which corresponded to lower alkylating activity and specificity.

3. Conclusion

We rationally designed and readily synthesized PI polyamide–Chl conjugates targeting CAG/CTG repeat sequences with different lengths or structures. In particular, by introducing the novel building blocks that were synthesized in solution phase to Fmoc SPPS, we prepared long-chain hairpin PI polyamides which otherwise would be difficult to synthesize. The sequencing gel analysis demonstrated that the long-chain hairpin PI polyamide
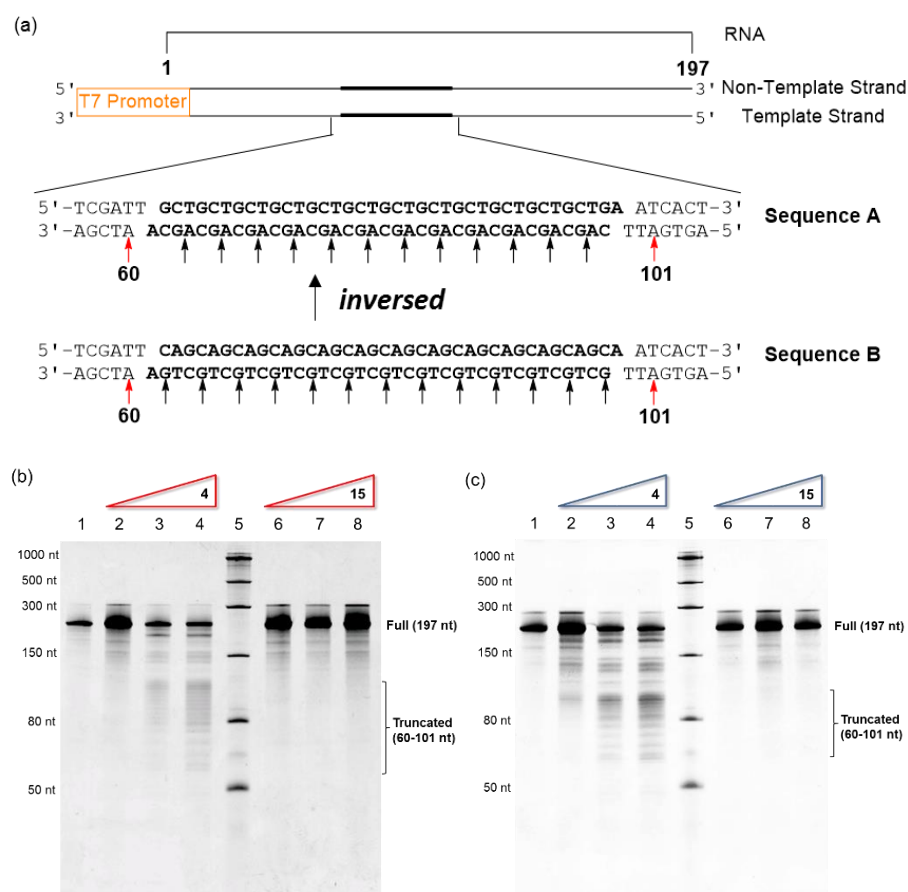
conjugates showed sequence-specific alkylation at the N3 adenines or guanines in the CAG/CTG repeat sequence. *In vitro* transcription assays demonstrated that sequence-specific alkylation by a PI polyamide–Chl conjugate terminated the progression of RNA polymerase at the alkylating sites. The chiral substitution on the γ-turn with an amino group enhanced the binding affinity observed in SPR assays. These *in vitro* experiments revealed that conjugate **4** was most appropriate for targeting CAG/CTG repeat sequences in terms of chemical activity. The propensity of conjugate **4** to specifically alkylate both strands suggests the possibility of selective DNA damage, which thus may lead to the deletion of expanded CAG/CTG repeat sequences within the nucleus. Based on its property of the transcriptional inhibition by specific alkylation at both the CAG and CTG repeat regions, conjugate **4** may have potency of efficiently preventing the elongation of CAG and CTG repeat sequences and the accumulation of expanded rCUG repeat transcripts and polyglutamine tracts. Moreover, as reported previously,[41] the concept that building blocks are prepared in solution phase to provide relatively long polyamides can be applied not only to this case, but also to various cases of targeting repeated sequences.

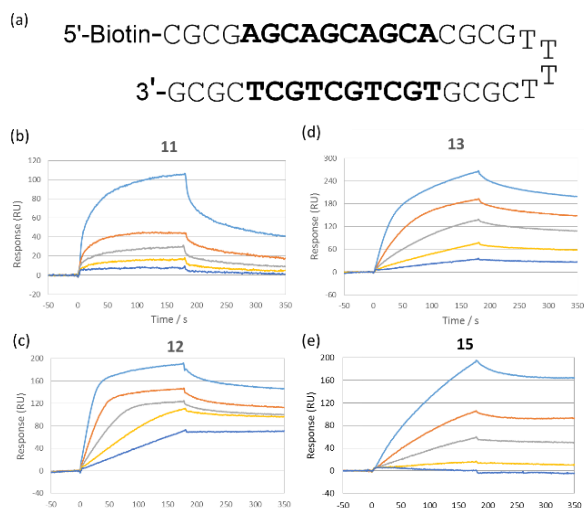## 4. Experimental section

4.1. General

Figure 5. SPR assays to evaluate the binding affinity of the parent polyamides **11–13** and **15** to that designed hairpin DNA oligomer containing CAG/CTG repeats which was immobilized on a sensor chip. (a) A biotinylated DNA oligomer used in this study. SPR response curves of the interaction of the parent polyamides at various concentrations with the hairpin DNA immobilized on a sensor chip. From the top, concentration: (b) 3200, 1600, 800, 400, and 200 nM; (c) 500, 250, 125, 62.5, and 31.25 nM; (d) 1000, 500, 250, 125, and 62.5 nM; (e) 200, 100, 50, 25, and 12.5 nM, respectively.

