Evaluation of the DNA Alkylation Property of a Chlorambucil-Conjugated Cyclic Pyrrole–Imidazole Polyamide

Yuki Hirose,[a] Kaori Hashiya,[a] Toshikazu Bando,*[a] and Hiroshi Sugiyama*[a, b]

[a] Y. Hirose, K. Hashiya, Dr. T. Bando, Prof. Dr. H. Sugiyama
Department of Chemistry, Graduate School of Science
Kyoto University
Kitashirakawa-oiwakecho, Sakyo-ku, Kyoto, 606-8502 (Japan)
E-mail: bando@kuchem.kyoto-u.ac.jp, hs@kuchem.kyoto-u.ac.jp

[b] Prof. Dr. H. Sugiyama
Institute for Integrated Cell-Material Science (iCeMS)
Kyoto University, Yoshida-ushinomiyacho, Sakyo-ku, Kyoto, 606-8501 (Japan)

Supporting information for this article is given via a link at the end of the document.

Abstract: Hairpin pyrrole–imidazole polyamides (hPIPs) and their chlorambucil (Chb) conjugates (hPIP–Chbs) can alkylate DNA in a sequence-specific manner, and have been studied as anticancer drugs. Here we conjugated Chb to a cyclic PIP (cPIP), which is known to have a higher binding affinity than the corresponding hPIP, and investigated the DNA alkylation properties of the cPIP–Chb using the optimized capillary electrophoresis method and conventional HPLC product analysis. cPIP–Chb conjugate **3** showed higher alkylation activity at its binding sites than did hPIP–Chb conjugates **1** and **2**. Subsequent HPLC analysis revealed that the alkylation site of conjugate **3**, which was identified by capillary electrophoresis, was reliable and that conjugate **3** alkylates the N3 position of adenine as do hPIP–Chbs. Moreover, conjugate **3** showed higher cytotoxicity against LNCaP prostate cancer cells than did conjugate **1** and cytotoxicity comparable to that of conjugate **2**. These results suggest that cPIP–Chbs could be novel DNA alkylating anticancer drugs.

Introduction

DNA alkylating agents such as duocarmycin A,[1,2] nitrogen mustard derivatives,[3,4] and CC-1065[5,6] form a covalent bond with DNA bases. Because of their ability to damage cells, these compounds have been studied as anticancer drugs.[7] However, their possible side effects derived from off-target alkylation remain a major problem.

To solve this problem, our group and others have tried to develop sequence-specific DNA alkylating agents using PIPs. PIPs are one type of well-studied DNA minor groove binders, which were developed by Dervan, Lown, and their coworkers based on the concept of Dickerson et al.[8–10] PIPs are principally composed of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) connected by amide bonds. The unique pairing rules enable PIPs to bind sequence-selectively to double-stranded B-DNAs: an antiparallel pairing of Py/Py recognizes an A/T or T/A base pair, while a Py/Im pair recognizes a C/G base pair. Such binding properties of PIPs motivated us to conjugate them to alkylating agents such as chlorambucil (Chb) to improve the sequence specificity of its alkylation. Hairpin-shaped PIP–Chb conjugates (hPIP–Chbs) have been widely studied and have shown good cell permeability, suppression of gene expression, and cytotoxicity against cancer cells.[11–14] Moreover, some of the hPIP–Chbs have shown anticancer effects in tumor xenograft mouse models.[15,16]

In the field of peptide drug discovery, cyclic peptides are known to exhibit higher *in vivo* stability, target binding ability, and membrane permeability than acyclic peptides.[17] Similarly, cyclic PIPs (cPIPs) have been reported to exhibit higher binding affinity to target sequences than do hPIPs.[18–20] Thus, cPIP–Chbs are expected to exhibit a higher capacity for sequence-specific alkylation than conventional hPIP–Chbs. However, because of the poor efficiency of cPIP synthesis,[19,20] research has not progressed, and cPIPs conjugated to alkylating agents have not been reported previously. Because cPIPs can now be synthesized more easily,[21] we decided to synthesize a cPIP–Chb (**3**) and compare its alkylation ability with that of hPIP–Chbs (**1** and **2**). We propose this compound as a new sequence-specific DNA alkylating agent (Figure 1).

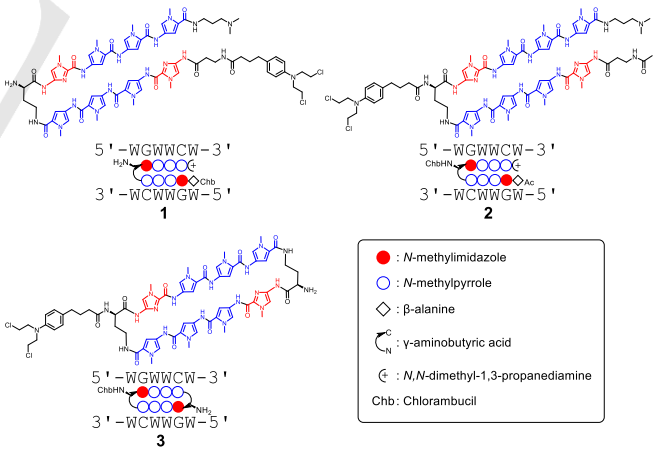


Figure 1. Chemical structures and ball-and-stick notations of hPIP–Chbs (**1** and **2**) and a cPIP–Chb (**3**). A Py/Py pair recognizes a W/W (W = A or T) base pair and a Py/Im pair recognized a C/G pair.

To determine the alkylation ability of PIP–Chbs, previous studies have identified DNA cleavage sites by slab-gel electrophoresis.[11–16] By contrast, our group has observed DNA cleavage at higher resolution using capillary electrophoresis and we have investigated nucleosome accessibility.[22] In this study, we optimized this high-resolution method to observe the DNA sites alkylated by PIP–Chbs and found that the cPIP–Chb **3** produced

a larger peak in the target sequence than did hPIP–Chbs **1** and **2**. Furthermore, HPLC and MALDI-TOF MS studies suggested that **3** alkylates the N3 position of the purine base adjacent to the target sequence. Finally, a cytotoxicity assay was performed that demonstrated that conjugate **3** was sufficiently toxic to LNCaP prostate cancer cells. These results suggest that cPIP–Chbs can be utilized as a novel sequence-specific DNA alkylating agent.

Results and Discussion

Design and Synthesis of Conjugates 1–3

Compounds **1**–**3** were designed to target the sequence 5′-WGWWCW-3′ (W = A or T). Conjugate **1** has the Chb moiety at the N-terminus and conjugates **2** and **3** have the Chb on the γ-aminobutyric acid turn (γ-turn). The (*R*)-α-substituted γ-turn is present in each compound, and this modification is reported to enhance the binding capacity and binding orientation specificity of PIPs.[21, 23] Each PIP was synthesized according to the synthetic procedures described in the Experimental section. Briefly, after machine-assisted solid-phase synthesis, cleavage from the resin was carried out and the final product was obtained by Chb-coupling and deprotection for conjugate **1**, deprotection and Chb-coupling for conjugate **2**, and cyclization, deprotection (Boc), Chb-coupling, and deprotection (Cbz) for conjugate **3**. The resulting crude products were purified by reverse-phase HPLC. The HPLC profiles and mass spectra of conjugates **1**−**3** are shown in Figures S1–S3.

Determination of Alkylation Sites by Capillary Electrophoresis

The alkylation sites of compounds **1**–**3** were identified by capillary electrophoresis using the DNA described in a previous report,[22] which is 383 bp long and has its 5′ end fluorescently labeled (Figure 2). There are two binding sites in the sequence, which are underlined in Figure 2.

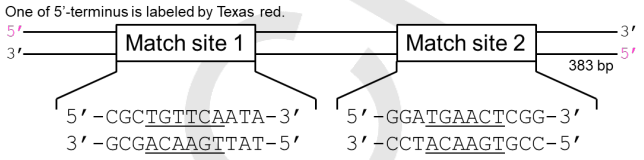


Figure 2. Diagram of the 383-bp DNA fragment. The binding sites of PIPs are underlined.

The alkylating reaction was carried out at 23 °C for 18 h. The samples were heated at 95 °C for 10 min and then treated with piperidine at 95 °C for 25 min. This treatment produces DNA fragments with 3'-phosphate termini. These DNA fragments were purified by spin column and lyophilized samples were diluted by Hi-Di formamide. Then, they were subjected to capillary electrophoresis and the results shown in Figure 3a were obtained. The determination of alkylation sites was performed by mixing the alkylated fragments with DNA markers prepared using the

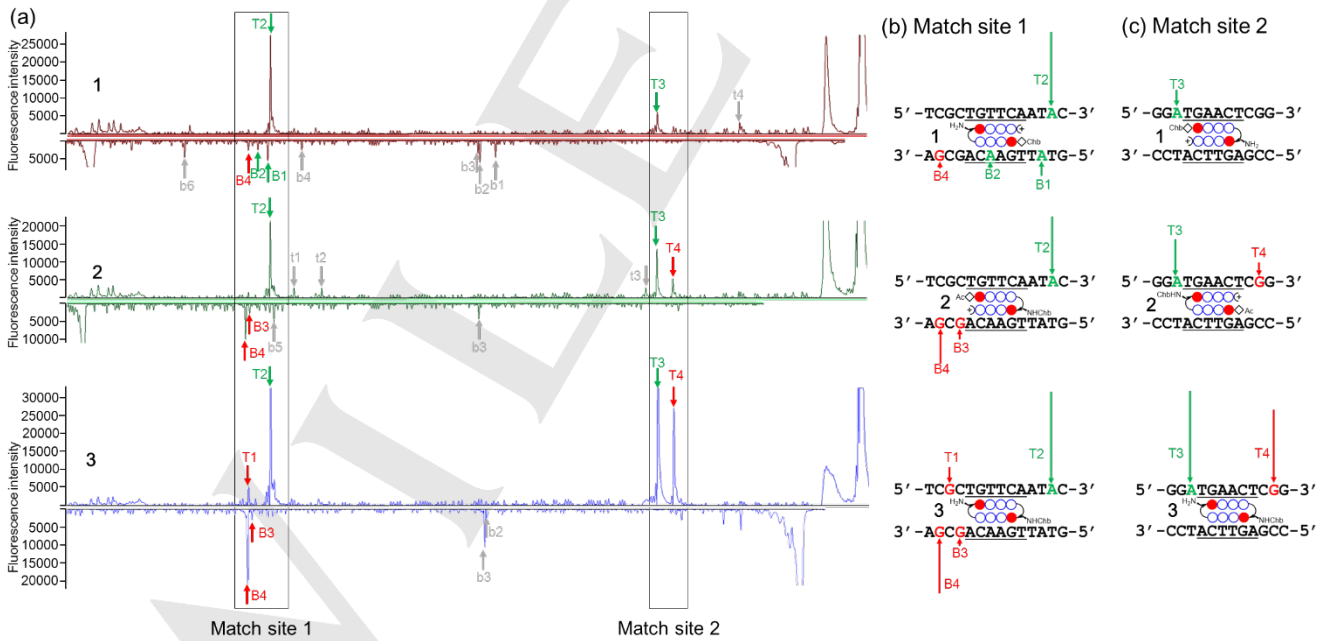


Figure 3. (a) Thermally induced strand cleavage of 5′-Texas Red-labeled 383-bp DNA fragments by conjugates **1**–**3**. Alkylation sites are indicated by arrows. (b, c) Schematic representations of the recognition of alkylation sites by conjugates **1**–**3** at match site 1 (b) and match site 2 (c). The length of each arrow in (b) and (c) corresponds to the ratio of the fluorescence intensity of the alkylation peak at each site. Peaks labeled as "T" and "B" indicate alkylation at the target sites of conjugates **1**–**3** on the "top strand" and "bottom strand," respectively. Peaks labeled as "t" and "b" indicate alkylation at nontarget sites of conjugates **1**–**3** on the "top strand" and "bottom strand," respectively.

