

Cite this: *RSC Adv.*, 2017, 7, 25378

Benzene-glycol nucleic acid (BGNA)–DNA chimeras: synthesis, binding properties, and ability to elicit human RNase H activity†

Nazuki Niwa,^a Saki Shimizu,^d Yusuke Maeda,^a Hidekazu Hiroak^d
and Yoshihito Ueno^{*abc}

This paper describes the synthesis and properties of benzene-glycol nucleic acid (BGNA)–DNA chimeras containing four nucleoside analogs – thymidine, cytidine, adenosine, and guanosine – with a base-benzene-glycol structure. We found that the BGNA–DNA chimeras are able to form thermally and thermodynamically stable duplexes with complementary RNAs, and have base-discriminating abilities. The BGNA–DNA chimeras were 20-fold more stable in a buffer containing 30% bovine serum than unmodified DNA. Furthermore, BGNA–DNA chimera/RNA duplexes were found to be good substrates for human RNase H. Thus, BGNA–DNA chimeras are good candidates for the development of therapeutic antisense molecules.

Received 6th April 2017
Accepted 28th April 2017

DOI: 10.1039/c7ra03896h

rsc.li/rsc-advances

Introduction

Antisense oligonucleotides (AONs) are single-stranded DNAs consisting of 17–30 bases that can inhibit translation by binding to complementary target mRNAs.¹ AONs have recently attracted attention as chemotherapeutic agents, since they can be rationally designed and synthesized given the sequences of disease-causing genes.² Fomivirsen, a drug for cytomegalovirus (CMV) retinitis in patients with acquired immune deficiency syndrome (AIDS), is the first AON drug approved by the U.S. Food and Drug Administration (FDA).³ Mipomersen is an AON drug for patients with homozygous familial hypercholesterolemia. Eteplirsen, a drug used to treat patients with Duchenne muscular dystrophy, has also recently been approved by the FDA.⁴

As phosphodiester linkages in DNA are easily hydrolyzed by the nucleases present inside and outside cells, chemical modifications are needed to impart nuclease-resistant properties to DNAs for them to function as antisense drugs. For example, fomivirsen is composed of phosphorothioate linkages instead of natural phosphate linkages.³ Mipomersen is partially composed of nucleoside analogs modified with methoxyethyl

(MOE) groups at the 2'-hydroxyl positions. Eteplirsen has a backbone composed of methylenemorpholine rings and phosphorodiamidate linkages.

Recently, we reported the synthesis of two nucleoside analogs, a cytidine analog (1) and an adenosine analog (2), comprising a base-benzene-glycol structure (Fig. 1).⁵ We found that DNAs partially incorporating the benzene-glycol nucleic acids (BGNAs), termed BGNA–DNA chimeras, form thermally stable duplexes with complementary DNA and RNA, and these duplexes are more resistant to hydrolysis by a snake venom phosphodiesterase (a 3'-exonuclease) than unmodified DNA. As an extension of our study on the synthesis of BGNAs, we report in this paper the synthesis of a thymidine analog (3) and a guanosine analog (4), and properties of BGNA–DNA chimeras containing these four analogs as antisense molecules.

Results and discussion

Synthesis

We previously reported the synthesis of the cytidine and adenosine analogs using the Chan–Lam–Evans reaction between an arylboronic acid and nucleobases.^{5–7} Synthesis of the thymidine and guanosine analogs was achieved using the same reaction. To prevent the side reaction between the N³ position of thymine and the arylboronic acid, we used an N³-protected thymine, N³-benzoylthymine, as a substrate for the synthesis of the thymidine analog. To solve the problem of the insolubility of guanine, a protected purine base, 2-bis(*tert*-butoxycarbonyl)amino-6-chloropurine, was used as a substrate for the synthesis of the guanosine analog.^{5b} The synthetic routes to the thymidine and guanosine analogs, and the corresponding phosphoramidite units (8 and 14, respectively), are shown in Schemes 1 and 2.

^aCourse of Applied Life Science, Faculty of Applied Biological Sciences, Gifu University, Japan. E-mail: uenoy@gifu-u.ac.jp; Fax: +81-58-293-2919; Tel: +81-58-293-2919

^bUnited Graduate School of Agricultural Science, Gifu University, Japan

^cCenter for Highly Advanced Integration of Nano and Life Sciences (G-CHAIN), Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan

^dDivision of Structural Biology, Graduate School of Pharmaceutical Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

† Electric supplementary information (ESI) available. See DOI: 10.1039/c7ra03896h



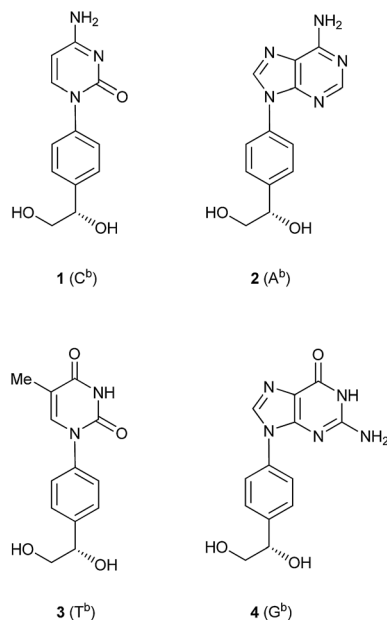


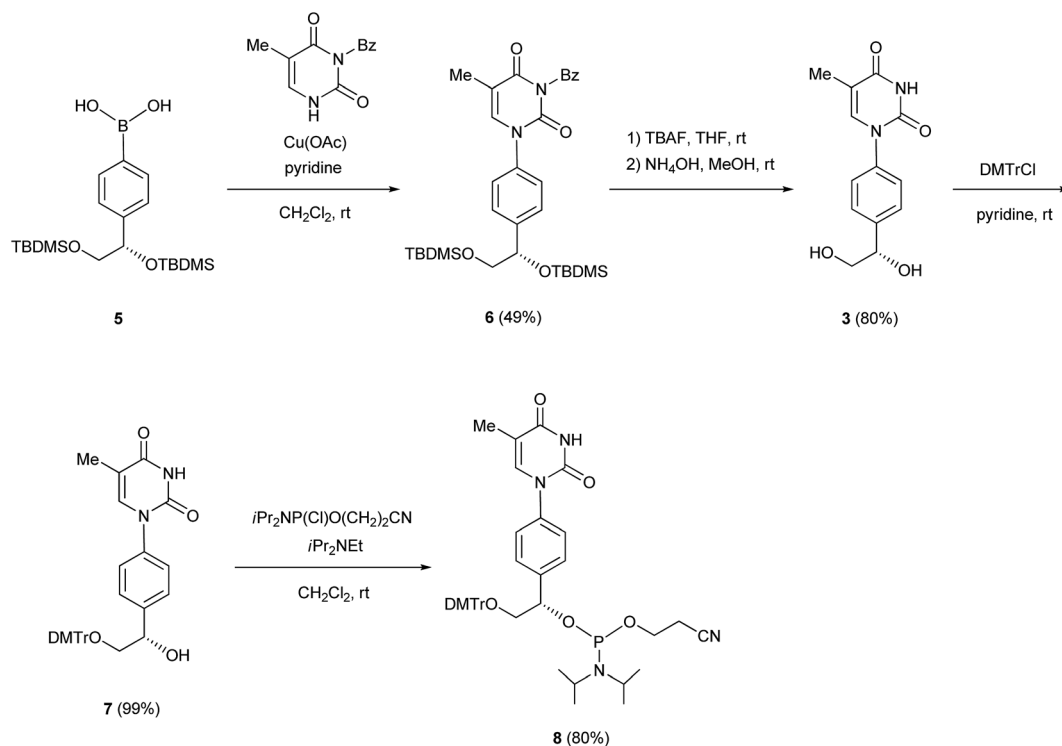
Fig. 1 Structures of the nucleoside analogs.