Table 1. The values of affinities and kinetics parameters by curve fitting for each sensorgam.

PI polyamide	$k_a/M^{-1}s^{-1}$	k_d/s^{-1}	K_D/M	χ^{2d}
11[c]	$k_{a1} = 8.8 \times 10^3$	$k_{d1} = 5.4 \times 10^{-3}$	$K_{D1} = 6.1 \times 10^{-7}$	2.3
	$k_{a2} = 4.4 \times 10^2$	$k_{d2} = 5.7 \times 10^{-6}$	$K_{D2} = 1.3 \times 10^{-8}$	
12[a]	1.2×10^5	1.6×10^{-3}	1.4×10^{-8}	7.5
13[b]	2.3×10^4	1.2×10^{-3}	5.3×10^{-8}	4.1
15[b]	8.5×10^4	9.5×10^{-4}	1.1×10^{-8}	3.8

Determined by fitting to [a]1:1 binding model with drifting baseline, [b]1:1 binding model with mass transfer, or [c]2:1 binding with mass transfer using BIAevaluation 4.1 program to give better fitting. [d]The closeness of fit is described by the statistical value χ^2.

^{1}H NMR spectra were recorded on 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), brs (broad singlet), d (doublet), t (triplet), q (quartet), and m (multiplet). HPLC analysis was performed on a Jasco Engineering PU-2080 plus series system using a 150×4.6 mm X-Terra MS C_{18} reversed-phase column in 0.1% TFA in water with acetonitrile as the eluent at a flow rate of 1.0 mL/min and a linear gradient elution of 0 to 100% acetonitrile in 20 or 40 min with detection at 254 nm. Collected fractions were analyzed byESI-TOF-MS (Bruker). Reversed-phase flash column chromatography was performed on CombiFlash Rf (Teledyne Isco, Inc.) using a 4.3 g reversed-phase flash column (C18 RediSep Rf) in 0.1% TFA in water with acetonitrile as the eluent at a flow rate of 18.0 mL/min and a linear gradient elution of 0 to 35% acetonitrile in 5 to 40 min with detection at 254 nm. HPLC purification was performed with a Jasco Engineering UV2075 HPLC UV/VIS detector and a PU- 2080 plus series system using a 150×4.6 mm Chemcobond 5-ODS-H reversed-phase column in 0.1% TFA in water with acetonitrile as the eluent at a flow rate of 1.0 mL/min and a linear gradient elution of 30 to 75% acetonitrile in 30 min with detection at 254 nm. Polyacrylamide gel electrophoresis was performed on a HITACHI SQ5500-S DNA sequencer. Long Ranger gel solution (50%) was purchased from FMC bioproducts. Calf thymus DNA (10 mg/mL, 1 μL) was purchased from Invitrogen and Thermo Sequence core sequencing

kit was from GE Healthcare. Loading dye was composed of 10 mL of formamide, 200 μL of H_2O, 300 μL of a 0.5 M EDTA 2Na in H_2O (Nacalai Tesque Inc.), and 2.5 mg of new fuchsin (Merck). In vitro transcription assays were carried out by *in vitro* Transcription T7 Kit (Takara Bio Inc.) and the transcripts were investigated on 10 % TBE-Urea Gels containing 7 M urea (Invitrogen). 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. Chlorambucil (Chl) was purchased from Sigma, and BocHN-β-alanine, 3,3′-diamino-*N*-methyldipropylamine (DMDPA) and 10% Pd/C were from Aldrich. 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. FmocHN-P-CO_2H, FmocHN-I-CO_2H, O_2N-I-COCl_3, *N*,*N*-dimethylformamide (DMF), 1-methyl-2-pyrrolidone (NMP), and piperidine were purchased from Wako, and Boc-D-Dab-OH was from Watanabe Chemical Ind., Ltd. PyBOP were purchased from Novabiochem. Diisopropylethylamine (DIEA), 4-amino-n-butyric carboxylic acid (GABA) were purchased from Nacalai Tesque, Inc. Trifluoroacetic acid (TFA) was purchased from Kanto Chemical Co., Inc. Dichloromethane (DCM) was purchased from Sasaki chemical co., Ltd. The other reagents and solvents were purchased from standard suppliers and used without further purification.

4.2. Syntheses of compound 8-10

4.2.1. FmocHN-P-γ-CO_2H (5)

Compound **5** was synthesized from FmocHN-P-CO_2H. To a solvent of DMF was added compound **5** (1.50 g, 4.1 mmol), HCTU (1.71 g, 4.1 mmol) and DIEA (1.44 mL, 8.3 mmol), and the mixture was stirred at room temperature for 15 min. The reaction mixture was added slowly to a solution of GABA (858 mg, 8.3 mmol) in 1.4 mL DIEA and 10 mL DMF, and it was stirred at room temperature for 3 h. After concentration, the oil residue was dissolved in the minimum amount of DCM, washed with diethyl ether. After the supernatant was removed, to the residue was added a large amount of water to afford the precipitation. It was filtered and then dried in *vacuo* to obtain **5** as a brown powder (1.73 g, 94%). The compound was used in the solid-phase synthesis without further purification. ^{1}H NMR (600 MHz, DMSO-d_6) δ 9.42 (s, 1H, NH), 8.31 (brs, 1H, NH), 7.90 (d, *J* = 7.1 Hz, 2H, CH), 7.73 (d, *J* = 7.5 Hz, 2H, CH),7.42 (t, *J* = 7.5 Hz, 2H, CH), 7.34 (t, *J* = 7.6 Hz, 2H, CH), 6.85 (s, 1H, CH), 6.65 (s, 1H, CH), 4.40 (d, *J* = 6.8 Hz, 2H, CH_2), 4.27 (t, *J* = 6.8 Hz, 1H, CH), 3.76 (s, 3H, NCH_3), 3.14 (m, 2H, CH_2), 2.16 (brs, 2H, CH_2), 1.67 (m, 2H, CH_2), ESI-MS m/z calcd for $C_{25}H_{25}N_3O_5$ [M+H]+ 448.1872; found 448.1869.