Maxam–Gilbert sequencing method[24] with the results shown in Figures 3b and 3c and Figures S4–S6. Currently, the Sanger sequencing method[25] is most commonly used, but the difference in the 3′-termini of alkylated samples (3′-phosphate termini)[26] and Sanger sequencing samples (3′-dideoxyribose termini) prevented us from identifying alkylation sites accurately using that method. Therefore, we used the Maxam–Gilbert method, which can produce the same 3′-termini as in the alkylation samples (Figure S7). The results shown in Figure 3 demonstrate that each PIP–Chb efficiently alkylated adenines or guanines adjacent to their target sites, although they also produced some off-target alkylations. The data also demonstrated that the alkylation sites and activities differ depending on the position of Chb (N-terminus or γ-turn)[13] and the structure (hairpin or cyclic). Conjugate **1**, which has the Chb on its N-terminus, strongly alkylated the adenine opposite to its Chb moiety, whereas conjugates **2** and **3**, which have the Chb on their γ-turn units, effectively alkylated the adenines and guanines on both strands. Focusing on the structure of the conjugates, cyclic compound **3** produced larger alkylation peaks than did the hairpin compounds **1** and **2**. This result suggested that the cPIP–Chb conjugates have higher alkylation activity than hPIP–Chbs.

HPLC Product Analysis of DNA Alkylation by cPIP–Chb 3

Because this study is the first to report alkylation by a cPIP–Chb conjugate, we undertook detailed investigations into the alkylation site of compound **3** by analyzing the HPLC product.[14, 26, 27] We designed a 16-bp duplex oligonucleotide (5′-AGGGATGATCATCCCT-3′) based on the results of capillary electrophoresis. HPLC analysis revealed that peaks derived from single-stranded (ss) DNA and conjugate **3** diminished and new peaks appeared at about 14 min and 20 min after alkylation for 18 h at 23 °C (Figure 4a). Based on the previous report[14], in which the hPIP–Chb–ssDNA adduct appeared at around 20 min under similar HPLC conditions, and the MALDI-TOF MS result for the peak at 14 min (Figure S8), the peaks at 14 min and 20 min were thought to be derived from the double-stranded (ds) DNA–3 adduct, in which one strand is alkylated by conjugate 3, and from the ssDNA–3 adduct, respectively.

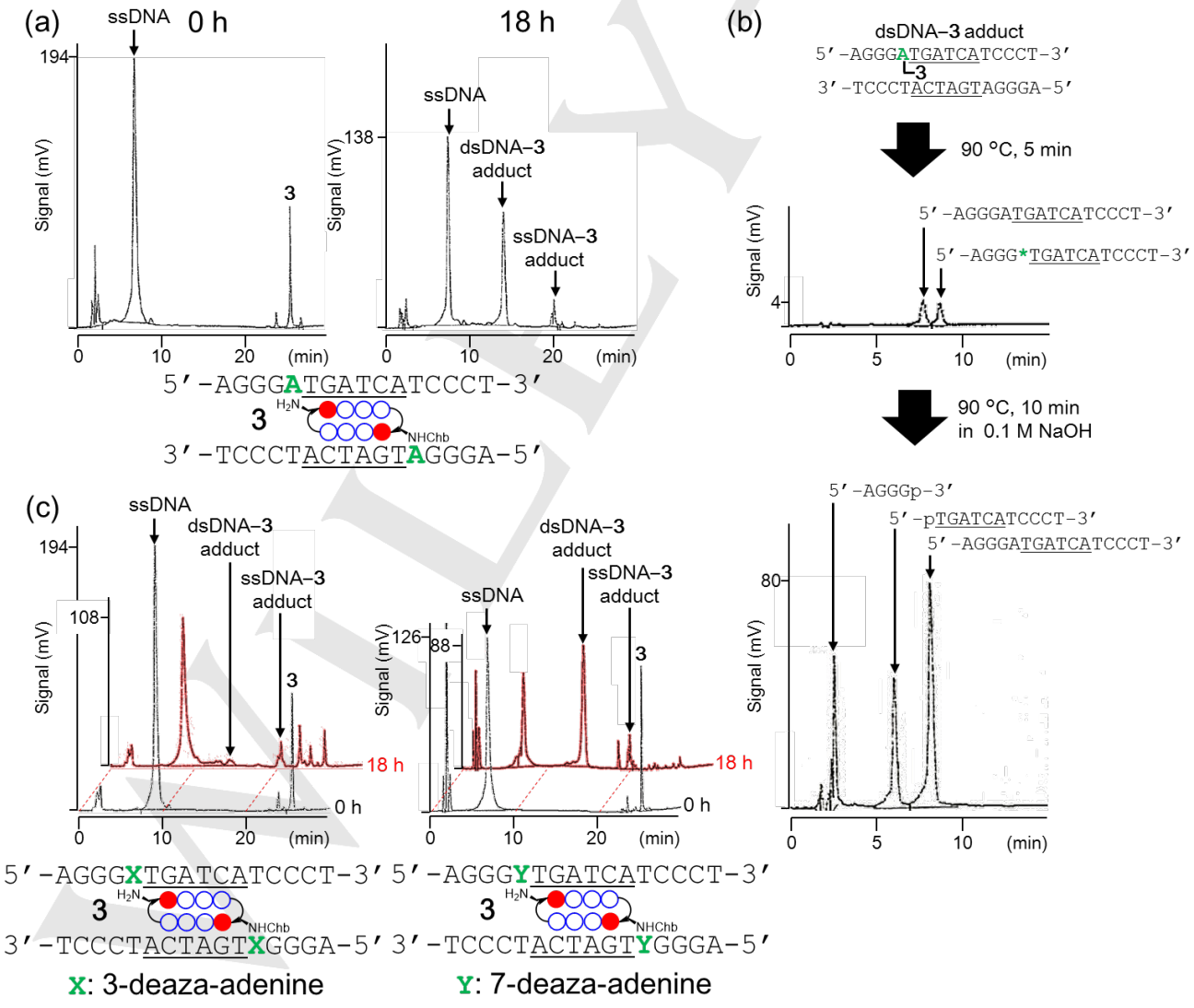


Figure 4. (a) HPLC analysis of the alkylation of a short DNA oligomer with conjugate **3**. (b) Determination of the alkylation sites by heat and alkali treatment. (c) HPLC analysis of the alkylation of the short DNA oligomer containing 3-deaza-adenine (X) and 7-deaza-adenine (Y) by conjugate **3**.

The fraction containing the dsDNA–**3** adduct was heated at 90 °C for 5 min and a part of this sample was subjected to HPLC analysis. The result of this HPLC analysis and MALDI-TOF MS results for each peak demonstrated that the alkylated strand was converted to a strand with an abasic site produced by the depurination of an adenine (upper HPLC chart in Figure 4b and Figure S9). Then, the remainder of the heated sample was lyophilized and heated at 90 °C for 10 min in 0.1 M NaOH solution. As a result of this treatment, the abasic site-containing strand was cleaved and short fragments appeared (lower HPLC chart in Figure 4b and Figure S10). We also detected cPIP–Chb–A or –G adducts after this treatment (Figure S11). The results shown in Figure 4b suggested that the alkylation site was consistent with the result obtained by the capillary electrophoresis.

Moreover, we confirmed that alkylation by conjugate 3 occurs at the N3 position of the adenine. We prepared oligonucleotides containing 3-deaza-adenine (X) or 7-deaza-adenine (Y). The HPLC analysis using the X-containing sequence (5′-AGGGXTGATCATCCCT-3′) shown in Figure 4c indicated that the new peak that appeared around 14 min became much smaller than the corresponding peak in Figure 4a, although the same experiment with the Y-containing sequence (5′-AGGGYTGATCATCCCT-3′) showed the formation of the new peak of similar intensity to that in Figure 4a. These results indicate that cPIP–Chb **3** alkylates at the N3 position of adenine as hPIP–Chbs.[14]

Cytotoxicity Assay of Conjugates 1–3 Against LNCaP Prostate Cancer Cells

Finally, to evaluate the potential of conjugate **3** as an anticancer drug, we compared the cytotoxicity of conjugates **1**–**3** against LNCaP prostate cancer cells. These cells were treated with various concentrations of each conjugate for 72 h. Cell viability and IC_{50} values are shown in Figure 5. These results suggested that the cytotoxicity of conjugates **2** and **3** was at least six to eight times higher than that of conjugate **1**, and that conjugate **2** was slightly more toxic than conjugate **3**. The results of the cytotoxicity assay were not completely consistent with the results of capillary electrophoresis, in which conjugate **3** produced larger peaks than the other two conjugates. This might be caused by factors other than alkylation properties *in vitro*, such as cell permeabilities and stabilities in the cellular condition. Even with such an inconsistency, the fact that the toxicity of the cPIP–Chb was comparable to that of hPIP–Chbs is a good indication to warrant further studies in mice and other animals.

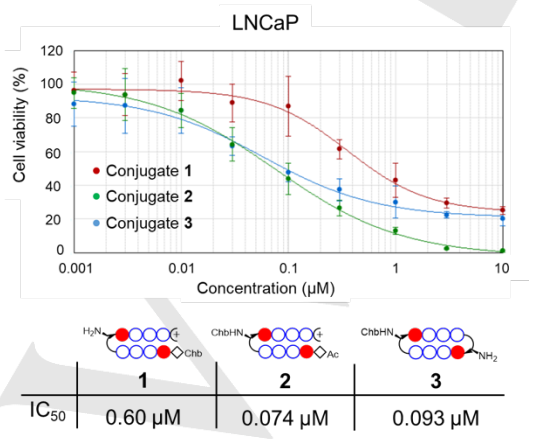


Figure 5. Cell viability and IC_{50} values of conjugates **1**–**3** against LNCaP cells (n = 5). Error bars indicate standard deviations.

Conclusion

In this report, we investigated the alkylation activity of a cPIP–Chb conjugate, **3**. The results of the newly optimized capillary electrophoresis study with 383-bp DNA fragments revealed that conjugate **3** showed higher alkylation activity at its target sites than did hPIP-Chb conjugates **1** and **2**. Furthermore, the HPLC analysis with 16-bp oligonucleotides suggested that the alkylation sites determined by capillary electrophoresis were reliable and that conjugate **3** alkylated the same position on adenine as previously reported for hPIP–Chbs. Additionally, conjugate **3** showed comparable cytotoxicity against cancer cells to that of conjugates **1** and **2**. These results suggested that cPIP–Chbs are new candidates for sequence-specific DNA alkylating drugs.