N^3 -Benzoylthymine was coupled with the arylboronic acid derivative **5** in the presence of $\text{Cu}(\text{OAc})_2$ and pyridine as a base in CH_2Cl_2 to give an N^1 -aryl thymine derivative (**6**) with a yield of 49%. Deprotection of the silyl and benzoyl groups of **6** with tetra-*n*-butylammonium fluoride (TBAF) and NH_4OH produced the thymine analog **3** with a yield of 80%. The primary hydroxy group of **3** was protected with a 4,4'-dimethoxytrityl (DMTr)

group to give a mono-DMTr derivative (**7**) with a yield of 99%. Compound **7** was phosphitylated using the standard procedure, leading to the formation of the corresponding phosphoramidite (**8**) with a yield of 80%.

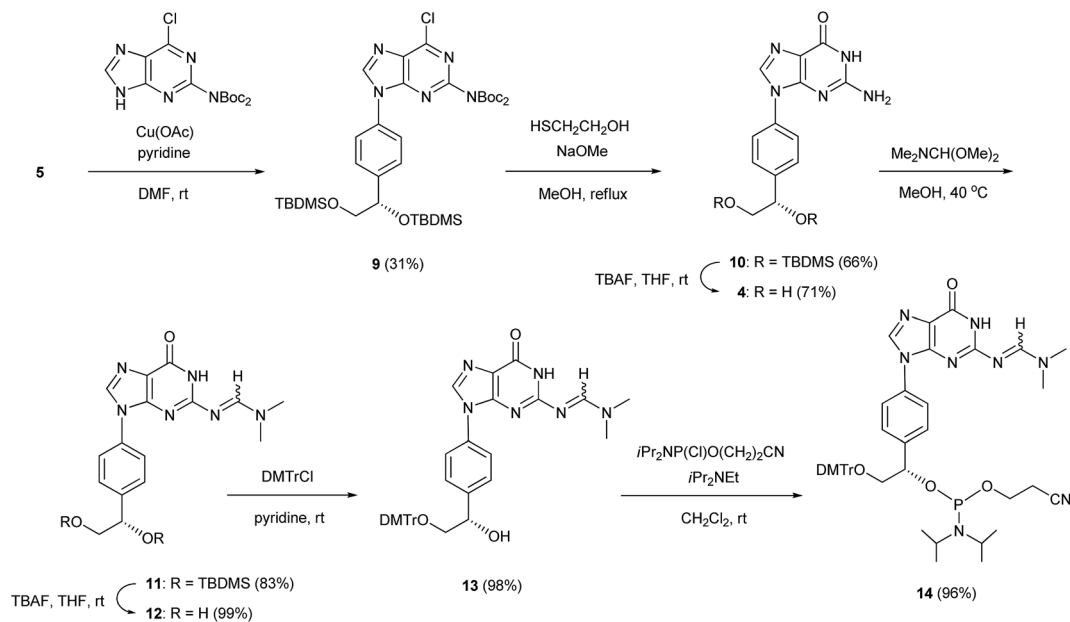
In a similar manner, 2-bis(*tert*-butoxycarbonyl)amino-6-chloropurine was coupled with the arylboronic acid derivative **5** in the presence of $\text{Cu}(\text{OAc})_2$ and pyridine in DMF to form an N^9 -aryl purine derivative (**9**) with a yield of 31%. The treatment of **9** with 2-mercaptoethanol and NaOMe in MeOH gave an N^9 -arylguanine derivative (**10**) with a yield of 66%. The silyl groups of **10** were removed by treatment with TBAF to give the guanosine analog **4** with a yield of 71%. The *exo*-amino function of **10** was protected with a dimethylaminomethylene group to produce a fully protected guanosine derivative (**11**) with a yield of 83%. After deprotection of the silyl groups with TBAF, the primary hydroxy group of **12** was protected with a DMTr group to obtain a mono-DMTr derivative (**13**) with a yield of 98%. Compound **13** was phosphitylated by the standard procedure to lead to the formation of the corresponding phosphoramidite (**14**) with a yield of 96%.

All oligonucleotides (ONs) containing the analogs **1–4** were synthesized using the phosphoramidites **8** and **14**, and the phosphoramidites of the cytidine and adenosine analogs, which were synthesized using the previously published procedure.^{5c} The ON sequences used are given in Table 1. The ONs were analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/MS), and the observed molecular weights obtained from these analyses were consistent with the ON structures.



Scheme 1



Table 1 Sequences of duplexes and ONs, and T_m s and thermodynamic parameters of the duplexes