4.2.2. Boc-D-Dab (FmocHN-P)-CO_2H (6)

Compound **8** was synthesized from FmocHN-P-CO_2H. To a solvent of 5 mL DMF was added FmocHN-P-CO_2H (300 mg, 0.83 mmol), PyBOP (431 mg, 0.83 mmol) and DIEA (0.40 mL), and the mixture was stirred at room temperature for 1 h. To the reaction mixture was added Boc-D-Dab-OH (190 mg, 0.87 mmol) and 2 mL DMF, and then it was stirred at room temperature for 3 h. After concentration, the oil residue was dissolved in the minimum amount of ethyl acetate, washed with diethyl ether. After the supernatant was removed, to the residue was added a large amount of water to afford the precipitation. It was filtered and then dried in *vacuo* to obtain **6** as a brown powder (478 mg, quant). The compound was used in the solid-phase synthesis without further purification. ^{1}H NMR (600 MHz, CDCl_3), δ 7.77 (d, 2H, *J* = 7.6 Hz, CH), 7.61 (d, 2H, *J* = 6.9 Hz, CH), 7.40 (t, 2H, *J* = 7.2 Hz, CH), 7.31 (t, 2H, *J* = 7.2 Hz, CH), 6.85 (s, 1H), 6.63 (s, 1H), 6.54 (s, 1H), 5.57 (d, 1H, *J* = 6.2 Hz, NH), 4.48 (d, 2H, *J* = 6.2 Hz, CH_2),

4.34 (d, 2H, J = 6.9 Hz), 4.23 (t, 1H, J = 6.2 Hz, CH), 3.83 (s, 3H, NCH$_3$), 3.76 (m, 1H, CH), 3.68 (m, 1H, CH), 3.17 (m, 1H, CH), 2.02 (m, 1H, CH), 1.95 (m, 1H, CH), 1.43 (s, 9H, CH$_3$), ESI-MS m/z calcd for C$_{30}$H$_{34}$N$_4$O$_7$ [M+H]$^+$ 563.2506; found 563.2491.

4.2.3. BocHN-β-I-CO₂Me (8)

Compound **8** was synthesized from O$_2$N-I-COCl$_3$. Compound **7** was prepared from O$_2$N-I-COCl$_3$ by 2 steps using much the same method as a previously reported one.[54] To a DMF solvent (12 mL) was added compound **7** (1.37 g, 8.83 mmol), BocHN-β-alanine (1.68g, 8.86 mmol), HCTU (3.85 g, 9.30 mmol), DIEA (3.10 mL, 17.7 mmol), and the solution was mixed and stirred at room temperature for 1.5 h. After concentration, a large amount of water was added to the oil residue to produce the precipitation. It was filtered and dried in *vacuo* to afford **8** as a white powder (2.44 g, 7.46 mmol, 84%). The compound was used in the next reaction step without further purification. ^{1}H NMR (600 MHz, CDCl$_3$) δ 8.45 (s, 1H, NH), 7.67 (d, J = 1.4 Hz, 1H, CH), 7.49 (s, 1H, NH), 4.00 (s, 3H, OCH$_3$), 3.95 (s, 3H, NCH$_3$), 3.49 (q, J = 6.2 Hz, 2H,CH$_2$), 2.63 (t, J = 6.22 Hz, 2H CH$_2$), 1.42 (s, 9H, CH$_3$).

4.2.4. FmocHN-P-β-I-CO₂H (10)

TFA (2.4 mL) was used in the deprotection of a Boc group that belongs to compound **8** (1.23 g, 3.75 mmol), and it was then evaporated. To a solution of obtained crude in H$_2$O/MeOH (1:1, total 20 mL) was added NaOH (1.26 g), and the solution was stirred at 40℃ for 2 h. After MeOH was removed from the reaction mixture by an evaporator, it was neutralized by 6N HClaq and concentrated in *vacuo* to afford crude gelatinous **9**. To a solution of an intact **9** in DMF (5 mL) was added a solution of FmocHN-P-CO$_2$H (816 mg, 2.25 mmol), HCTU (932 mg, 2.25 mmol) and DIEA (1.3 mL) in DMF (10 mL) which was stirred at room temperature for 1 h, and the reaction mixture was stirred at room temperature for 5 h. After a sault was removed from the reaction mixture by filtration, a large amount of water was added to the oil residue obtained to produce the precipitation. The crude obtained by filtration was dissolved in a minimum amount of DCM and added diethyl ether to be precipitated again. The pellet by a centrifuge was collected and dried in *vacuo* to afford **10** as a pale brown powder (1.07 g, 1.93 mmol, 73%). The compound was used in the solid-phase synthesis without further purification. ^{1}H NMR (600 MHz, DMSO-d_6) δ 9.50 (s, 1H, NH), 9.40 (s, 1H, NH), 8.01 (s, 1H, NH), 7.90 (d, J = 7.6 Hz, 2H, CH), 7.72 (d, J = 7.5 Hz, 2H, CH), 7.42 (t, J = 7.6 Hz, 2H, CH), 7.33 (t, J = 7.6 Hz, 2H, CH), 7.20 (s, 1H, CH), 6.86 (s, 2H, CH), 4.41 (d, J = 6.2 Hz, 2H, CH$_2$), 4.26 (t, J = 6.8 Hz, 1H, CH), 3.95 (s, 3H, NCH$_3$), 3.77 (s, 3H, NCH$_3$), 3.48 (m, 2H, CH$_2$), 2.52 (t, J = 2.0 Hz, 2H, CH$_2$), ESI-MS m/z calcd for C$_{29}$H$_{28}$N$_6$O$_6$ [M+H]$^+$ 557.2149; found 557.2137.