Experimental Section

General procedure

Reagents and solvents were purchased from standard suppliers and used without further purification. Automated polyamide synthesis was performed on a PSSM-8 system (Shimadzu). HPLC analysis of each synthesized compound was performed on a Jasco Engineering PU-2089 plus series system using a Chemcobond 5-ODS-H 4.6 mm × 150 mm column (Chemco Plus Scientific) 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 0-100% acetonitrile in 40 min with detection at 254 nm. Collected fractions were analyzed by MALDI-TOF-MS Microflex-KS II (Bruker). HPLC purification was carried out by Jasco engineering PU-2080 plus series using a COSMOSIL 150 × 10 mm $5C_{18}$-MS-II Packed Column (Nacalai Tesque Inc.) in 0.1% TFA in water with acetonitrile as the eluent at a flow rate of 3.8 mL/min and a linear gradient elution of +1%/min acetonitrile for 20 min with detection at 254 nm. The concentration of PIP solution in DMF was prepared based on the following formula.

ε= 9900 × (sum number of Py and Im)

Abs = εcl

ε, Abs, c, and l are molar extinction coefficients of PIPs in DMSO solution at around 302 nm, absorbance at 302 nm measured by Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific), molar concentration, and the path length, respectively.

Fmoc Solid-Phase Synthesis of PIPs

The solid-phase synthesis of each PIP was performed as described previously[28]. The building blocks used in this study were FmocHN-Py-CO_2H, FmocHN-Im-CO_2H, FmocHN-β-alanine-CO_2H, FmocHN-Cbz-(*R*)-α-aminobutyric acid, FmocHN-Boc-(*R*)-α-aminobutyric acid. Each of them was introduced sequentially to FmocHN-Py-trityl resin (for **3**) or FmocHN-Py-oxime resin (for **1** and **2**). The terminal Im-NH_2 of **2** was capped by an Ac-group.

Synthesis of Chb-HN-β-ImPyPyPy-(*R*)$^{α\text{-}NH2}$γ-ImPyPyPy-Dp (1)

By using 80 mg of FmocHN-Py-oxime resin (0.426 mmol/g) and proper building blocks, H_2N-β-ImPyPyPy-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-oxime resin was prepared by solid-phase synthesis. Following cleavage reaction with 1000 µL of *N, N*-dimethyl-1,3-propanediamine at 55 °C for 3 hours, filtration and Et_2O precipitation gave 61.9 mg of H_2N-β-ImPyPyPy-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-Dp (analytical HPLC: t_R = 16.7 min. MALDI-TOF MS: *m/z* calcd for $C_{63}H_{82}N_{23}O_{12}^+$[M+H]$^+$1352.59, found; 1352.66). A part of this crude polyamide (9.8 mg) dissolved in dimethylformamide (DMF, 0.1 mg/µL) was mixed with chlorambucil (Chb, 2 equiv.), benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP, 4 equiv.) and *N, N*-diisopropylethylamine (DIEA, 12 equiv.) and stirred at room temperature for 3.5 hours. After Et_2O precipitation, 22.4 mg of Chb-HN-β-ImPyPyPy-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-Dp was obtained (analytical HPLC: t_R = 23.1 min. MALDI-TOF MS: *m/z* calcd for $C_{77}H_{99}Cl_2N_{24}O_{13}^+$[M+H]$^+$1637.71, found; 1638.18). The Boc protecting group on the γ-turn was deprotected by adding 500 µL of dichloromethane (DCM) and 200 µL of trifluoroacetic acid (TFA) and stirring at room temperature for 30 min. After Et_2O precipitation, 13.8 mg of Chb-HN-β-ImPyPyPy-(*R*)$^{α\text{-}NH2}$γ-ImPyPyPy-Dp (**1**) was obtained (analytical HPLC: t_R = 19.6 min. MALDI-TOF MS: *m/z* calcd for $C_{72}H_{91}Cl_2N_{24}O_{11}^+$[M+H]$^+$1537.66, found; 1537.99). This crude sample of **1** was dissolved in DMF and purified by HPLC to obtain pure **1** as an off-white powder (3.7 mg, 2.4 µmol, 45 % yield for 12 steps, analytical HPLC: t_R = 19.6 min. MALDI-TOF MS: *m/z* calcd for $C_{72}H_{91}Cl_2N_{24}O_{11}^+$[M+H]$^+$1537.66, found; 1537.71). For the next experiments, this purified sample was used.

Synthesis of Ac-HN-β-ImPyPyPy-(*R*)$^{α\text{-}NHChb}$γ-ImPyPyPy-Dp (2)

By using 81 mg of FmocHN-Py-oxime resin (0.402 mmol/g) and proper building blocks, Ac-β-ImPyPyPy-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-oxime resin was prepared by solid-phase synthesis. The same cleavage reaction as **1** gave 39.5 mg of Ac-β-ImPyPyPy-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-Dp (analytical HPLC: t_R = 17.7 min. MALDI-TOF MS: *m/z* calcd for $C_{65}H_{84}N_{23}O_{13}^+$[M+H]$^+$1394.65, found; 1395.09). After deprotection of the Boc group on the γ-turn, 58.3 mg of Ac-β-ImPyPyPy-(*R*)$^{α\text{-}NH2}$γ-ImPyPyPy-Dp was obtained (analytical HPLC: t_R = 14.3 min. MALDI-TOF MS: *m/z* calcd for $C_{60}H_{76}N_{23}O_{11}^+$[M+H]$^+$1294.60, found; 1295.26). Then Chb was conjugated to a part of this compound (8.0 mg) by the same procedure as **1** for 3 hours and 10.2 mg of Ac-HN-β-ImPyPyPy-(*R*)$^{α\text{-}NHChb}$γ-ImPyPyPy-Dp (**2**) was obtained (analytical HPLC: t_R = 22.0 min. MALDI-TOF MS: *m/z* calcd for $C_{74}H_{93}Cl_2N_{24}O_{12}^+$[M+H]$^+$1579.67, found; 1580.71). This crude sample of **2** was dissolved in DMF and purified by HPLC to obtain pure **2** as an off-white powder (2.5 mg, 1.6 µmol, 35 % yield for 13 steps, analytical HPLC: t_R = 21.9 min. MALDI-TOF MS: *m/z* calcd for $C_{74}H_{93}Cl_2N_{24}O_{12}^+$[M+H]$^+$1579.67, found; 1579.78). For the next experiments, this purified sample was used.

Synthesis of cyclo-(-ImPyPyPy-(*R*)$^{α\text{-}NH2}$γ-ImPyPyPy-(*R*)$^{α\text{-}NHChb}$γ-) (3)

By using 82 mg of FmocHN-Py-trityl resin (0.426 mmol/g) and proper building blocks, H_2N-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-(*R*)$^{α\text{-}NHCbz}$γ-ImPyPyPy-trityl resin was prepared by solid-phase synthesis. Following cleavage reaction with 30% hexafluoroisopropanol (HFIP) in DCM at room temperature for 3 hours, filtration and Et_2O precipitation gave 41.8 mg of H_2N-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-(*R*)$^{α\text{-}NHCbz}$γ-ImPyPyPy-COOH (analytical HPLC: t_R = 20.2 min. MALDI-TOF MS: *m/z* calcd for $C_{67}H_{79}N_{22}O_{15}^+$[M+H]$^+$1431.60, found; 1431.98). This crude polyamide was dissolved in DMF to obtain 1 mM solution and stirred with pentafluorophenyl diphenylphosphinate (FDPP, 3 equiv.) and DIEA (6 equiv.) for 39 hours. The solvent was evaporated and dried in vacuo. The residue was dissolved in the minimum volume of MeOH/DCM 1:1 mixture and powdered in Et_2O. After vacuum drying, 42.7 mg of cyclo-(-ImPyPyPy-(*R*)$^{α\text{-}NHBoc}$γ-ImPyPyPy-(*R*)$^{α\text{-}NHCbz}$γ-) was obtained (analytical HPLC: t_R = 21.4 min. MALDI-TOF MS: *m/z* calcd for $C_{67}H_{77}N_{22}O_{14}^+$[M+H]$^+$1413.59, found; 1413.62). Then the Boc deprotection same as compounds **1** or **2** was performed and 39.4 mg of cyclo-(-ImPyPyPy-(*R*)$^{α\text{-}NH2}$γ-ImPyPyPy-(*R*)$^{α\text{-}NHCbz}$γ-) was obtained (analytical HPLC: t_R = 18.4 min. MALDI-TOF MS: *m/z* calcd for $C_{62}H_{69}N_{22}O_{12}^+$[M+H]$^+$1313.54, found; 1313.60). After that Chb was conjugated to a part of this compound (10.1 mg) by the same procedure as **1** and **2** for 1.5 hours and 14.7 mg of cyclo-(-ImPyPyPy-(*R*)$^{α\text{-}NHChb}$γ-ImPyPyPy-(*R*)$^{α\text{-}NHCbz}$γ-) was obtained (analytical HPLC: t_R = 25.4 min. MALDI-TOF MS: *m/z* calcd for $C_{76}H_{86}Cl_2N_{23}O_{13}^+$[M+H]$^+$1598.61, found; 1599.00). Finally, the compound was treated with 50 µL of trifluoromethanesulfonic acid (TfOH) in 500 µL of TFA to deprotect the Cbz protecting group on γ-turn. After stirring for 4 min at room temperature, the mixture was dropped into cold Et_2O (-79 °C). The precipitate was dried *in vacuo* and 15.4 mg of cyclo-(-ImPyPyPy-(*R*)$^{α\text{-}NHChb}$γ-ImPyPyPy-(*R*)$^{α\text{-}NH2}$γ-) (**3**) was obtained (analytical HPLC: t_R = 22.0 min. MALDI-TOF MS: *m/z* calcd for $C_{68}H_{80}Cl_2N_{23}O_{11}^+$[M+H]$^+$1464.57, found; 1465.04). This crude sample of **3** was dissolved in DMF and purified by HPLC to obtain pure **3** as an off-white powder (1.6 mg, 1.1 µmol, 12 % yield for 14 steps, analytical HPLC: t_R = 21.5 min. MALDI-TOF MS: *m/z* calcd for $C_{68}H_{80}Cl_2N_{23}O_{11}^+$[M+H]$^+$1464.57, found; 1464.58). For the next experiments, this purified sample was used.

Preparation of 5'–Texas Red–labeled 383 bp DNA fragment

The 383 bp dsDNA was prepared in our previous report. By using this dsDNA fragment for template DNA and the 5'-Texas Red-labeled ssDNA (5'-Texas Red-TAATACGACTCACTATAGGGCGAATTCGAGCTCGGTAC-3') and the non-labeled ssDNA (5'-ATTTAGGTGACACTATAGAATACTCAAGCTTGCATGCCTGC-3') for primer, 5'-Texas Red-labeled 383 bp DNA fragment was prepared by PCR amplification (95 °C for 5 min followed by 30 thermal cycles of 95 °C for 20 s, 55 °C for 30 s and 72 °C for 1 min with final extension step of 72 °C 7 min). For preparing the opposite strand labeled DNA fragment, the 5'-Texas Red-labeled ssDNA (5'-Texas Red-ATTTAGGTGACACTATAGAATAC-3') and the non-labeled ssDNA (5'-TAATACGACTCACTATAGGGCGAATTCGAGCTCGGTAC-3') for primer were used. The PCR amplification reaction was performed by using GOTaq® Green Master Mix. Then amplified DNA was purified by Wizard® SV Gel and PCR Clean-Up System (Promega) and the final concentration of DNA was measured by Nano Drop One (Thermo Fisher Scientific).