Duplex	ON	Sequences ^a	T_m (°C)	ΔT_m (°C)	ΔH° (kcal mol ⁻¹)	ΔS° (cal mol ⁻¹ K ⁻¹)	ΔG_{310}° (kcal mol ⁻¹)
Duplex 1	DNA 1	5'-d(AAAGTTCACTACTCAA)-3'	52.3		-98.4	-275.2	-13.1
	RNA 1	3'-r(UUUCAAGUGAUGAGUUU)-5'					
Duplex 2	ON 1	5'-d(AA ^b A ^b GTTCACTACTCA ^b A ^b A ^b)-3'	55.6	+3.3	-94.7	-259.8	-14.1
	RNA 1	3'-r(UUUCAAGUGAUGAGUUU)-5'					
Duplex 3	DNA 2	5'-d(ACAGTTCACTACTCACA)-3'	58.0		-144.1	-406.6	-18.0
	RNA 2	3'-r(UGUCAAGUGAUGAGUGU)-5'					
Duplex 4	ON 2	5'-d(AC ^b A ^b GTTCACTACTCA ^b C ^b A ^b)-3'	57.9	-0.1	-113.6	-315.0	-16.0
	RNA 2	3'-r(UGUCAAGUGAUGAGUGU)-5'					
Duplex 5	DNA 3	5'-d(AGAGTTCACTACTCAGA)-3'	57.5		-132.5	-373.5	-16.8
	RNA 3	3'-r(UCUCAAGUGAUGAGUCU)-5'					
Duplex 6	ON 3	5'-d(AG ^b A ^b GTTCACTACTCA ^b G ^b A ^b)-3'	56.4	-1.1	-112.3	-313.0	-15.3
	RNA 3	3'-r(UCUCAAGUGAUGAGUCU)-5'					
Duplex 7	DNA 4	5'-d(ATAGTTCACTACTCATA)-3'	54.8		-122.0	-343.9	-15.4
	RNA 4	3'-r(UAUCAAGUGAUGAGUAU)-5'					
Duplex 8	ON 4	5'-d(AT ^b A ^b GTTCACTACTCA ^b T ^b A ^b)-3'	56.0	+1.2	-117.6	-329.6	-15.5
	RNA 4	3'-r(UAUCAAGUGAUGAGUAU)-5'					
Duplex 9	DNA 5	5'-d(AAATCACTACAAA)-3'	36.8		-89.6	-260.9	-8.7
	RNA 5	3'-r(UUUAGUGAUGUUU)-5'					
Duplex 10	ON 5	5'-d(AA ^b A ^b TCACTACA ^b A ^b A ^b)-3'	35.0	-1.8	-66.0	-186.1	-8.4
	RNA 5	3'-r(UUUAGUGAUGUUU)-5'					
Duplex 11	DNA 6	5'-d(ACATCACTACACA)-3'	47.7		-108.6	-310.5	-12.4
	RNA 6	3'-r(UGUAGUGAUGUGU)-5'					
Duplex 12	ON 6	5'-d(AC ^b A ^b TCACTACA ^b C ^b A ^b)-3'	41.8	-5.9	-99.9	-288.8	-10.4
	RNA 6	3'-r(UGUAGUGAUGUGU)-5'					
	DNA 7	5'-F-d(AAACACCGACGCGAAA)-3'					
	ON 7	5'-F-d(AA ^b A ^b CACCGACGCGCA ^b A ^b A ^b)-3'					
	RNA 7	5'-F-r(UUUGAGUGAUGAACUUU)-3'					
	RNA 8	5'-F-r(UUUGAGUGAUUU)-3'					

^a F indicates a fluorescein.

Thermal and thermodynamic stabilities of duplexes

The thermal stabilities of DNA/RNA and ON/RNA duplexes containing the analogs were compared by measuring their melting temperatures (T_m s) in a 10 mM sodium phosphate buffer (pH 7.0) containing 100 mM NaCl (Table 1 and Fig. S1†). Furthermore, to examine the stability of the duplexes in detail, we calculated the thermodynamic parameters of duplex formation based on the slope of the plot of $1/T_m$ vs. $\ln(C_T/4)$, where C_T is the total concentration of single-stranded ON and RNA (Table 1 and Fig. S3†). Duplexes 1–8 are composed of 17 base pairs, while duplexes 9–12 comprise 13 base pairs. All ON/RNA duplexes contain two and three analogs at the 5'- and 3'-regions of the ONs, respectively. Duplex 2 has five A^b:rU base pairs among its 17 base pairs, whereas duplexes 4, 6, and 8 contain two C^b:rG, G^b:rC, and T^b:rA base pairs in addition to the A^b:rU base pairs in the sequences, respectively. Duplex 10 has five A^b:rU base pairs among its 13 base pairs, while duplex 12 contains two C^b:rG base pairs in addition to the A^b:rU base pairs in the sequence.

The thermal and thermodynamic stabilities of the duplexes containing the analogs were found to be dependent on their sequences. Duplexes 2 and 8, containing the A^b:rU and T^b:rA base pairs, were thermally and thermodynamically more stable than the corresponding DNA/RNA duplexes 1 and 7, respectively, while duplexes 4 and 6, containing the C^b:rG and G^b:rC base pairs, were thermally and thermodynamically less stable than the corresponding DNA/RNA duplexes 3 and 5, respectively. However, the differences in the thermal and

thermodynamic stabilities between the duplexes containing the analogs and the unmodified DNA/RNA duplexes were not large ($\Delta T_m = -1.1 \sim +3.3$ °C, $\Delta\Delta G_{310}^\circ = +2.0 \sim -1.0$ kcal mol⁻¹). On comparison of the thermodynamic parameters, it was found that duplex formation between ONs containing the analogs and the complementary RNAs was less favorable in terms of enthalpy but more favorable in terms of entropy than that of the corresponding unmodified DNA/RNA duplexes.

The ΔT_m and $\Delta\Delta G_{310}^\circ$ values between duplexes 1 and 2 were +3.3 °C and -1.0 kcal mol⁻¹, respectively, while those between duplexes 9 and 10 were -1.8 °C and +0.3 kcal mol⁻¹, respectively. Similarly, the ΔT_m and $\Delta\Delta G_{310}^\circ$ values between duplexes 3 and 4 were -0.1 °C and +2.0 kcal mol⁻¹, respectively, while those between duplexes 11 and 12 were -5.9 °C and +2.0 kcal mol⁻¹, respectively. Thus, incorporation of the analogs into duplexes comprising 13 base pairs destabilized the duplexes more than their incorporation into duplexes composed of 17 base pairs, both thermally and thermodynamically.

The base-discriminating abilities of the analogs 1–4 in DNA/RNA duplexes were then assessed (Fig. 2 and Table S1†). The ΔT_m (T_m of duplex containing mismatched base pair - T_m of complementary duplex) values for the unmodified duplexes were -0.6 ~ -12.3 °C, whereas those for the duplexes containing the analogs were -0.3 ~ -5.2 °C. Thus, the analogs 1–4 have base-discriminating abilities in DNA/RNA duplexes, though weaker than those of the corresponding unmodified nucleosides.

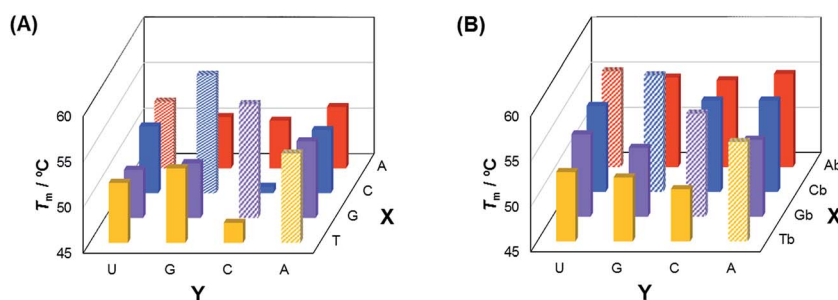


Fig. 2 Comparison of the T_m values of the duplexes composed of matched and mismatched sequences. Sequences are shown in Table S1.†

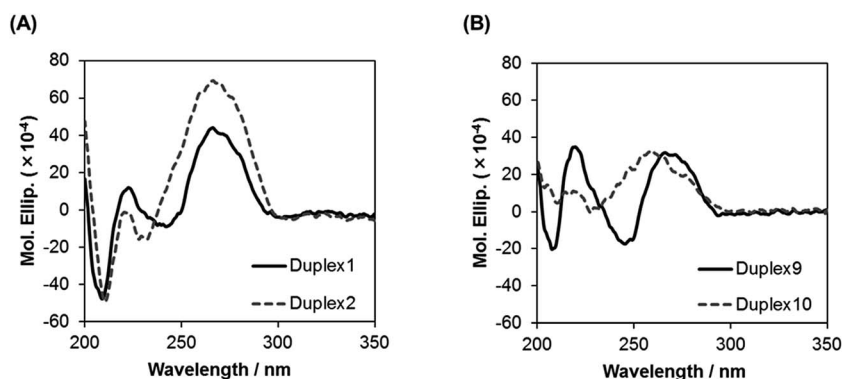


Fig. 3 CD spectra of duplexes in a buffer of 10 mM sodium phosphate (pH 7.0) and 0.1 M NaCl at 15 °C. Concentration of duplexes: 4 μ M.