4.3. Syntheses of parent PI polyamides and their chlorambucil conjugate

4.3.1. 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 step was set up to solve by NMP on the synthetic line. The synthetic procedure of all PI polyamides was as follows; twice deblocking for 4 min with 20% piperidine/DMF (0.6 mL), activating for 2 min with HCTU (88 mg, 0.21 mmol) in NMP (1 mL) and 10% DIEA/DMF (0.4 mL), coupling for 60 min, and washing with DMF. All coupling were carried out with a single-coupling cycle. Building blocks used in this study are FmocHN-P-CO$_2$H (77 mg), FmocHN-I-CO$_2$H (77 mg), and compound **10** (70 mg) for synthesis of polyamide **11**; FmocHN-P-CO$_2$H (77 mg),

FmocHN-I-CO$_2$H (77 mg), compound **5** (70 mg), and compound **10** (70 mg) for synthesis of polyamide **12** and **13**; FmocHN-P-CO$_2$H (77 mg), FmocHN-I-CO$_2$H (77mg), compound **6** (70 mg), and compound **10** (70 mg) for synthesis of polyamide **14**. At the last capping process, the samples were washed with 20% acetic acid in DMF (1 mL). All lines are purged with solution transfers and bubbled by N$_2$ gas for stirring the resin. After the completion of the synthesis, the resin was washed with DMF (2 mL) and methanol (2 mL) and then dried in a desiccator at room temperature in *vacuo*.

4.3.2. AcIP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH₂ (11)

Using a 43 mg Fmoc-β-Wang resin (0.55 mmol/g) and a novel building block **10**, compound **11** was synthesized in a stepwise reaction by Fmoc solid-phase protocol described above. After resulting AcIP-β-IP-β-wang resin was cleaved with 200 μL of DMDPA for 3 h at 55°C, the solvent was evaporated. The residue was dissolved in the minimum amount of DCM and subsequently washed with diethyl ether to yield a 15.1 mg white solid, which was used in next coupling reaction without further purification. For SPR assays, the crude compound was purified by reversed-phase flash column chromatography (water with 0.1% TFA containing 0‑35% CH$_3$CN over a linear gradient in 5 to 40 min). The peak around 23 min was collected and lyophilized to produce **11** (6.6 mg, 7.9 μmol, 23%) as a yellow solid. ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.28 (s, 1H, NH) 10.23 (s, 1H, NH) 9.93 (s, 1H, NH) 9.90 (s, 1H, NH) 8.09-8.05 (m, 3H, NH) 7.45 (s, 1H, CH) 7.41 (s, 1H, CH) 7.22 (d, 1H, J = 1.4 Hz, CH) 7.20 (s, 1H, CH) 6.93 (s, 2H, CH) 3.94 (s, 1H, NCH$_3$) 3.93 (s, 1H, NCH$_3$) 3.81 (s, 1H, NCH$_3$) 3.80 (s, 1H, NCH$_3$) 3.42 (m, 6H, CH$_2$) 2.87 (br, 3H, CH$_3$) 2.77-2.73 (m, 6H, CH$_2$) 2.59 (t, J = 6.8 Hz, 2H, CH$_2$) 2.35 (t, 2H, J = 7.5 Hz, CH$_2$) 2.02 (s, 3H, CH$_3$) 1.77 (m, 4H, CH$_2$), ESI-MS m/z calcd for C$_{37}$H$_{53}$N$_{15}$O$_7$ [M+2H]$^{2+}$ 410.7205; found 410.7201.

4.3.3. AcIP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH-Chl (1)

To a solution of the crude compound **11** (2.0 mg) in DMF (200 μL), PyBOP (5.1 mg, 10 μmol), DIEA (1.7 μL, 10 μmol) and chlorambucil (3.1 mg, 10 μmol) were added. The reaction mixture was incubated for overnight at room temperature. Evaporation of the solvent gave a yellow crude, which was washed with ether and DCM and subsequently was purified by HPLC using a Chemcobond 5-ODS-H column (water with 0.1% TFA/MeCN 30-75% linear gradient, 0-30 min, detected at 254 nm), to produce **1** (1 mg, 0.91 μmol) as a pale yellow powder. ESI-MS m/z calcd for C$_{51}$H$_{72}$Cl$_2$N$_{16}$O$_8$ [M+2H]$^{2+}$ 553.2548; found 553.2527.