Alkylation by PIP and capillary electrophoresis of alkylated DNA

The 5'-Texas Red-labeled DNA fragments (50 nM) were alkylated by conjugates **1**–**3** in 50 µL of 5 mM sodium phosphate buffer (pH 7.0) containing 10% DMF at 23 °C for 18 hours. Alkylated DNA fragments were heated at 95 °C for 10 min. 5 µL of piperidine was added to each sample and heated at 95 °C for 25 min. The samples were purified by kit and lyophilized. Dried pellets were dissolved by Hi-Di Formamide™ 15 µL. These mixtures were heated at 95 °C for 5 min and immediately cooled to room temperature and subjected to Kyoto University Medical Research Support Center for capillary electrophoresis. Capillary electrophoresis was performed on Applied Biosystem® 3500xl Genetic Analyzer (Thermo Fisher Scientific) in standard run mode. Data was viewed and analyzed by sequence-scanner 2 software.

Preparation of sequencing sample by Maxam-Gilbert method (A+G and T+C reaction)[29]

10 µL of 383 bp dsDNA fragment (125 nM) was prepared and 22 µL of 99% formic acid was added for A+G specific cleavage reaction. And 5 µL of 383 bp dsDNA fragment (500 nM) was prepared and 45 µL of hydrazine

monohydrate was added for T+C specific cleavage reaction. These two mixtures were incubated at 20 °C for 5 min and 7 min, respectively. After incubation 200 µL of acetate-RNA mix (0.3 M of sodium acetate (pH 5.2), 0.1 mM of EDTA, 0.4 mg/mL of yeast tRNA (Thermo Fisher Scientific)) was added to each mixture at 0 °C. Then 750 µL of EtOH (99.5) (-30 °C) was added and precipitated the DNA at -80 °C for 15 min. After precipitation, the sample was centrifuged for 15 min at 13200 r.p.m. and the supernatant was discarded. Then the pellet was dissolved in 200 µL of 0.3 M sodium acetate and 600 µL of EtOH (99.5) was added. After precipitation and centrifugation again, the pellet was washed by 1000 µL of 80% EtOH. After centrifugation and discarding the supernatant, the pellet was washed by 1000 µL of 96% EtOH and centrifuged. After discarding the supernatant, the sample was dried *in vacuo*. The dry pellet was dissolved in 100 µL of 10% piperidine and heated at 95 °C for 30 min. After heat treatment, the sample was lyophilized until a dry pellet was obtained. The dry pellet was dissolved in 50 µL of H_2O. This sample was lyophilized again until a dry pellet was obtained. The dry pellet was dissolved in 10 µL of H_2O. This sample was lyophilized again until a dry pellet was obtained. The dry pellet was used for capillary electrophoresis.

Preparation of sequencing sample by Sanger method

ddNTP mixtures, containing 300 µM dATP, dTTP, dGTP and dCTP and 3 µM ddNTP, were prepared by mixing dNTP mixture (Takara) and ddNTP (Roche) for ddATP, ddTTP, ddGTP and ddCTP. And 18 µL of the mixture containing 1.0 µL of ~20 ng/µL non-labeled 383 bp DNA template, 2.2 µL of Thermo Sequenase™ Reaction Buffer (Thermo Fisher Scientific), 2.0 µL of 1 µM labeled primer (5'-Texas Red-TAATACGACTCACTATAGGGCGAATTCGAGCTCGGTAC-3' or 5'-Texas Red-ATTTAGGTGACACTATAGAATAC-3') and 1.0 µL of Thermo Sequense™ DNA polymerase w/Pyrophosphatase (Thermo Fisher Scientific) and 11.8 µL of nuclease-free H_2O was prepared. Then the solution was divided into 4 samples (4 µL × 4) and 1 µL of the ddNTP mix was added to each sample. Various length of DNA fragments with dideoxyribose on their 3'-terminus were prepared by PCR amplification (95 °C for 3 min followed by 34 thermal cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min with final extension step of 72 °C 5 min). After amplification, the samples were lyophilized and the dry pellets were used for capillary electrophoresis.

Capillary electrophoresis of alkylated DNA with sequencing sample

The dry pellets of DNA fragments prepared by the Maxam-Gilbert or the Sanger sequencing method were dissolved by 30 µL of 80% formamide in 1 × TBE. Each of these solutions was added to the pellet of alkylated DNA prepared by the method mentioned above. These mixtures were heated at 95 °C for 5 min and immediately cooled to room temperature and subjected to Kyoto University Medical Research Support Center for capillary electrophoresis. Before the capillary electrophoresis by Applied Biosystem® 3500xl Genetic Analyzer, samples were purified by gel filtration spin column. Data were viewed and analyzed by sequence-scanner 2 software.

HPLC and MALDI-TOF MS analysis of alkylation by compound 3

All DNA fragments were purchased from Sigma. Palindromic 16 bp DNA was annealed in a final volume of 200 µL containing 40 µM of dsDNA (5'-AGGGATGATCATCCCT-3'/3'-TCCCTACTAGTAGGGA-5'). Annealing was performed by heating at 65 °C for 10 min and then the sample was gradually cooled to room temperature. The reaction mixture containing conjugate **3** (30 µM) and DNA annealing solution (30 µM) in 5 mM sodium phosphate buffer (pH 7.0) was incubated at 23 °C. The progress of the reaction was monitored by HPLC just after starting the reaction and after 18 hours using a Chemcobond 5-ODS-H column (4.6 × 150 mm). The analysis was performed with a mobile phase containing 50 mM Triethylamine acetate (TEAA) in gradient combination with acetonitrile (9-12%: 0-15 min, 12-100%: 15-30 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm. Each of the peaks around 8 min, 14 min, 20 min were collected and analyzed by MALDI-TOF MS after lyophilization. The peak at about 14 min was used for the following heat treatments.

After alkylation for 18 hours, the dsDNA-**3** peak at about 14 min was collected and heated at 90 °C for 5 min. The heated solution was analyzed by HPLC. The analysis was performed with a mobile phase containing 50 mM TEAA in gradient combination with acetonitrile (9-12%: 0-15 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm. Then two peaks around 8 min and 9 min were analyzed by MALDI-TOF MS after lyophilization.

For hot alkali treatment, the dsDNA-**3** peak at about 14 min was collected, heated at 90 °C for 5 min, and lyophilized. The dry pellet was dissolved in 0.1 M NaOH 45 µL and heated at 90 °C for 10 min. After neutralization by adding 0.1 M AcOH, the solution was analyzed by HPLC. The analysis was performed with a mobile phase containing 50 mM TEAA in gradient combination with acetonitrile (9-12%: 0-15 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm. Then three peaks around 3 min, 8 min, and 9 min were analyzed by MALDI-TOF MS after lyophilization. For detecting cPIP-Chb-base adducts, the same treatments were performed and the last HPLC analysis was performed with mobile phase containing 50 mM TEAA in gradient combination with acetonitrile (9-12%: 0-15 min, 12-100%: 15-30 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm. Then the peak around 25 min was analyzed by MALDI-TOF MS after lyophilization.

HPLC analysis by using 3-deaza-adenine containing oligonucleotides

3-Deaza-dA-CE and 7-Deaza-dA-CE Phosphoramidite were purchased from Glen Research and the synthesis of the DNA fragment containing 3-deaza-dA was performed on M-2-MX_NP DNA synthesizer (Nihon Techno Service) and the synthesis of the DNA fragment containing 7-deaza-dA was performed on 3400 DNA synthesizer (Applied Biosystems). Synthesized palindromic 16 bp DNA was annealed in a final volume of 200 µL containing 40 µM of dsDNA (5'-AGGG[X or Y]TGATCATCCCT-3'/3'-TCCCTACTAGT[X or Y]GGGA-5', X: 3-deaza-A, Y: 7-deaza-A). Annealing was performed by heating at 65 °C for 10 min and then the sample was gradually cooled to room temperature. The reaction mixture containing conjugate **3** (30 µM) and DNA annealing solution (30 µM) in 5 mM sodium phosphate buffer (pH 7.0) was incubated at 23 °C. The progress of the reaction was monitored by HPLC just after starting the reaction and after 18 hours under the same condition as above.

Cytotoxicity assay

LNCaP cells were purchased from American Type Culture Collection (ATCC, CRL-1740™) and cultured in Roswell Park Memorial Institute 1640 Medium (RPMI-1640, Wako, 187-02705) supplemented with 10% FBS (Sigma) at 37 °C under 5% CO_2. Cytotoxicity of each conjugate was determined by colorimetric assay using cell count reagent SF (nacalai tesque Inc.). Cells were plated on 96-well plates in 100 µL of the medium at a density of 5 × 10^3 cells/well. After 48 h, the medium was renewed and PIP-Chb conjugates were added to give specified concentrations in 0.2% DMSO. After PIP-Chbs treatment for 72 h, 10 µL of cell count reagent SF was added to each well. Absorbance at 450 nm and 650 nm was measured by using a SpectraMax M2e microplate reader (Molecular Devices) after 1h incubation. Cell viability was calculated by following formula: {$Abs_{450nm-650nm}$(treated) - $Abs_{450nm-650nm}$(blank)} / {$Abs_{450nm-650nm}$(0.2% DMSO) - $Abs_{450nm-650nm}$(blank)}. The IC_{50}values (the concentration required for 50% inhibition of cell-growth) of each conjugate were calculated by SoftMax Pro 6.5.

Acknowledgements

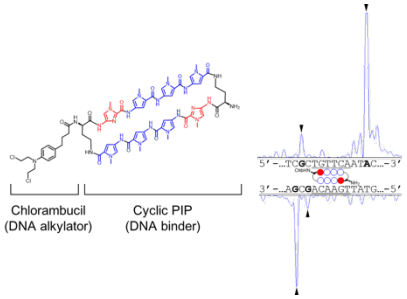
Capillary electrophoresis was performed at the Medical Research Support Center, Graduate School of Medicine, Kyoto University. This research was supported by AMED under Grant No. JP18am0301005 (Basic Science and Platform Technology Program for Innovative Biological Medicine) and JSPS KAKENHI (Grant No. JP16H06356 to H.S).