Circular dichroism

To study the global conformations of the duplexes, we analyzed the circular dichroism (CD) spectra of the duplexes composed of ON 1 (17-mer) and of ON 5 (13-mer), containing five analogs (Fig. 3). The ON 1/RNA 1 duplex showed a positive CD band around 270 nm and a negative CD band at 210 nm, which was similar to the spectrum of the unmodified DNA 1/RNA 1 duplex, whereas the shape of the CD spectrum of the ON 5/RNA 5 duplex was different from that of the unmodified DNA 5/RNA 5 duplex. These results suggest that the global conformation of the ON 1/RNA 1 duplex composed of 17 base pairs is similar to that of the unmodified DNA 1/RNA 1 duplex, while the global conformation of the ON 5/RNA 5 duplex comprising 13 base pairs is different from that of the unmodified DNA 5/RNA 5 duplex.

Nuclease resistance

Previously, we reported that BGN–DNA chimeras are more resistant to hydrolysis by a snake venom phosphodiesterase (a 3'-exonuclease) than unmodified DNA.^{5c} To be able to function as drugs, AONs must be stable in blood, which contains various enzymes, including nucleases. Therefore, we investigated the stability of BGNADNA chimeras in bovine serum (BS). DNA 7 or ON 7, labeled with fluorescein at the 5'-end, was incubated in a buffer containing 30% BS and analyzed by polyacrylamide gel electrophoresis under denaturing conditions (Fig. 4). The half-life of DNA 7 was 35 min, while that of ON 7 was 11.7 h. Thus, the BGN–DNA chimera (ON 7) was approximately 20-fold more stable in the buffer containing 30% BS than the unmodified DNA 7.

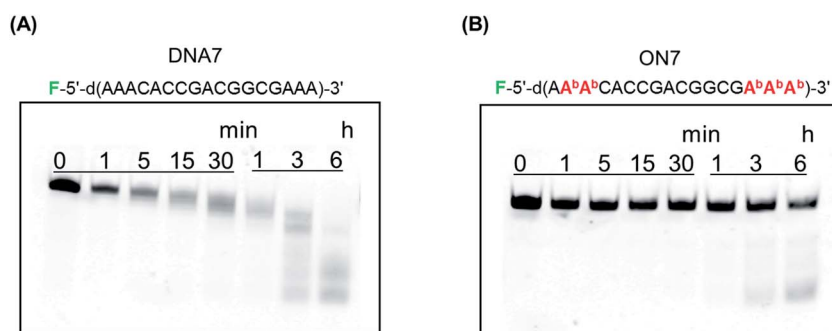


Fig. 4 Polyacrylamide gel electrophoresis of (A) DNA 7 and (B) ON 7 treated in a buffer containing 30% BS. F denotes a fluorescein. Experimental conditions are described in the Experimental.

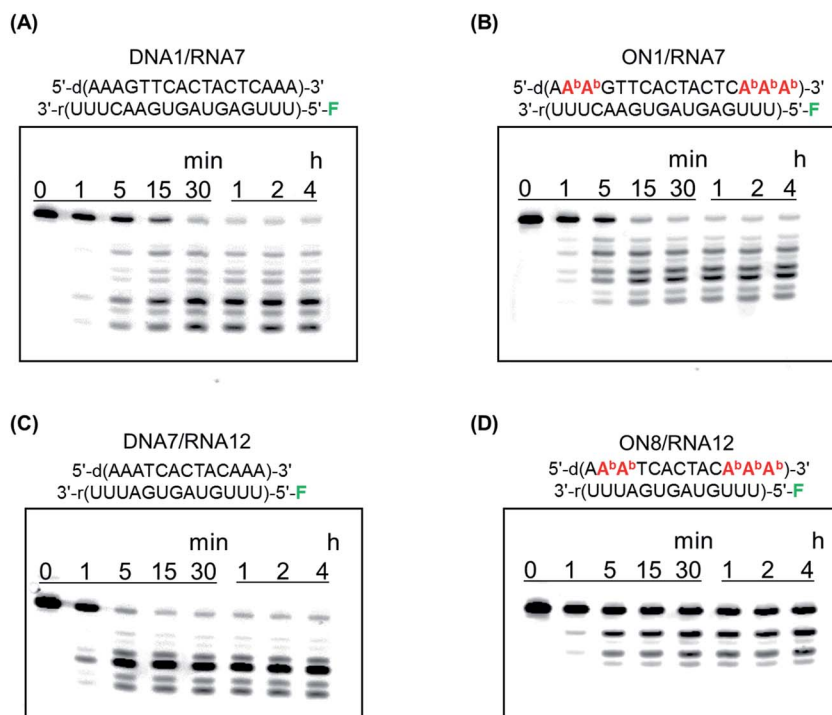


Fig. 5 Polyacrylamide gel electrophoresis of (A) DNA 1/RNA 7 duplex, (B) ON 1/RNA 7 duplex, (C) DNA 7/RNA 12 duplex, and (D) ON 8/RNA 12 duplex hydrolyzed by human recombinant RNase H1. F denotes a fluorescein. Experimental conditions are described in the Experimental.