4.3.4. AcIP-β-IP-γ-IP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH₂ (12)

Using a 45 mg Fmoc-β-Wang resin (0.55 mmol/g) and two novel building blocks **5** and **10**, compound **12** was synthesized in a stepwise reaction by Fmoc solid-phase protocol described above. A subsequent synthetic procedure similar to that used for the preparation of compound **11** provided compound **12** (3.6 mg, 2.5 μmol, 10%) as a yellow powder. ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.29 (s, 1H, NH) 10.27 (s, 1H, NH) 10.25 (s, 1H, NH) 10.23 (s, 1H, NH) 9.96-9.92 (m, 4H, NH) 8.09-8.06 (m, 5H, NH) 7.45(s, 2H, CH) 7.43 (s, 1H, CH) 7.40 (s, 1H, CH) 7.22 (s, 1H, CH) 7.21 (s, 1H, CH) 7.20 (s, 2H, CH) 6.95 (s, 1H, CH) 6.93 (s, 3H, CH) 3.94-3.93 (m, 12H, NCH$_3$) 3.81 (s, 6H, NCH$_3$) 3.80 (s, 6H, NCH$_3$) 3.42 (m, 8H, CH$_2$) 3.19 (t, J = 6.2 Hz, 2H, CH$_2$) 2.86 (s, 3H, NCH$_3$) 2.75-2.73 (m, 6H, CH$_2$) 2.59 (t, J = 7.6 Hz, 6H, CH$_2$) 2.38 (m, 1H, CH$_2$) 2.01 (s, 3H, CH$_3$) 1.78-1.76 (m, 6H, CH$_2$), ESI-MS m/z calcd for C$_{66}$H$_{87}$N$_{27}$O$_{13}$ [M+2H]$^{2+}$ 733.8567; found 733.8549.

4.3.5. AcIP-β-IP-γ-IP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH-Chl (2)

To a solution of the crude compound **12** (1.9 mg) in DMF (200 μL), PyBOP (5.9 mg, 12 μmol), DIEA (1.9 μL, 11 μmol) and

chlorambucil (3.3 mg, 11 μmol) were added. A subsequent synthetic procedure similar to that used for the preparation of compound **1** provided **2** (0.5 mg, 0.28 μmol) as a pale yellow powder. ESI-MS m/z calcd for $C_{80}H_{104}Cl_2N_{28}O_{14}$ $[M+2H]^{2+}$ 876.3910; found 876.3892.

4.3.6. AcIP-β-IP-β-IP-γ-IP-β-IP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH₂ (13)

Using a 50 mg Fmoc-β-Wang resin (0.60 mmol/g) and two novel building blocks **5** and **10**, compound **13** was synthesized in a stepwise reaction by Fmoc solid-phase protocol described above. A subsequent synthetic procedure similar to that used for the preparation of compound **11** provided compound **13** (4.9 mg, 2.4 μmol, 8%) as a yellow solid. ^{1}H NMR (600 MHz, DMSO-d_6) δ 10.29 (s, 3H, NH) 10.26 (s, 1H, NH) 10.25 (s, 1H, NH) 10.23 (s, 1H, NH) 9.94 (s, 1H, NH) 9.93 (s, 2H, NH) 9.92 (s, 2H, NH) 9.91 (s, 1H, NH) 8.07 (m, 7H, NH) 7.44 (s, 3H, CH) 7.43 (s, 1H, CH)7.41 (s, 2H, CH) 7.21-7.19 (m, 6H, CH) 6.95-6.93 (m, 6H,CH) 3.94 (s, 12H, NCH₃) 3.93 (s, 6H, NCH₃) 3.81 (s, 12H,NCH₃) 3.80 (s, 6H, NCH₃) 3.41 (m, 14H, CH₂) 2.77-2.72 (m, 6H, CH₂) 2.58 (t, J = 6.8 Hz, 10H, CH₂) 2.37-2.33 (m, 2H, CH₂) 2.08 (s, 3H, CH₃) 2.02 (s, 3H, CH₃) 1.79-1.75 (m, 6H, CH₂), ESI-MS m/z calcd for $C_{94}H_{119}N_{39}O_{19}$ $[M+3H]^{3+}$ 700.3260 ; found 700.3239.

4.3.7. AcIP-β-IP-β-IP-γ-IP-β-IPy-β-IPy-β-(CH₂)₃-NCH₃-(CH₂)₃-NH-Chl (3)

To a solution of the crude compound **13** (2.1 mg) in DMF (200 μL), PyBOP (3.9 mg, 7.6 μmol), DIEA (1.3 μL, 7.5 μmol) and chlorambucil (2.3 mg, 7.6 μmol) were added. A subsequent synthetic procedure similar to that used for the preparation of compound **1** provided **3** (0.6 mg, 0.25 μmol) as a pale yellow powder. ESI-MS m/z calcd for $C_{108}H_{136}Cl_2N_{40}O_{20}$ $[M+3H]^{3+}$ 795.3489 ; found 795.3477.

4.3.8. Boc-D-Dab (AcIP-β-IP-β-IP)-IP-β-IP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH₂ (14)

Using a 44 mg Fmoc-β-Wang resin (0.60 mmol/g) and two novel building blocks **6** and **10**, compound **14** was synthesized in a stepwise reaction by Fmoc solid-phase protocol described above. Resulting Boc-D-Dab (AcImPy-β-ImPy-β-ImPy)-ImPy-β-ImPy-b-ImPy-β-wang resin was cleaved with 200 μL of DMDPA and 100 μL DMF for 3 h at 45°C, the solvent was evaporated and the residue was dissolved in the minimum amount of DCM, washed with diethyl ether to yield a 18.5 mg white solid, which was used in next reactions without further purification.

4.3.9. AcIP-β-IP-β-IP-D-Dab-IP-β-IP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH₂ (15).

To a solution of the crude compound **14** (16.3 mg) in DCM (600 μL) was added TFA (400 μL) and the mixture was shook at room temperature for 30 min. The mixture was concentrated in *vacuo* and the residue was dissolved in the minimum amount of DCM, washed with diethyl ether to afford the oil residue. The residue was purified by reversed-phase flash column as mentioned above to obtain **15** (3.5 mg, 1.7 μmol) as a yellow solid. ESI-MS m/z calcd for $C_{94}H_{120}N_{40}O_{19}$ $[M+3H]^{3+}$ 705.3296 ; found 705.3255.