Keywords: DNA binder • DNA alkylator • anticancer drug

[1] I. Takahashi, K. Takahashi, M. Ichimura, M. Morimoto, K. Asano, I. Kawamoto, F. Tomita, H. Nakano, *J. Antibiot.* **1988**, *41*, 1915-1917.
[2] H. Sugiyama, M. Hosoda, I. Saito, A. Asai, H. Saito, *Tetrahedron Lett.* **1990**, *31*, 7197-7200.
[3] W. B. Mattes, J. A. Hartley, K. W. Kohn, *Nucleic Acids Res.* **1986**, *14*, 2971-2987.
[4] J. H. Yoon, C. S. Lee, *Arch. Pharm. Res.* **1997**, *20*, 550-554.
[5] L. J. Hanka, A. Dietz, S. A. Gerpheide, S. L. Kuentzel, D. G. Martin, *J. Antibiot.* **1978**, *31*, 1211-1217.
[6] L. H. Hurley, V. L. Reynolds, D. H. Swenson, G. L. Petzold, T. A. Scahill, *Science* **1984**, *226*, 843-844.
[7] S. R. Rajski, R. M. Williams, *Chem. Rev.* **1998**, *98*, 2723-2795.
[8] M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 1376-1380.
[9] J. W. Lown, *Drug Dev. Res.* **1995**, *34*, 145-183.
[10] P. B. Dervan, B. S. Edelson, *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.
[11] N. R. Wurtz, P. B. Dervan, *Chem. Biol.* **2000**, *7*, 153-161.
[12] Y. D. Wang, J. Dziegielewski, N. R. Wurtz, B. Dziegielewska, P. B. Dervan, T. A. Beerman, *Nucleic Acids Res.* **2003**, *31*, 1208-1215.
[13] M. Minoshima, T. Bando, K. Shinohara, G. Kashiwazaki, S. Nishijima, H. Sugiyama, *Bioorg. Med. Chem.* **2010**, *18*, 1236-1243.
[14] R. Maeda, S. Sato, S. Obata, T. Ohno, K. Hashiya, T. Bando, H. Sugiyama, *J. Am. Chem. Soc.* **2019**, *141*, 4257-4263.
[15] L. A. Dickinson, R. Burnett, C. Melander, B. S. Edelson, P. S. Arora, P. B. Dervan, J. M. Gottesfeld, *Chem. Biol.* **2004**, *11*, 1583-1594.
[16] K. Morita, K. Suzuki, S. Maeda, A. Matsuo, Y. Mitsuda, C. Tokushige, G. Kashiwazaki, J. Taniguchi, R. Maeda, M. Noura, M. Hirata, T. Kataoka, A. Yano, Y. Yamada, H. Kiyose, M. Tokumasu, H. Matsuo, S. Tanaka, Y. Okuno, M. Muto, K. Naka, K. Ito, T. Kitamura, Y. Kaneda, P. P. Liu, T. Bando, S. Adachi, H. Sugiyama, Y. Kamikubo, *J. Clin. Invest.* **2017**, *127*, 2815-2828.
[17] J. S. Choi, S. H. Joo, *Biomol. Ther.* **2020**, *28*. 18-24.
[18] D. M. Herman, J. M. Turner, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129.
[19] D. M. Chenoweth, D. A. Harki, J. W. Phillips, C. Dose, P. B. Dervan, *J. Am. Chem. Soc.* **2009**, *131*, 7182-7188.
[20] B. C. Li, D. C. Montgomery, J. W. Puckett, P. B. Dervan, *J. Org. Chem.* **2013**, *78*, 124-133.
[21] Y. Hirose, S. Asamitsu, T. Bando, H. Sugiyama, *J. Am. Chem. Soc.* **2019**, *141*, 13165-13170.
[22] T. Zou, S. Kizaki, H. Sugiyama, *ChemBioChem* **2018**, *19*, 664-668.
[23] D. M. Herman, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 1382-1391.
[24] A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 560-564.
[25] F. Sanger, S. Nicklen, R. Coulson, *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 5463-5467.
[26] H. Sugiyama, T. Fujiwara, A. Ura, T. Tashiro, K. Yamamoto, S. Kawanishi, I. Saito, *Chem. Res. Toxicol.* **1994**, *7*, 673-683.
[27] T. Bando, A. Narita, I. Saito, H. Sugiyama, *Chem. Eur. J.* **2002**, *8*, 4781-4790.
[28] S. Asamitsu, Y. Kawamoto, F. Hashiya, K. Hashiya, M. Yamamoto, S. Kizaki, T. Bando, H. Sugiyama, *Bioorg. Med. Chem.* **2014**, *22*, 4646-4657.
[29] G. Gaastra, in *Nucleic Acids. Methods in Molecular Biology*, *Vol. 2* (Eds.: J. M. Walker), Humana Press, Totowa, **1984**, pp. 333-341.

Entry for the Table of Contents

Insert graphic for Table of Contents here.



The DNA alkylation property of a cyclic pyrrole-imidazole polyamide, which was conjugated with a DNA-alkylator chlorambucil, was reported. The conjugate showed good alkylation ability *in vitro* and anti-cancer effects against cancer cells.

CONTENTS

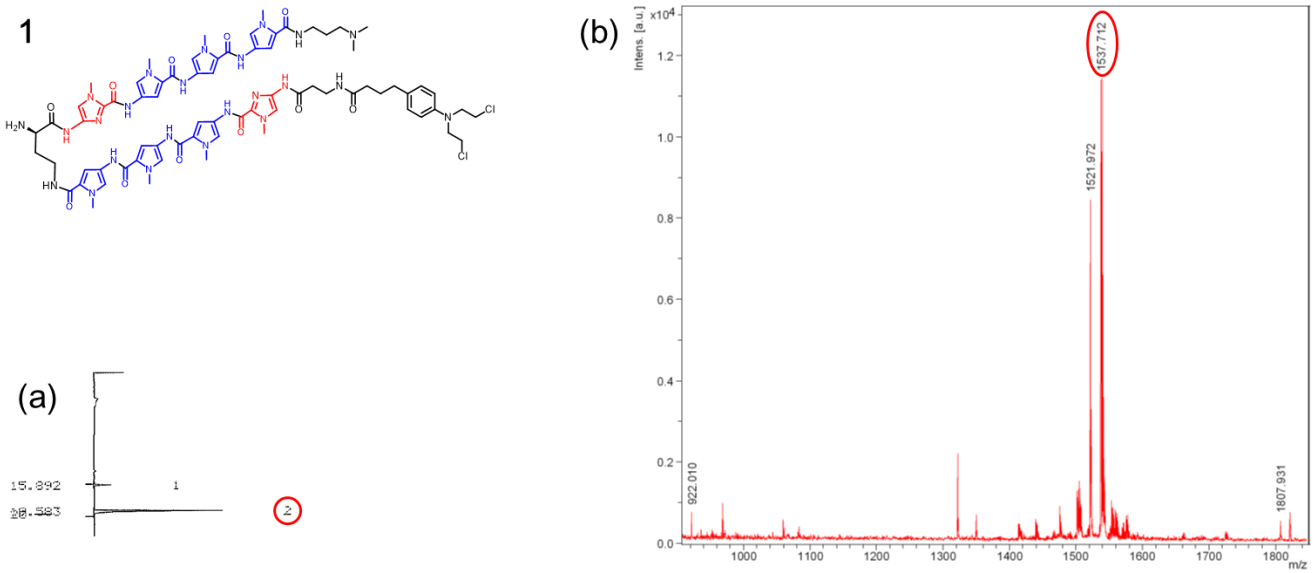


Figure S1. (a) HPLC spectrum of compound 1. Conditions: 0.1% TFA containing 0−100% acetonitrile over a linear gradient for 40 min at a flow rate of 1.0 mL/min detected at 254 nm. The retention time was 19.583 min. (b) MALDI-TOF MS spectrum of compound 1. *m/z* calcd for $C_{72}H_{91}Cl_2N_{24}O_{11}^+$[M+H]$^+$1537.66, found; 1537.71. The three peaks, 922.01, 1521.97, and 1807.93, are calibration peaks.

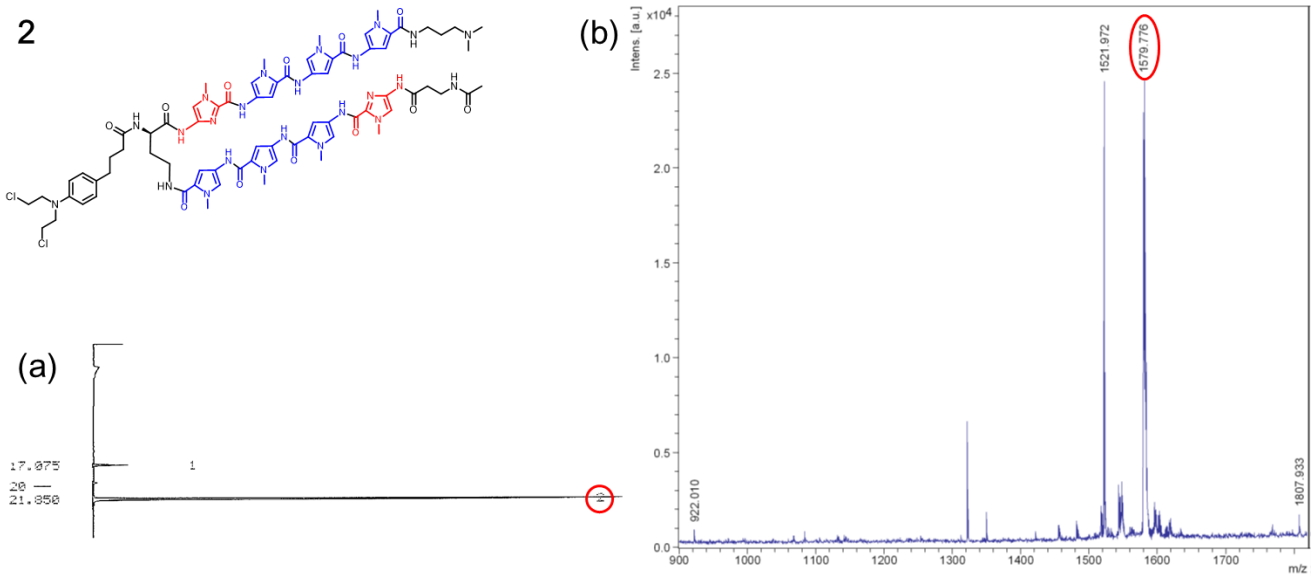


Figure S2. (a) HPLC spectrum of compound 2. Conditions: 0.1% TFA containing 0−100% acetonitrile over a linear gradient for 40 min at a flow rate of 1.0 mL/min detected at 254 nm. The retention time was 21.850 min. (b) MALDI-TOF MS spectrum of compound 2. *m/z* calcd for $C_{74}H_{93}Cl_2N_{24}O_{12}^+$[M+H]$^+$1579.67, found; 1579.78. The three peaks, 922.01, 1521.97, and 1807.93, are calibration peaks.