Degradation by human recombinant RNase H1

To function as therapeutics, AONs need to elicit RNase H activity in human cells. We previously reported that a duplex between a BGNA–DNA chimera and a complementary RNA could act as a substrate for *Escherichia coli* RNase H.^{5c} In this study, we examined cleavage of RNAs by human recombinant RNase H1. A DNA fragment of *E. coli mal E* gene encoding maltose binding protein (MBP) and a fragment encoding catalytic domain of human RNase H1 (residues 136–286, H-domain) were subcloned into the multiple cloning site of pET15b, resulting in pET-MBP-hRNase H1 (H-domain) vector for expressing MBP-hRNase H1 (H-domain) fusion protein. Human recombinant RNase H1 was expressed from the pET-MBP-hRNase H1 (H-domain) vector in *E. coli* BL21(DE3) competent cells. The duplex between DNA 1 or ON 1 and RNA 7, labeled with fluorescein at the 5'-end, or DNA 5 or ON 5 and RNA 8, also labeled with fluorescein at the 5'-end, was incubated in a buffer containing human recombinant RNase H, and the products were analyzed by polyacrylamide gel electrophoresis under denaturing conditions (Fig. 5). The initial rates of degradation of the DNA 1/RNA 7 and ON 1/RNA 7 duplexes (17-mers) were 206 pmol min⁻¹ and 215 pmol min⁻¹, respectively, while those of the DNA 5/RNA 8 and ON 5/RNA 8 duplexes (13-mers) were 302 pmol min⁻¹ and 100 pmol min⁻¹, respectively. Thus, the ability of ON 5 (13-mer) containing five analogs to elicit human RNase H activity was found to be weaker than that of the unmodified DNA 5, while the ability of the ON 1 (17-mer) containing five analogs to elicit human RNase H activity was comparable with that of the unmodified DNA 1.

X-ray co-crystal structural analysis has shown that human RNase H recognizes eight base-pairs in a DNA/RNA duplex.⁸ ON 1 contains eleven contiguous unmodified nucleoside residues, while ON 5 has seven consecutive unmodified nucleoside moieties. This may partially explain why the ON 5/RNA duplex is a poor substrate for human RNase H. Furthermore, CD spectra suggested that the global conformation of the ON 1/RNA duplex was similar to that of the unmodified DNA/RNA duplex, while the global conformation of the ON 5/RNA duplex was different from that of the unmodified DNA/RNA duplex. This is also likely to influence the abilities of the ONs to elicit human RNase H activity.

Conclusions

We have synthesized nucleoside analogs with a thymine- or guanine-benzene-glycol structure. We have synthesized DNAs incorporating four nucleoside analogs – thymidine, cytidine, adenosine, and guanosine – at the 5'- and 3'-regions of the DNAs. We found that the benzene-glycol nucleic acid (BGNA)–DNA chimeras formed thermally and thermodynamically stable duplexes with complementary RNAs, and these are more stable in bovine serum than unmodified DNA. Furthermore, we found that duplexes formed by BGNA–DNA chimeras and complementary RNAs are good substrates for human RNase H. Based on these results, we believe that BGNA–DNA chimeras are good candidates for the development of therapeutic antisense molecules.

Experimental

General remarks

The chemicals were purchased from Sigma-Aldrich, Kanto Chemical Co. Inc., Tokyo Chemical Industry Co., Ltd., Nacalai Tesque Inc., Wako Pure Chemical Industries, Ltd., and Glen Research Co. All chemicals were used without further purification. Gibco® bovine serum was purchased from Thermo Fisher Scientific Inc. A pET15b expression vector was purchased from Novagen, Merck Millipore, Japan. *E. coli* BL21(DE3) competent cells were purchased from NEW ENGLAND BioLabs® Japan Inc. CDCl₃ (CIL) or DMSO-*d*₆ (CIL) was used as a solvent for obtaining NMR spectra. Chemical shifts (δ) are given in parts per million (ppm) downfield from (CH₃)₄Si (δ 0.00 for ¹H NMR in CDCl₃), or a solvent (for ¹³C NMR and ¹H NMR in DMSO-*d*₆) as an internal reference with coupling constants (*J*) in Hz. The abbreviations s, d, and q signify singlet, doublet, and quartet, respectively.

(S)-N³-Benzoyl-1-{4-[1,2-bis(*tert*-butyldimethylsilyloxy)ethyl]phenyl}thymine (6). To a mixture of (S)-4-[1,2-bis(*tert*-butyldimethylsilyloxy)ethyl]phenylboronic acid (5) (0.56 g, 1.35 mmol) in CH₂Cl₂ (5 mL) was added Cu(OAc)₂ (0.25 g, 1.35 mmol, 1 eq.), N³-benzoylthymine (0.62 g, 2.71 mmol, 2 eq.), and pyridine (0.21 mL, 2.71 mmol, 2 eq.). The mixture was vigorously stirred at room temperature for 30 min. The resulting mixture was filtered through a Celite pad, and the eluent was partitioned between CHCl₃ and brine. The organic layer was dried (Na₂SO₄), and concentrated. The residue was purified by column chromatography (SiO₂, 20% EtOAc in hexane) to give 6 (0.39 g, 0.65 mmol, 49%): ¹H NMR (600 MHz, CDCl₃): δ = -0.05 to -0.03 (m, 9H), 0.07 (s, 3H), 0.85 (s, 9H), 0.88 (s, 9H), 2.02 (d, *J* = 1.4 Hz, 3H), 3.51 (dd, *J* = 4.8 and 10.2, 1H), 3.64 (dd, *J* = 6.9 and 10.3, 1H), 4.72 (t, *J* = 5.8, 1H), 7.29 (s, 1H), 7.32 (d, *J* = 8.9, 2H), 7.45 (d, *J* = 8.2, 2H), 7.50 (t, *J* = 7.9, 2H), 7.64 (t, *J* = 6.8, 1H), 7.98 (d, *J* = 6.8, 2H). ¹³C NMR (151 MHz, CDCl₃): δ = -5.6, -5.5, -4.9, -4.8, 12.4, 18.2, 18.4, 25, 8, 25.9, 69.7, 75.3, 111.1, 125.7, 127.5, 129.1, 130.5, 131.7, 134.9, 137.2, 140.5, 143.8, 149.3, 163.2, 168.9. HRMS (ESI): calcd for C₃₂H₄₆N₂O₅Si₂, 633.2582 [M + K]⁺, found 633.2600.

(S)-1-[4-(1,2-Dihydroxyethyl)phenyl]thymine (3). Compound 6 (0.51 g, 0.86 mmol) was dissolved in THF (5 mL). TBAF (1 M in THF, 2.58 mL) was added to the solution, and the mixture was stirred at room temperature for 16 h. The solvent was evaporated *in vacuo*, and the resulting residue was dissolved in MeOH (5 mL). Concentrated NH₄OH (15 mL) was added to the solution, and the mixture was stirred at room temperature for 4 h. The solvent was evaporated *in vacuo*, and the resulting residue was purified by column chromatography (SiO₂, 5–10% MeOH in CHCl₃) to give 3 (0.18 g, 0.69 mmol, 80%): ¹H NMR (600 MHz, DMSO-*d*₆): δ = 3.19 (s, 1H), 4.02 (s, 1H), 4.13 to 4.15 (m, 3H), 5.27 (dd, *J* = 5.5 and 10.3, 1H), 5.46 (t, *J* = 5.9, 1H), 6.02 (d, *J* = 4.1, 1H), 8.03 (d, *J* = 8.2, 2H), 8.12 (d, *J* = 8.2, 2H), 8.28 (s, 1H), 12.09 (s, 1H). ¹³C NMR (151 MHz, DMSO-*d*₆): δ = 11.8, 67.3, 73.3, 109.2, 126.2, 126.9, 137.6, 141.4, 143.4, 150.4, 164.3. HRMS (ESI): calcd for C₁₃H₁₄N₂O₄, 263.1032 [M + H]⁺, found 263.1049.