4.3.10. AcIP-β-IP-β-IP-D-Dab-IP-β-IP-β-IP-β-(CH₂)₃-NCH₃-(CH₂)₃-NH-Chl (4)

To a solution of the crude compound **14** (16.3 mg) in DMF (1 mL), PyBOP (41.8 mg, 80 μmol), DIEA (15 μL) and chlorambucil (22.6 mg, 74 μmol) were added. The reaction mixture was mixed for 9 h at room temperature. Evaporation of the solvent gave a yellow residue, which was dissolved in the minimum amount of DCM, washed with diethyl ether. It was purified by reverse-phase flash column chromatography as mentioned above and the boc

group was then deprotected with TFA-DCM (1:1, total 2.4 mL) to produce the oil residue. It was dissolved in the minimum amount of DCM, washed with diethy ether, and purified by HPLC as mentioned above to produce **4** (0.7 mg, 0.29 μmol) as a pale yellow powder. ESI-MS m/z calcd for $C_{108}H_{137}Cl_2N_{41}O_{20}$ $[M+3H]^{3+}$ 800.3525 ; found 800.3488.

4.4. Preparation of plasmid containing (CAG/CTG)₁₂ repeat sequences

All DNA fragments and primers for cloning or DNA amplification were purchased from Sigma. The 38 bp DNA fragments were annealed in a final volume of 50 μL containing 10 μM of each strand (5'-(CAG)₁₂CA-3' and 5'-G(CTG)₁₂A-3') and ligated into pGEM-T Easy vectors (Promega). *Escherichia coli* JM109 competent cells (Toyobo) were transformed and cultured on LB plates with 50 μg/mL ampicillin and 32 mg X-gal/400 mg IPTG overnight at 37 °C. White colonies were identified by colony direct polymerase chain reaction (PCR) in 15 μL of $1 \times$ Go Taq Green Master Mix (Promega) containing 300 nM of primer set (T7: 5'-TAATACGACTCACTATAGG-3', sp6: 5'-ATTTAGGTGACACTATAGAATAC-3'). The reaction mixture was incubated at 95 °C for 3 min then followed by 34 incubation cycles of 95 °C for 30 sec, 55 °C for 30 sec, and 72 °C for 1 min with final extension step of 72 °C for 5 min. The appropriate colony was selected for transfer to 5 mL of LB medium with 50 μg/mL ampicillin and cultured overnight at 37 °C overnight. The plasmids with inserts were extracted using GenElute Plasmid Miniprep Kit (Sigma Aldrich) and their sequences were identified by 3130*xL* Genetic Analyzer (Applied Biosystems).

4.5. Preparation of 5'-Texas Red-labeled DNA fragments and high resolution gel electrophoresis

The 5'-Texas Red labeled DNA fragments containing the CAG/CTG repeat sequence was prepared by PCR using a primer set of 5'-Texas Red-labeled T7 and sp6 promoter primer or that of T7 and 5'-Texas Red-labeled sp6 promoter primer under the same condition and reagents as above. The PCR products were purified by GenElute PCR cleanup kit (Sigma Aldrich). The 5'-Texas Red-labeled DNA fragments were alkylated by various concentrations of alkylating polyamides in 10 μL of 5 mM sodium phosphate buffer (pH 7.0) containing 10% DMF at 23 °C, for 18 h. The reaction was quenched by the addition of calf thymus DNA (10 mg/mL, 1 μL) and the mixtures were heated at 95 °C for 10 min to cleave the DNA strands at the alkylating sites. The DNA was recovered by vacuum centrifugation. The pellet was dissolved in 7 μL of loading dye (formamide with New fuchsin), heated at 95 °C for 25 min, and then immediately chilled on ice. The 1.8 μL aliquot was subjected to electrophoresis on a 6% denaturing polyacrylamide gel using a HITACHI SQ5500-E DNA Sequencer.

4.6. *In vitro* transcription assays

DNA fragments including (CAG/CTG)₁₂ repeats for RNA transcription was prepared from the plasmids, which is described above, by PCR-amplification using 1X Go Taq Green Master Mix containing T7 promoter primer (5'-TAATACGACTCACTATAGG-3') and sp6 promoter primer (5'-ATTTAGGTGACACTATAGAATAC-3'). The reaction mixture was incubated at 95 °C for 3 min then followed by 30 incubation cycles of 95 °C for 30 sec, 50 °C for 30 sec, and 72 °C for 30 sec. After incubation, the PCR products were purified by QIAquick Gel Extraction Kit (Qiagen) and QIAquick Nucleotide Removal Kit (Qiagen). The obtained DNA fragments (8 nM) were incubated with various concentrations of **4** and **15** in a 5 mM sodium phosphate buffer (pH 7.0) containing 10 % DMF at the room temperature for 18 h. The alkylation reagents were quenched by the addition of DNA oligomer (200 μM, 2 μL).The reaction

mixtures were purified by illustra MicroSpin G-25 Columns (GE Healthcare) and then lyophilized for 3 h. The DNA fragments were transcribed with *in vitro* Transcription T7 Kit (Takara Bio Inc.). After transcriptions, the remaining DNA fragments were digested by RNase free DNase I (10 Units) included in the transcription kit for 30 min at 37 °C and the obtained RNAs were purified by RNeasy Mini Kit (Qiagen). Transcription products were analyzed by PAGE at 180 V for 80 min using 10 % TBE-Urea Gels containing 7 M urea (Invitrogen) and detected by SYBR Gold (Applied Biosystems). Before loading, samples were dissolved in TBE Urea Sample Buffer (Invitrogen), heated for 5 min at 95 °C, and then immediately cooled on ice for a few minutes. Low Range ssRNA Ladder (New England Biolabs) was used as a RNA marker and this was heated and then cooled similarly in loading buffer included in the ladder set. The bands were photographed and analyzed with Fluoro Image Analyzer FLA-3000GF (Fujifilm).