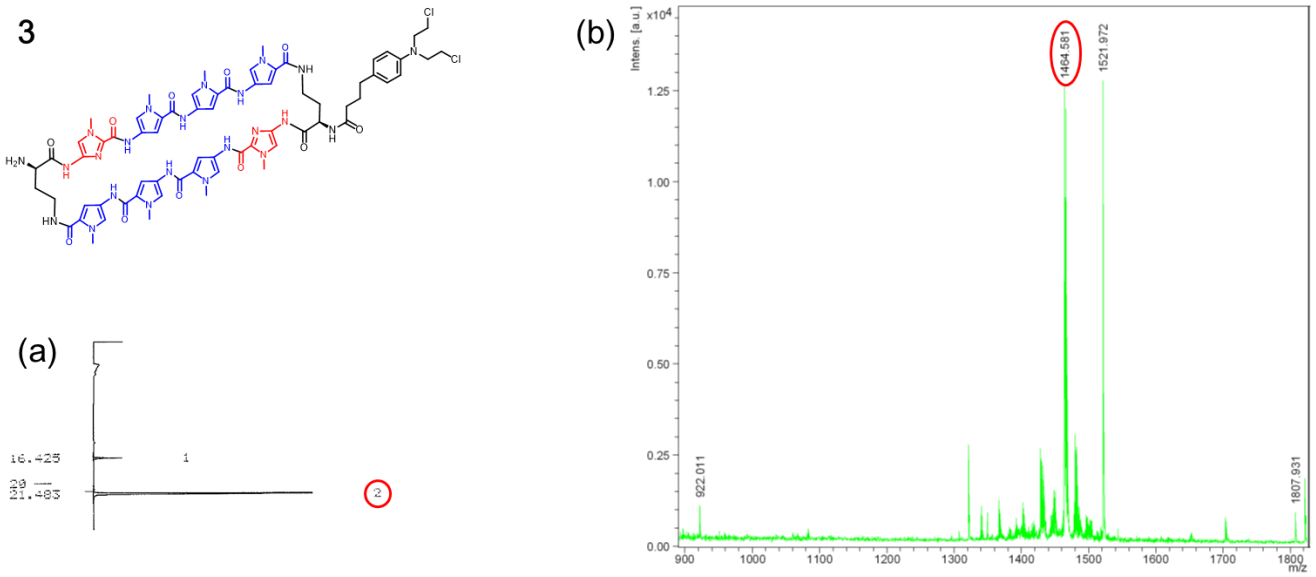


Figure S3. (a) HPLC spectrum of compound 3. Conditions: 0.1% TFA containing 0−100% acetonitrile over a linear gradient for 40 min at a flow rate of 1.0 mL/min detected at 254 nm. The retention time was 21.483 min. (b) MALDI-TOF MS spectrum of compound 3. *m/z* calcd for $C_{68}H_{80}Cl_2N_{23}O_{11}^+$[M+H]$^+$1464.57, found; 1464.58. The three peaks, 922.01, 1521.97, and 1807.93, are calibration peaks.

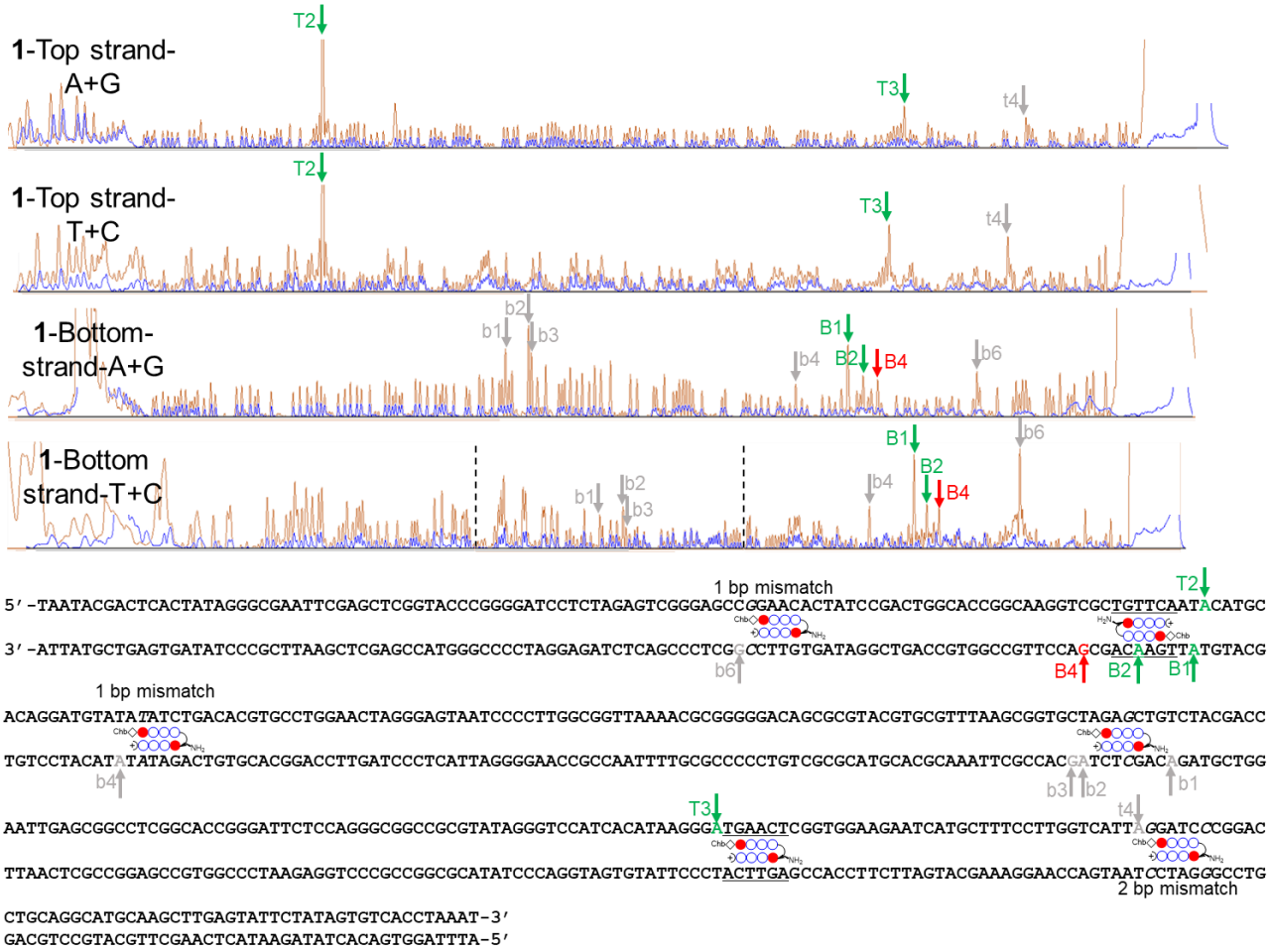


Figure S4. The alkylation sites of conjugate 1 identified by mixing alkylated fragments and sequencing fragments prepared by the Maxam-Gilbert sequencing method (blue waves: sequencing fragments prepared by the Maxam-Gilbert sequencing method, orange waves: mixtures of sequencing fragments and alkylated fragments). “t” and “b” mean “top strand” and “bottom strand”, respectively. Data of “1-Bottom-T+C” were pieced together to correct the misalignment; joints were shown as dashed lines.

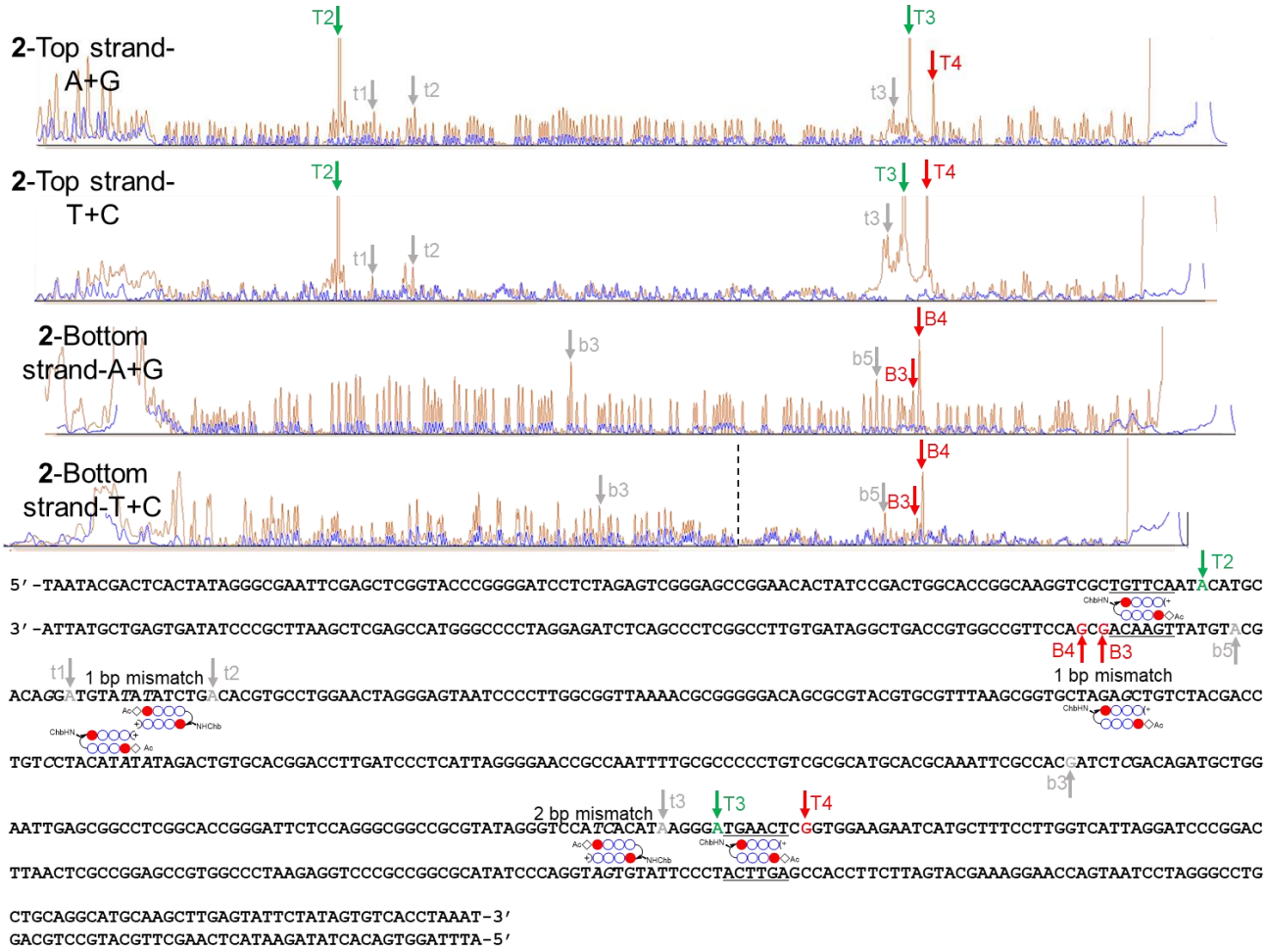


Figure S5. The alkylation sites of conjugate 2 identified by mixing alkylated fragments and sequencing fragments prepared by the Maxam-Gilbert sequencing method (blue waves: sequencing fragments prepared by the Maxam-Gilbert sequencing method, orange waves: mixtures of sequencing fragments and alkylated fragments). “t” and “b” mean “top strand” and “bottom strand”, respectively. Data of “2-Bottom-T+C” were pieced together to correct the misalignment; a joint was shown as dashed lines.