(S)-1-[4-[2-(4,4'-Dimethoxytrityloxy)-1-hydroxyethyl]phenyl]thymine (7). A mixture of **3** (0.15 g, 0.57 mmol) and DMTrCl (0.21 g, 0.63 mmol, 1.1 eq.) in pyridine (1.5 mL) was stirred at room temperature for 90 min. The mixture was partitioned between EtOAc and aqueous NaHCO₃ (saturated). The organic layer was washed with brine, dried (Na₂SO₄), and concentrated. The residue was purified by column chromatography (SiO₂, 66% EtOAc in hexane) to give **7** (0.32 g, 0.56 mmol, 99%): ¹H NMR (600 MHz, CDCl₃): δ = 1.94 (s, 3H), 2.92 (s, 1H), 3.23 (t, *J* = 9.3, 1H), 3.34 (dd, *J* = 3.4 and 9.6, 1H), 3.78 (s, 6H), 4.81 (dd, *J* = 3.4 and 8.3, 1H), 6.82 (d, *J* = 8.2, 4H), 7.11 (d, *J* = 1.4, 1H), 7.21 to 7.29 (m, 9H), 7.36 to 7.41 (m, 4H), 8.46 (s, 1H). ¹³C NMR (151 MHz, CDCl₃): δ = 12.3, 55.2, 69.0, 72.7, 86.7, 111.1, 113.2, 126.1, 126.9, 127.4, 127.9, 128.0, 130.0, 135.7, 137.8, 140.6, 141.4, 144.6, 150.0, 158.6, 163.7. HRMS (ESI): calcd for C₃₄H₃₂N₂O₆, 587.2158 [M + Na]⁺, found 587.2160.

(S)-1-[4-[1-(2-Cyanoethoxy)(*N,N*-diisopropylamino)phosphoryloxy]-2-(4,4'-dimethoxytrityloxy)ethyl]phenyl]thymine (8). Compound **7** (0.63 g, 1.10 mmol) was dissolved in CH₂Cl₂ (6 mL) containing *N,N*-diisopropylethylamine (0.57 mL, 3.32 mmol, 3 eq.). Chloro(2-cyanoethoxy)(*N,N*-diisopropylamino)phosphine (0.37 mL, 1.66 mmol, 1.5 eq.) was added to the solution, and the mixture was stirred at room temperature for 1 h. Aqueous NaHCO₃ (saturated) and CHCl₃ were added to the mixture, and the separated organic layer was washed with aqueous NaHCO₃ (saturated) and brine, dried (Na₂SO₄) and concentrated. The residue was purified by column chromatography (a neutralized SiO₂, 80% EtOAc in hexane) to give **8** (0.67 g, 0.87 mmol, 80%): ³¹P NMR (162 MHz, CDCl₃): δ = 149.6, 150.2. HRMS (ESI): calcd for C₄₃H₄₉N₄O₇P, 765.3417 [M + H]⁺, found 765.3423.

(S)-2-[*N,N*-Bis(*tert*-butoxycarbonyl)amino]-6-chloro-9-[4-[1,2-bis(*tert*-butyldimethylsilyloxy)ethyl]phenyl]purine (9). To a mixture of 2-[*N,N*-bis(*tert*-butoxycarbonyl)amino]-6-chloropurine (2.06 g, 5.58 mmol) and **5** (2.29 g, 5.58 mmol) in DMF (23 mL) was added Cu(OAc)₂ (2.03 g, 11.16 mmol, 2 eq.) and pyridine (0.9 mL, 11.16 mmol, 2 eq.). The mixture was vigorously stirred at room temperature for 63 h. The resulting mixture was partitioned between EtOAc and brine. The organic layer was dried (Na₂SO₄), and concentrated. The residue was purified by column chromatography (SiO₂, 20% EtOAc in hexane) to give **9** (1.28 g, 1.74 mmol, 31%): ¹H NMR (600 MHz, CDCl₃): δ = -0.03 (s, 3H), -0.01 (s, 6H), 0.10 (s, 3H), 0.86 (s, 9H), 0.91 (s, 9H), 1.44 (s, 18H), 3.55 (dd, *J* = 5.5 and 10.3, 1H), 3.70 (dd, *J* = 6.8 and 10.3, 1H), 4.77 (t, *J* = 6.2, 1H), 7.56 (d, *J* = 8.9, 2H), 7.65 (d, *J* = 8.2, 2H), 8.41 (s, 1H). ¹³C NMR (151 MHz, CDCl₃): δ = -5.5, -5.4, -4.8, -4.7, 18.3, 18.4, 25.8, 25.9, 27.9, 69.6, 75.2, 83.6, 122.8, 128.0, 130.4, 132.8, 144.0, 144.7, 150.6, 151.7, 152.1, 152.6. HRMS (ESI): calcd for C₃₅H₅₆ClN₅O₆Si₂, 734.3536 [M + H]⁺, found 734.3570.

(S)-9-[4-[1,2-Bis(*tert*-butyldimethylsilyloxy)ethyl]phenyl]guanine (10). A mixture of NaOMe (5 M in MeOH, 2.58 mL) and 2-mercaptoethanol (1.22 mL, 17.56 mmol, 10 eq.) was added to a solution of **9** (1.29 g, 1.75 mmol) in MeOH (26 mL). The mixture was refluxed for 26 h. The mixture was cooled and neutralized by CH₃CO₂H. The resulting mixture was partitioned between CHCl₃ and brine. The organic layer was dried (Na₂SO₄),

and concentrated. The residue was purified by column chromatography (SiO₂, 6% MeOH in CHCl₃) to give **10** (0.59 g, 1.15 mmol, 66%): ¹H NMR (600 MHz, CDCl₃): δ = -0.01 (s, 6H), 0.01 (s, 3H), 0.10 (s, 3H), 0.87 (s, 9H), 0.91 (s, 9H), 3.56 (dd, *J* = 5.5 and 10.3, 1H), 3.66 (dd, *J* = 7.6 and 9.7, 1H), 3.76 (t, *J* = 6.2, 1H), 6.56 (s, 2H), 7.48 (d, *J* = 8.2, 2H), 7.56 (d, *J* = 8.3, 2H), 7.75 (s, 1H), 12.32 (s, 1H). ¹³C NMR (151 MHz, CDCl₃): δ = -5.5, -5.3, -4.8, -4.7, 18.3, 18.4, 25.9, 26.0, 69.8, 75.4, 117.6, 123.0, 127.5, 133.8, 136.7, 142.7, 151.4, 154.0, 159.4. HRMS (ESI): calcd for C₂₅H₄₁N₅O₃Si₂, 516.2826 [M + H]⁺, found 516.2852.