4.7. SPR assays

SPR experiments were performed on a Biacore X instrument. A biotinylated hairpin DNA was purchased from Sigma and the hairpin DNA is shown in Figure 5a. A streptavidin-functionalized SA sensor chip was purchased from Biacore. The biotinylated DNA is immobilized to the chip to obtain the desired immobilization level (approximately 1200 RU rise). SPR assays were carried out using HBS-EP buffer (10 mM HPES pH 7.4, 150 mM NaCl, 3mM 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 ml/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 or with drafting baseline) were used for fitting the sensorgrams of polyamide **12**, **13**, and **15** to give better fitting. For polyamide **11**, in order to represent dynamics in forming the complexes with two polyamides bound side by side in the same region, the sensorgram was fitted with the following model (eq 1):

$$A + B \underset{k_{d1}}{\overset{k_{a1}}{\rightleftarrows}} AB \qquad A + AB \underset{k_{d2}}{\overset{k_{a2}}{\rightleftarrows}} A_2B$$

(A: polyamide, B: DNA, AB, A_2B: polyamide–DNA complex)

$$dA/dt = k_t \cdot (Conc\ A) - (k_{a1} \cdot A \cdot B - k_{d1} \cdot AB) - (k_{a2} \cdot A \cdot AB - k_{d2} \cdot A_2B) \quad \text{(1A)}$$

$$dB/dt = -(k_{a1} \cdot A \cdot B - k_{d1} \cdot AB) \quad \text{(1B)}$$

$$dAB/dt = (k_{a1} \cdot A \cdot B - k_{d1} \cdot AB) - (k_{a2} \cdot A \cdot AB - k_{d2} \cdot A_2B) \quad \text{(1C)}$$

$$dA_2B/dt = (k_{a2} \cdot A \cdot AB - k_{d2} \cdot A_2B). \quad \text{(1D)}$$

A and *Conc A* correspond to the polyamide close to the surface which is capable of binding to the DNA ligand and the bulk concentration of the polyamide. k_{a1} and k_{a2} represent the rates of association of the first and second polyamide with the ligand, respectively; k_{d1} and k_{d2} represent the rates of dissociation of the polyamide from a 1:1 complex and 2:1 complex, respectively. k_t is the empirical constant for mass transfer. The closeness of the fit is described by the statistical value χ^2 (eq 2):

$$\chi^2 = \frac{\sum_1^n (r_f - r_x)^2}{n - p} \quad \text{(2)}$$

where is the fitted value at a given point, r_x is the experimental value at the same point, n is the number of data points, and p is the number of fitted parameters

Acknowledgments

Supplementary Material

Supplementary data associated with this article can be found, in the online version, at http

References and notes

1. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235.
2. Dervan, P. B.; Edelson, B. S. *Curr. Biol. Curr. Opin. Str. Biol.* **2003**, *13, 284*.
3. Blackledge, M. S.; Melander, C. *Bioorg. Med. Chem.* **2013**, *21*, 6101.
4. Bando, T.; Sugiyama, H. *Acc. Chem. Res.* **2006**, *39*, 935-944.
5. Mirkin, S. M. *Nature* **2007**, *447*, 932.
6. Cummings, C. J.; Zoghbi, H. Y. *Hum. Mol. Gen.* **2000**, *9*, 909.
7. Kremer, B.; Goldberg, P.; Andrew, S. E.; Theilmann, J.; Telenius, H.; Zeisler, J.; Squitieri, F.; Lin, B.; Bassett, A.; Almqvist, E.; Bird, T. D.; Hayden, M. R. *N. Engl. J.* **1994**, *330*, 1401.
8. Reddy, P. H.; Williams, M.; Charles, V.; Garrett, L.; Pike-Buchanan, L.; Whetsell Jr, W. O.; Miller, G.; Tagle, D. A. *Nature Genet.* **1998**, *20*, 198.
9. Lin, X.; Cummings, C. J.; Zoghbi, H. Y. *Neuron* **1999**, *24*, 499.
10. Philips, A. V.; Timchenko, L. T.; Cooper, T. A. *Science* **1998**, *280*, 737.
11. Mankodi, A.; Logigian, E.; Callahan, L.; McClain, C.; White, R.; Henderson, R.; Krym, M.; Thornton, C. *Science* **2000**, *289*, 1769.
12. Orengo, J. P.; Chambon, P.; Metzger, D.; Mosier, D. R.; Snipes, G. J.; Cooper, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2646.
13. Wheeler, T. M.; Sobczak, K.; Lueck, J. D.; Osborne, R. J.; Lin, X.; Dirksen, R. T.; Thornton. C. A. *Science* **2009**, *325*, 336.
14. Guan, L.; Disney, M. D. *Angew. Chem. Int. Ed.* **2013**, *52*, 10010.
15. Nakatani, K.; Hagihara, S.; Goto, Y.; Kobori, A.; Hagihara, M.; Hayashi, G.; Kyo, M.; Nomura, M.; Mishima, M.; Kojima, C. *Nat. Chem. Biol.* **2005**, *1*, 39.
16. Bando, T.; Fujimoto, J.; Minoshima, M.; Shinohara, K.; Sasaki, S.; Kashiwazaki, G.; Mizumura, M.; Sugiyama, H. *Bioorg. Med. Chem.* **2007**, *15*, 6937.
17. Fujimoto, J.; Bando, T.; Minoshima, M.; Uchida, S.; Iwasaki, M.; Shinohara, K.; Sugiyama, H. *Bioorg. Med. Chem.* **2008**, *16*, 5899.
18. Ferguson, L. R.; Liu, A. P.; Denny, W. A.; Cullinane, C.; Talarico, T.; Phillips, D. R. *Chem. Bio. Int.* **2000**, *126*, 15.
19. Hashem, V. I.; Sinden, R. R *Mutat Res.* **2002**, *508*, 107.
20. Hashem, V. I.; Pytlos, M. J.; Klysik, E. A.; Tsuji, K.; Khajav, M.; Ashizawa, T.; Sinden, R. R. *Nucleic Acids Res.* **2004**, *32*, 6334.
21. Takahashi, R.; Bando, T.; Sugiyama, H.; *Bioorg. Med. Chem.* **2003**, *11*, 2503.
22. Sasaki, S.; Bando, T.; Minoshima, M.; Shimizu, T.; Shinohara, K.; Takaoka, T.; Sugiyama, H. *J. Am. Chem. Soc.* **2006**, *128*, 12162.
23. Yamamoto, M.; Bando, T.; Kawamoto, Y.; Taylor, R. D.; Hashiya, K.; Sugiyama, H. *Bioconjugate. Chem.* **2014**, *14*, 552.
24. Taylor, R. D.; Asamitsu, S.; Takenaka, T.; Yamamoto, M.; Hashiya, Kawamoto, Y.; Bando, T.; Nagase, H.; Sugiyama, H. *Chem.–Eur. J.* **2013**, *19*, 1.
25. Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Org. Lett.* **2001**, *3,* 1201.
26. Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 5723.
27. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
28. Faguet, G. B. *J. Clin. Oncol.* **1994**, *12,* 1974.
29. CLL Trialists' Collaborative Group. *J. Natl. Cancer Inst.* **1999**, *91*, 861.
30. Wurtz, N. R.; Dervan. P. B. *Chem. Biol.* **2000**, *7,* 153.
31. Wang, Y. D.; Dziegielewski, J.; Wurtz, N. R.; Dziegielewska, B.; Dervan, P. B.; Beerman, T. A. *Nucleic Acids Res.* **2003**, *31,* 1208.
32. Wyatt, M. D.; Lee, M.; Garbhas, B. J.; Souhami, R. L.; Hartley, J. A. *Biochemistry* **1995**, *34*, 13034.
33. Sunters, A.; Springer, C. J.; Bagshawe, K. D.; Souhami, R. L.; Hartley, J. A. *Biochem. pharmacol.* **1992**, *44*, 59.