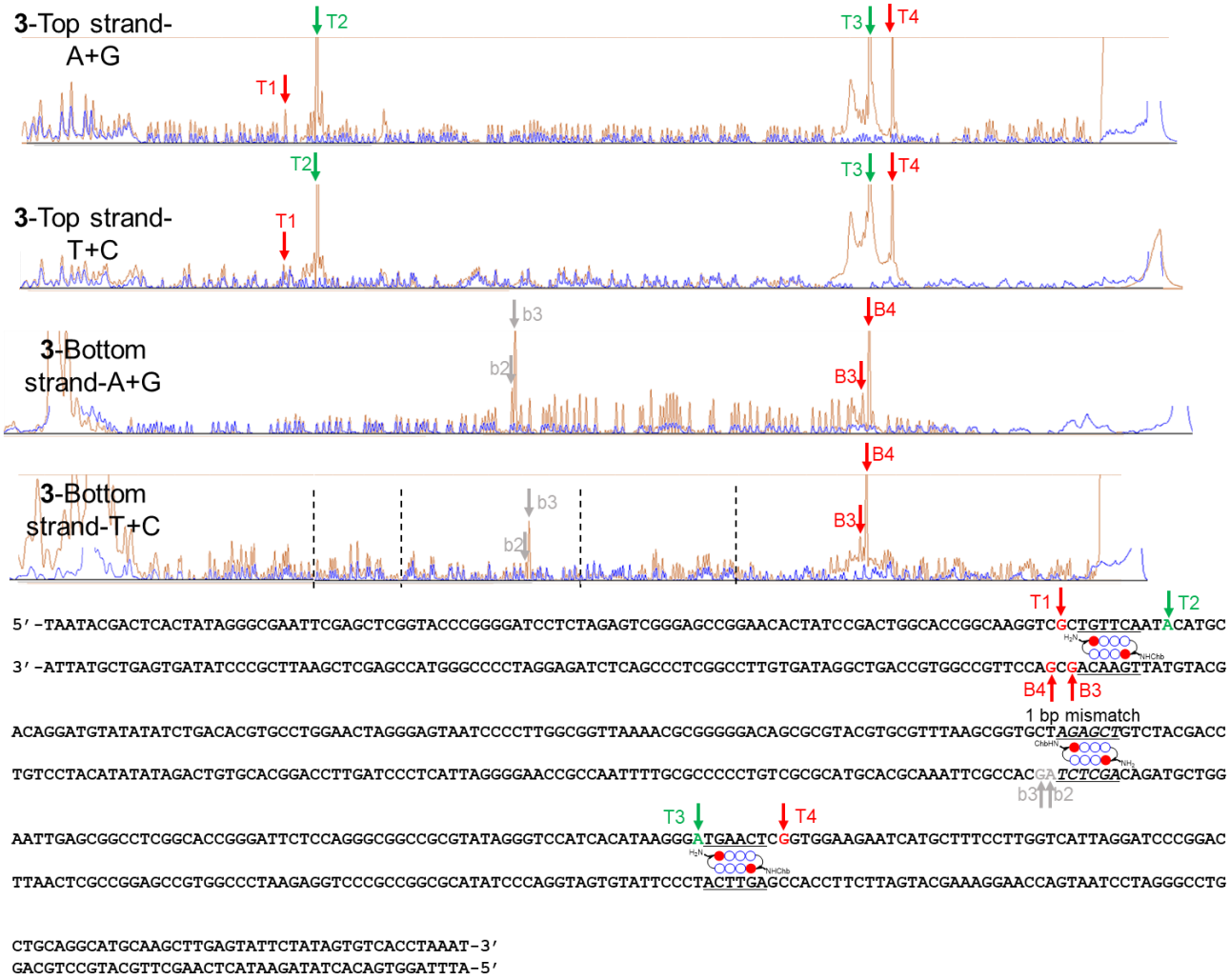


Figure S6. The alkylation sites of conjugate 3 identified by mixing alkylated fragments and sequencing fragments prepared by the Maxam-Gilbert sequencing method (blue waves: sequencing fragments prepared by the Maxam-Gilbert sequencing method, orange waves: mixtures of sequencing fragments and alkylated fragments). “t” and “b” mean “top strand” and “bottom strand”, respectively. Data of “3-Bottom strand-T+C” were pieced together to correct the misalignment; joints were shown as dashed lines.

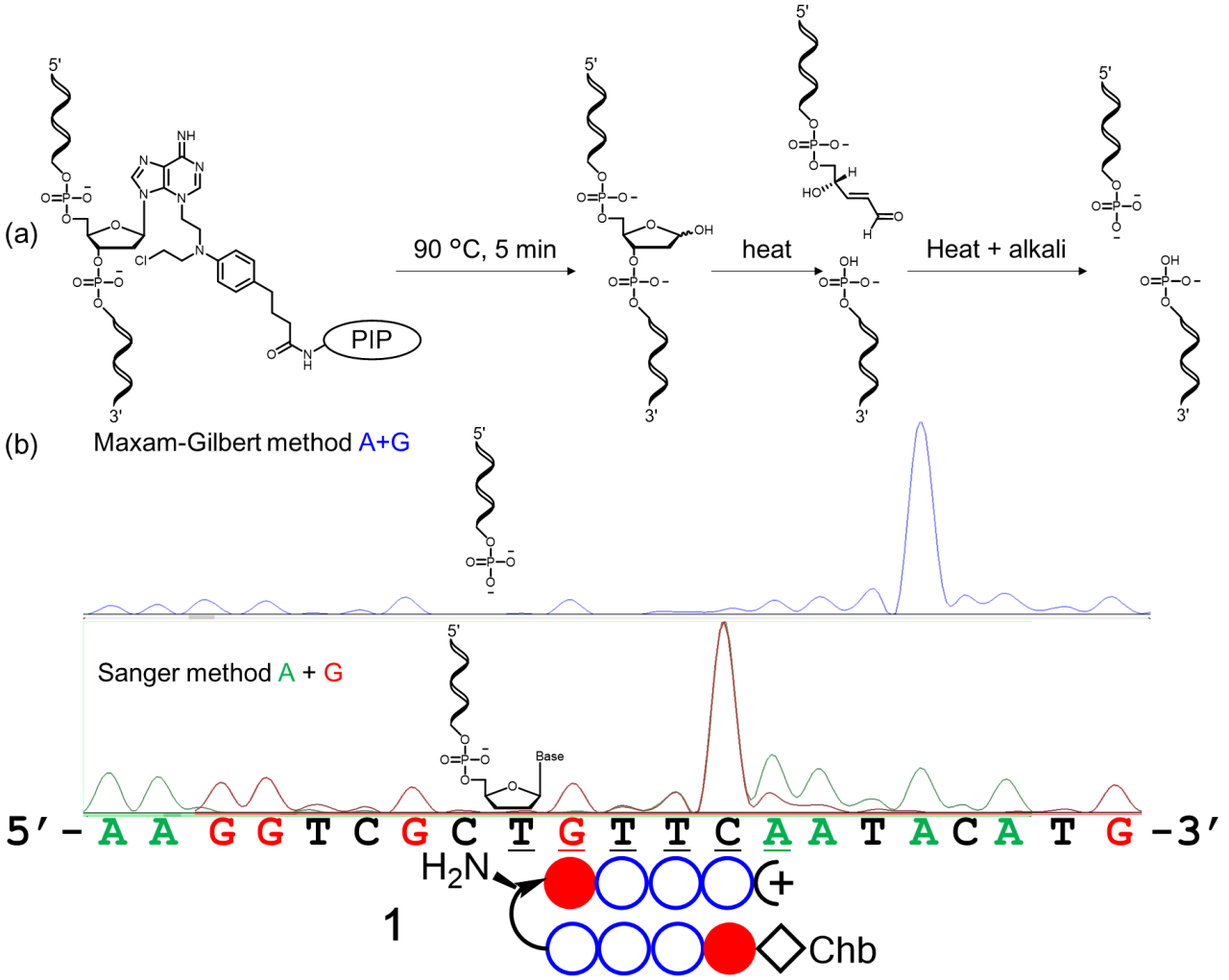


Figure S7. (a) The mechanism of the alkylation by Chb and cleavage by heat and alkali treatment. (b) Comparison of the alkylation site at one of the binding sites identified by mixing alkylated fragments and sequencing fragments prepared by the Maxam-Gilbert sequencing method (blue) or Sanger sequencing method (green and red).

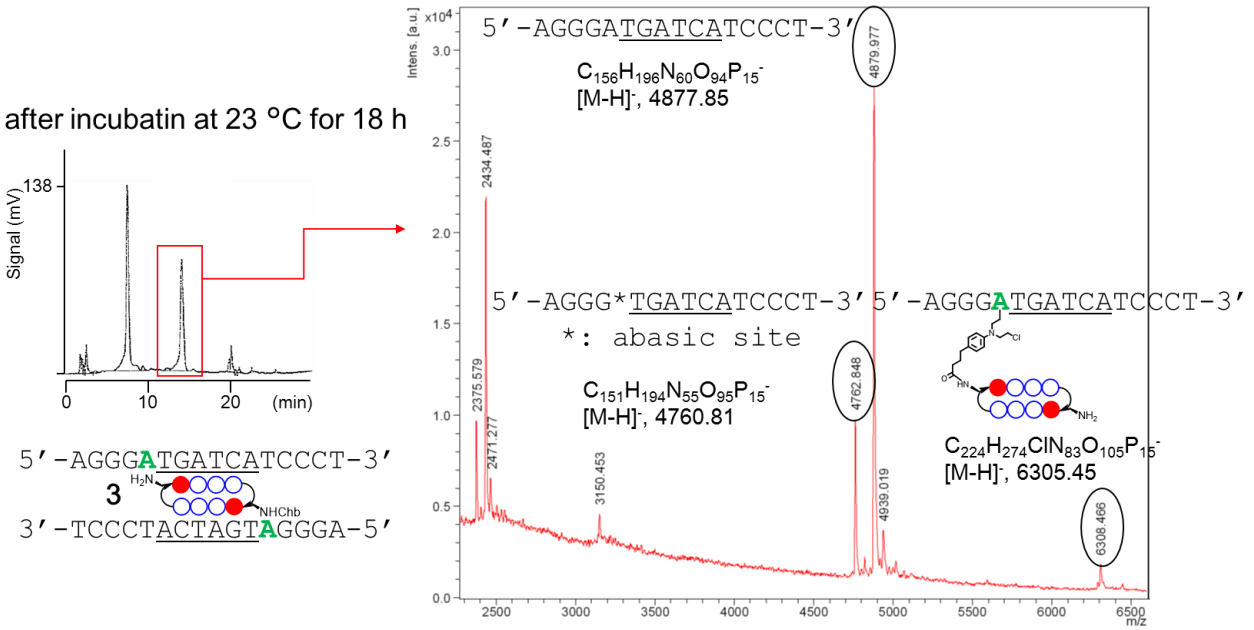


Figure S8. The HPLC profile after alkylation of 16 bp DNA fragments by conjugate 3 (same profile as figure 4a) and MALDI-TOFMS spectrum of the peak at 14.1 min. HPLC analysis was performed with a mobile phase containing 50 mM TEAA (Triethylamine acetate) in gradient combination with acetonitrile (9-12%: 0-15 min, 12-100%: 15-30 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm.