(S)-9-[4-(1,2-Dihydroxyethyl)phenyl]guanine (4). Compound **10** (0.47 g, 0.91 mmol) was dissolved in THF (5 mL). TBAF (1 M in THF, 2.74 mL) was added to the solution, and the mixture was stirred at 45 °C for 4 h. The solvent was cooled and the crystals were corrected to give **4** (0.19 g, 0.65 mmol, 71%): ¹H NMR (400 MHz, DMSO-*d*₆): δ = 3.45 (s, 1H), 4.59 (s, 1H), 4.76 (s, 1H), 5.34 (s, 1H), 6.52 (s, 2H), 7.46 (d, *J* = 7.8, 2H), 7.60 (d, *J* = 7.8, 2H), 7.98 (s, 1H), 10.73 (s, 1H). ¹³C NMR (101 MHz, DMSO-*d*₆): δ = 67.3, 73.2, 117.1, 123.4, 127.2, 133.7, 136.5, 143.0, 151.0, 154.0, 157.1. HRMS (ESI): calcd for C₁₃H₁₃N₅O₃, 310.0916 [M + Na]⁺, found 310.0938.

(S)-9-[4-[1,2-Bis(*tert*-butyldimethylsilyloxy)ethyl]phenyl]-*N*²-[[*N,N*-dimethyl]aminomethylene]guanine (11). A mixture of **10** (0.56 g, 1.08 mmol) and *N,N*-dimethylformamide dimethyl acetal (0.72 mL, 5.43 mmol, 5 eq.) in MeOH (5.6 mL) was stirred at 40 °C for 7 h. The resulting mixture was partitioned between CHCl₃ and brine. The organic layer was dried (Na₂SO₄), and concentrated. The residue was purified by column chromatography (SiO₂, 6% MeOH in CHCl₃) to give **11** (0.51 g, 0.90 mmol, 83%): ¹H NMR (600 MHz, CDCl₃): δ = -0.03 (s, 3H), -0.02 (s, 3H), -0.01 (s, 3H), 0.10 (s, 3H), 0.86 (s, 9H), 0.90 (s, 9H), 3.11 (s, 3H), 3.15 (s, 3H), 3.55 (dd, *J* = 5.5 and 9.6, 1H), 3.69 (dd, *J* = 6.8 and 10.3, 1H), 4.76 (t, *J* = 6.2, 1H), 7.49 (d, *J* = 8.9, 2H), 7.56 (d, *J* = 8.3, 2H), 7.83 (s, 1H), 8.54 (s, 1H), 9.42 (s, 1H). ¹³C NMR (151 MHz, CDCl₃): δ = -5.5, -5.4, -4.8, -4.7, 18.3, 18.4, 25.8, 25.9, 35.2, 41.3, 69.7, 75.3, 121.1, 123.2, 127.5, 134.0, 137.5, 142.7, 150.0, 156.8, 158.0. HRMS (ESI): calcd for C₂₈H₄₆N₆O₃Si₂, 571.3248 [M + H]⁺, found 571.3256.

(S)-*N*²-[[*N,N*-Dimethyl]aminomethylene]-9-[4-(1,2-dihydroxyethyl)phenyl]guanine (12). Compound **11** (0.22 g, 0.39 mmol) was dissolved in THF (2 mL). TBAF (1 M in THF, 1.18 mL, 3 eq.) was added to the solution, and the mixture was stirred at room temperature for 22 h. The solvent was evaporated *in vacuo*, and the resulting residue was purified by column chromatography (SiO₂, 6–16% MeOH in CHCl₃) to give **12** (0.13 g, 0.39 mmol, 99%): ¹H NMR (600 MHz, DMSO-*d*₆): δ = 3.01 (s, 3H), 3.13 (s, 3H), 3.47 to 3.50 (m, 2H), 4.60 (dd, *J* = 5.5 and 11.0, 1H), 4.77 (t, *J* = 5.5, 1H), 5.34 (d, *J* = 4.1, 1H), 7.50 (d, *J* = 8.3, 2H), 7.73 (d, *J* = 8.9, 2H), 8.16 (s, 1H), 8.48 (s, 1H), 11.42 (s, 1H). ¹³C NMR (151 MHz, DMSO-*d*₆): δ = 34.7, 40.6, 67.2, 73.3, 120.3, 122.7, 127.3, 133.7, 137.5, 142.7, 149.5, 157.4, 157.7, 157.9. HRMS (ESI): calcd for C₁₆H₁₈N₆O₃, 343.1519 [M + H]⁺, found 343.1521.

(S)-9-[4-[2-(4,4'-Dimethoxytrityloxy)-1-hydroxyethyl]phenyl]-*N*²-[[*N,N*-dimethyl]aminomethylene]guanine (13). A mixture of **12** (0.25 g, 0.74 mmol) and DMTrCl (0.38 g, 1.11 mmol, 1.5 eq.) in pyridine (2 mL) was stirred at room temperature for 2 h. The mixture was partitioned between EtOAc and aqueous NaHCO₃



(saturated). The organic layer was washed with brine, dried (Na_2SO_4), and concentrated. The residue was purified by column chromatography (SiO_2 , 5–16% MeOH in CHCl_3) to give **13** (0.47 g, 0.72 mmol, 98%): ^1H NMR (600 MHz, CDCl_3): δ = 2.88 (d, J = 2.7, 1H), 2.95 (s, 3H), 3.05 (s, 3H), 3.29 (t, J = 8.2, 1H), 3.38 (dd, J = 3.4 and 10.3, 1H), 3.78 (s, 6H), 4.86 (t, J = 3.4, 1H), 6.83 (d, J = 8.9, 4H), 7.21 to 7.30 (m, 7H), 7.40 (d, J = 7.6, 2H), 7.45 (d, J = 8.3, 2H), 7.53 (d, J = 8.3, 2H), 7.78 (s, 1H), 8.42 (s, 1H), 8.60 (s, 1H). ^{13}C NMR (151 MHz, CDCl_3): δ = 35.1, 41.1, 55.2, 69.0, 72.8, 86.6, 113.2, 121.0, 123.9, 126.9, 127.4, 127.9, 128.0, 130.0, 134.4, 135.8, 137.4, 140.8, 144.7, 150.1, 156.9, 157.9, 158.0, 158.6. HRMS (ESI): calcd for $\text{C}_{37}\text{H}_{36}\text{N}_6\text{O}_5$, 645.2825 $[\text{M} + \text{H}]^+$, found 645.2858.