34. Minoshima, M.; Bando, T.; Shinohara, K.; Kashiwazaki, G.; Nishijima, S.; Sugiyama, H. *Bioorg. Med. Chem*. **2010**, *18*, 1236.
35. Trauger, J. W.; Baird, E. E.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6160.
36. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 6219.
37. Wang, C.C.C.; Ellervik, U.; Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 653.
38. Bando, T.; Minoshima, M.; Kashiwazaki, G.; Shinohara, K.; Sasaki, S.; Fujimoto, J.; Ohtsuki, A.; Murakami, M.; Nakazono, S.; Sugiyama, H. *Bioorg Med. Chem.* **2008**, *16*, 2286.
39. Minoshima, M.; Bando, T.; Sasaki, S.; Fujimoto, J.; Sugiyama, H. *Nucleic Acids Res.* **2008**, *36*, 2889.
40. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**,*116*, 1983.
41. Kawamoto, Y.; Bando, T.; Kamada, F.; Li, Y.; Hashiya, K.; Maeshima, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2013**, *135*, 16468.
42. Wetzler, M.; Wemmer, D. *Org. lett*. **2010**, *12*, 3488.
43. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998***, 120*, 1382.
44. Bando, T.; Narita, A.; Saito, I.; Sugiyama, H. *Chem.–Eur. J.* **2002**, *8*, 4781.
45. Sasaki, S.; Bando, T.; Minoshima, M.; Shinohara, K.; Sugiyama, H. *Chem.–Eur. J.* **2008**, *14*, 864.
46. Shinohara, K.; Sasaki, S.; Minoshima, M.; Bando, T.; Sugiyama, H. *Nucleic Acids Res.* **2006**, *34*, 1189.
47. Sugiyama, H.; Fujiwara, T.; Ura, A.; Tashiro, T.; Yamamoto, K.; Kawanishi, S.; Saito, I. *Chem. Res. Toxicol.* **1994**, *7*, 673.
48. Tao, Z.; Fujiwara, T.; Saito, I.; Sugiyama. H. *Angew. Chem. Int. Ed.* **1999**, *38*, 650.
49. Minoshima, M.; Bando, T.; Sasaki, S.; Shinohara. K.; Shimizu, T.; Fujimoto, J.; Sugiyama, H. *J. Am. Chem. Soc.* **2007**, *129*, 5384.
50. Oyoshi, T.; Kawakami, M.; Narita, A.; Bando, T.; Sugiyama, H. *J. Am. Chem. Soc.* **2003**, *125,* 4752.
51. Gottesfeld, J. M.; Belitsky, J. M.; Melander, C.; Dervan, P. B.; Luger, K. J. *Mol. Biol.* **2002**, *321*, 249.
52. Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J .Am. Chem. Soc.* **2002**, *124*, 2153.
53. Henry, J. A.; Le, N. M.; Nguyen, B.; Howard, C. M.; Bailey, S. L.; Horick, S. M.; Buchmueller, K. L.; Kotecha, M.; Hochhauser. D.; Hartley, J. A.; Wilson, W. D.; Lee, M. *Biochemistry* **2004**, *43*, 12249.
54. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.