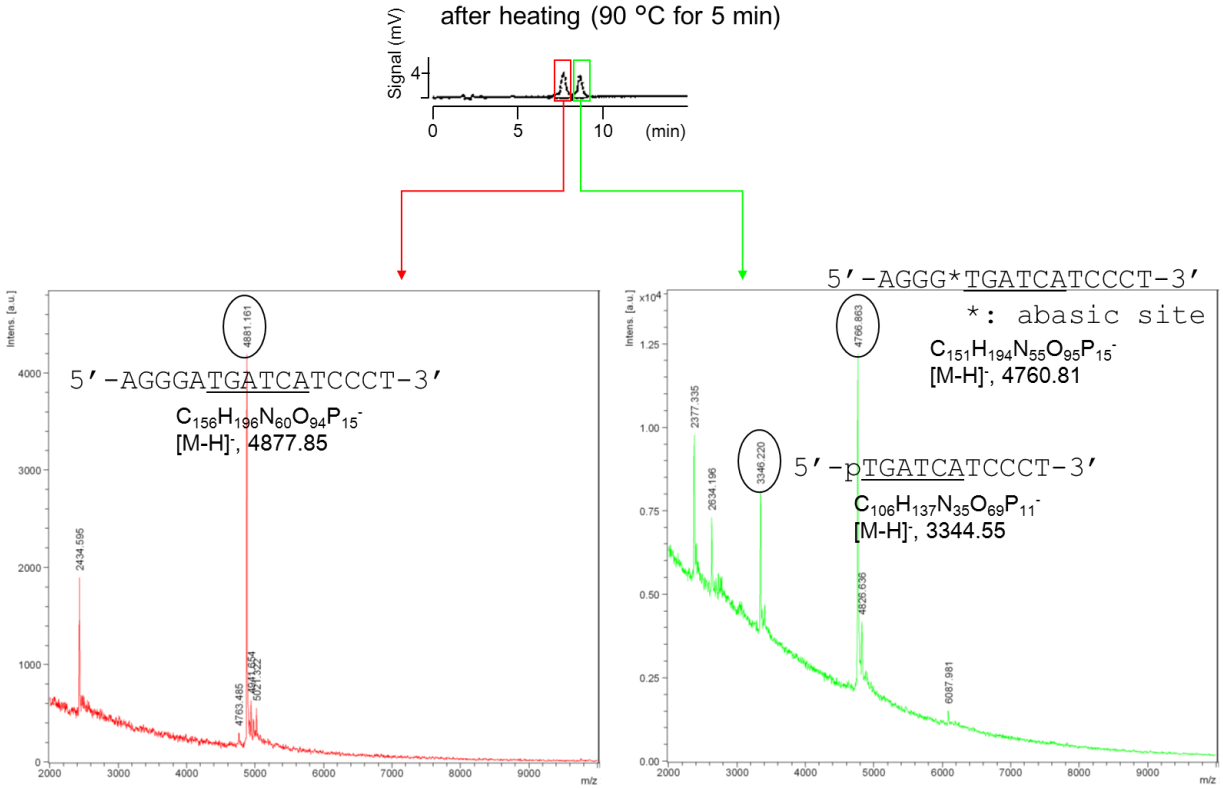


Figure S9. The HPLC profile after heating (90 °C for 5 min) the peak at about 14 min in figure 4a and S8 (same profile as the upper profile in figure 4b) and MALDI-TOFMS spectrum of the peak at about 7.7 min and 8.7 min. HPLC analysis was performed with a mobile phase containing 50 mM TEAA (Triethylamine acetate) in gradient combination with acetonitrile (9-12%: 0-15 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm.

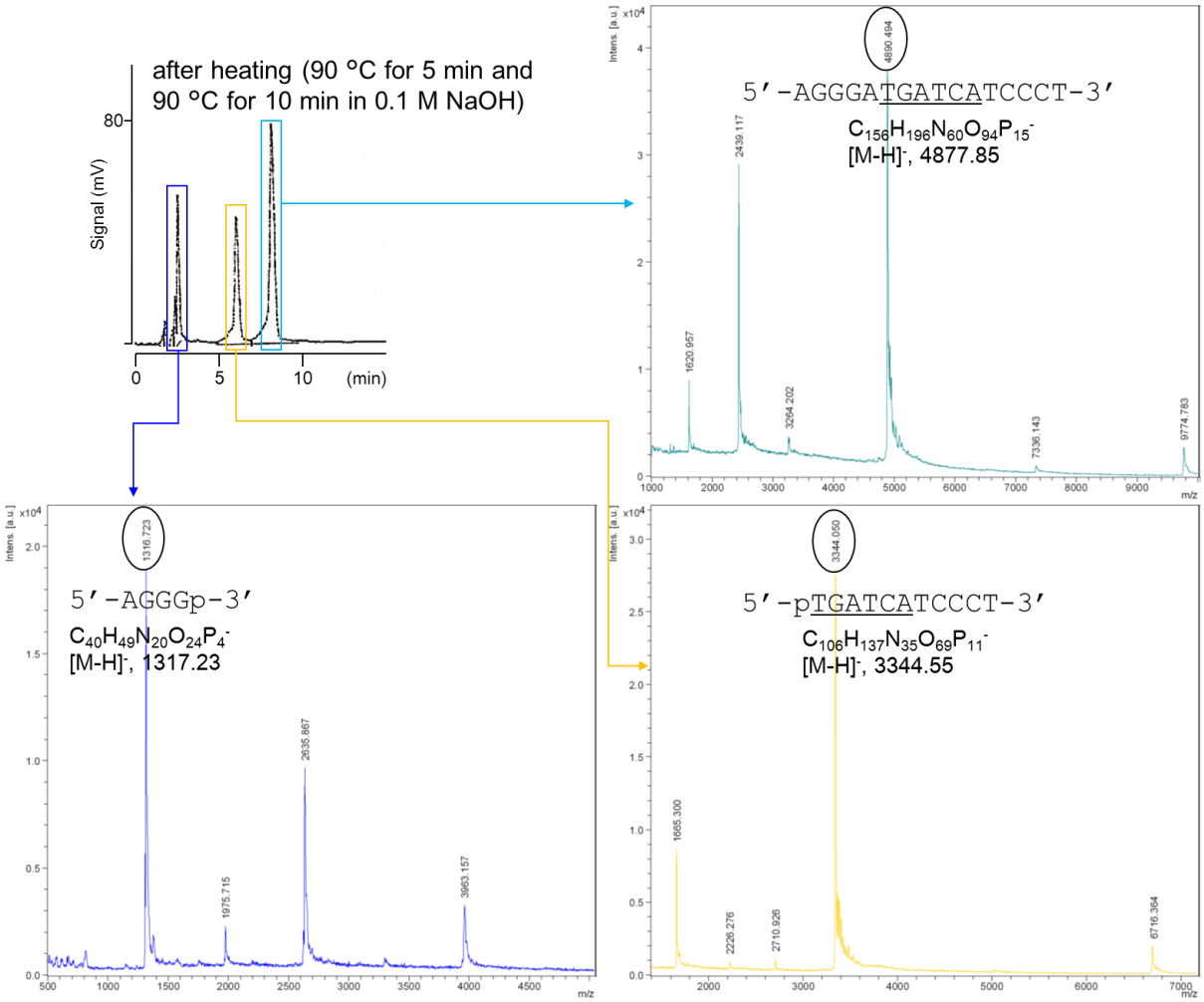


Figure S10. The HPLC profile after heating (90 °C for 5 min), lyophilizing, and additionally heating (90 °C for 10 min in 0.1 M NaOH) the peak at about 14 min in figure 4a and S8 and MALDI-TOFMS spectrum of the peak at 2.6 min, 6.1 min, and 8.2 min. HPLC analysis was performed with a mobile phase containing 50 mM TEAA (Triethylamine acetate) in gradient combination with acetonitrile (9-12%: 0-15 min) at a flow rate of 1.0 mL/min and the eluent was detected at 254 nm.

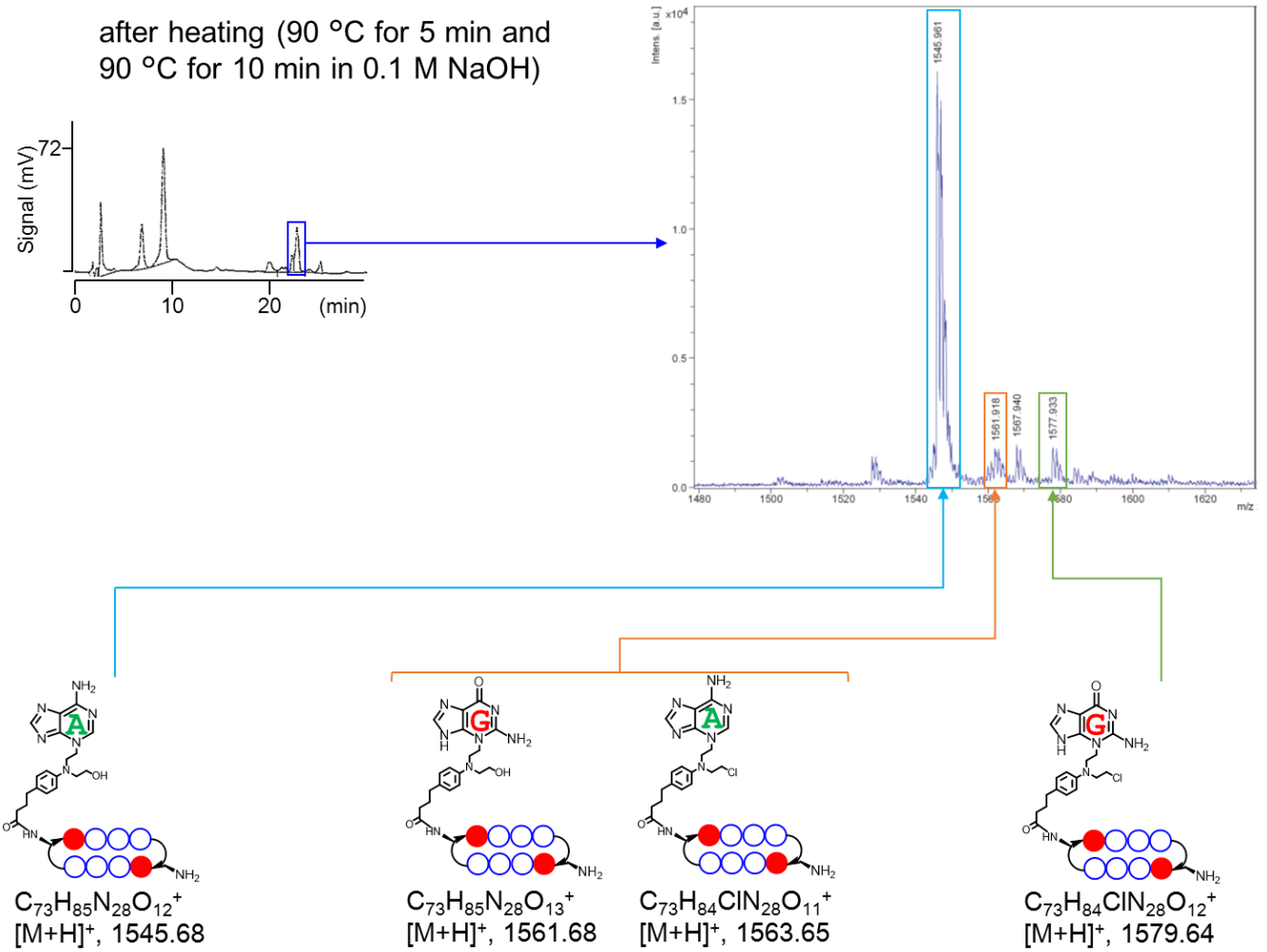


Figure S11. The HPLC profile after heating (90 °C for 5 min), lyophilizing, and additionally heating (90 °C for 10 min in 0.1 M NaOH) the peak at about 14 min after alkylation and MALDI-TOFMS spectrum of the peak at about 22.8 min.