(S)-9-(4-{1-[(2-Cyanoethoxy)(diisopropylamino)phosphanyloxy]-2-(4,4'-dimethoxytrityloxy)ethyl}phenyl)-*N,N*-dimethylaminomethylene]guanine (14). Compound **13** (0.43 g, 0.67 mmol) was dissolved in CH_2Cl_2 (4 mL) containing *N,N*-diisopropylethylamine (0.35 mL, 2.02 mmol, 3 eq.). Chloro(2-cyanoethoxy)(*N,N*-diisopropylamino)phosphine (0.23 mL, 1.01 mmol, 1.5 eq.) was added to the solution, and the mixture was stirred at room temperature for 1 h. Aqueous NaHCO_3 (saturated) and CHCl_3 were added to the mixture, and the separated organic layer was washed with aqueous NaHCO_3 (saturated) and brine, dried (Na_2SO_4) and concentrated. The residue was purified by column chromatograph (a neutralized SiO_2 , 5–9% MeOH in CHCl_3) to give **14** (0.54 g, 0.64 mmol, 96%): ^{31}P NMR (162 MHz, CDCl_3): δ = 149.3, 149.8. HRMS (ESI): calcd for $\text{C}_{46}\text{H}_{53}\text{N}_8\text{O}_6\text{P}$, 845.3904 $[\text{M} + \text{H}]^+$, found 845.3903.

Oligonucleotide synthesis

The synthesis was carried out with a DNA/RNA synthesizer by the phosphoramidite method. Deprotection of DNAs was performed in concentrated NH_4OH at 55 °C for 16 h. Deprotection of RNAs was performed in concentrated $\text{NH}_4\text{OH}/\text{EtOH}$ (3 : 1, v/v) at 55 °C for 4 h. 2'-*O*-TBDMS groups were removed by $\text{Et}_3\text{N} \cdot 3\text{HF}$ in DMSO at 65 °C for 90 min. The reaction was quenched with 0.1 M TEAA buffer (pH 7.0) and the mixture desalted using a Sep-Pak C18 cartridge. The oligonucleotides were purified by 20% PAGE containing 7 M urea and/or reversed-phase C-18 HPLC using a gradient of solution A (5% MeCN in 0.1 M TEAA buffer, pH 7.0) and solution B (50% MeCN in 0.1 M TEAA buffer, pH 7.0).

MALDI-TOF/MS analysis of ONs

The spectra were obtained with a time-of-flight mass spectrometer equipped with a nitrogen laser (337 nm, 3 ns pulse). A solution of 3-hydroxypicolinic acid (3-HPA) and diammonium hydrogen citrate in H_2O was used as the matrix. Data of synthetic ONs: DNA **1**: m/z = 5145.40 (calcd for $\text{C}_{166}\text{H}_{208}\text{N}_{65}\text{O}_{96}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5145.40); DNA **2**: m/z = 5098.51 (calcd for $\text{C}_{164}\text{H}_{208}\text{N}_{61}\text{O}_{98}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5097.35); DNA **3**: m/z = 5177.60 (calcd for $\text{C}_{166}\text{H}_{208}\text{N}_{65}\text{O}_{98}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5177.40); DNA **4**: m/z = 5127.60 (calcd for $\text{C}_{166}\text{H}_{210}\text{N}_{59}\text{O}_{100}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5127.37); DNA **5**: m/z = 3919.72 (calcd for $\text{C}_{127}\text{H}_{158}\text{N}_{53}\text{O}_{70}\text{P}_{12}$ $[\text{M} - \text{H}]^-$ 3919.62); DNA **6**: m/z = 3871.68 (calcd for $\text{C}_{125}\text{H}_{158}\text{N}_{49}\text{O}_{72}\text{P}_{12}$ $[\text{M} - \text{H}]^-$ 3870.57); DNA **7**: m/z = 5743.48 (calcd for

$\text{C}_{192}\text{H}_{229}\text{N}_{76}\text{O}_{101}\text{P}_{17}$ $[\text{M} - \text{H}]^-$ 5743.88); ON **1**: m/z = 5245.06 (calcd for $\text{C}_{181}\text{H}_{208}\text{N}_{65}\text{O}_{91}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5245.57); ON **2**: m/z = 5198.15 (calcd for $\text{C}_{179}\text{H}_{208}\text{N}_{61}\text{O}_{93}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5197.53); ON **3**: m/z = 5277.04 (calcd for $\text{C}_{181}\text{H}_{208}\text{N}_{65}\text{O}_{93}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5277.57); ON **4**: m/z = 5228.18 (calcd for $\text{C}_{181}\text{H}_{210}\text{N}_{59}\text{O}_{95}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5227.55); ON **5**: m/z = 4018.86 (calcd for $\text{C}_{142}\text{H}_{158}\text{N}_{53}\text{O}_{65}\text{P}_{12}$ $[\text{M} - \text{H}]^-$ 4018.79); ON **6**: m/z = 3970.97 (calcd for $\text{C}_{140}\text{H}_{158}\text{N}_{49}\text{O}_{67}\text{P}_{12}$ $[\text{M} - \text{H}]^-$ 3970.75); ON **7**: m/z = 5844.20 (calcd for $\text{C}_{207}\text{H}_{229}\text{N}_{76}\text{O}_{96}\text{P}_{17}$ $[\text{M} - \text{H}]^-$ 5844.05); RNA **1**: m/z = 5389.64 (calcd for $\text{C}_{161}\text{H}_{196}\text{N}_{59}\text{O}_{121}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5389.19); RNA **2**: m/z = 5467.74 (calcd for $\text{C}_{163}\text{H}_{198}\text{N}_{65}\text{O}_{119}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5467.27); RNA **3**: m/z = 5387.00 (calcd for $\text{C}_{161}\text{H}_{198}\text{N}_{61}\text{O}_{119}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5387.22); RNA **4**: m/z = 5436.10 (calcd for $\text{C}_{163}\text{H}_{198}\text{N}_{65}\text{O}_{117}\text{P}_{16}$ $[\text{M} - \text{H}]^-$ 5435.27); RNA **5**: m/z = 4080.14 (calcd for $\text{C}_{122}\text{H}_{148}\text{N}_{41}\text{O}_{95}\text{P}_{12}$ $[\text{M} - \text{H}]^-$ 4080.39); RNA **6**: m/z = 4158.26 (calcd for $\text{C}_{124}\text{H}_{150}\text{N}_{47}\text{O}_{93}\text{P}_{12}$ $[\text{M} - \text{H}]^-$ 4158.47); RNA **7**: m/z = 5927.30 (calcd for $\text{C}_{188}\text{H}_{221}\text{N}_{60}\text{O}_{130}\text{P}_{17}$ $[\text{M} - \text{H}]^-$ 5927.65); RNA **8**: m/z = 4618.50 (calcd for $\text{C}_{149}\text{H}_{173}\text{N}_{42}\text{O}_{104}\text{P}_{13}$ $[\text{M} - \text{H}]^-$ 4618.85).

Thermal denaturation study

The solution containing the duplex in a buffer comprising 10 mM sodium phosphate (pH 7.0) and 100 mM NaCl was heated at 100 °C for 5 min, cooled gradually to an appropriate temperature, and then used for the thermal denaturation study. The thermally induced transition of each mixture was monitored at 260 nm with a UV/vis spectrometer fitted with a temperature controller in quartz cuvettes with a path length of 1.0 cm and a 3.0 μM duplex concentration in a buffer of 10 mM sodium phosphate (pH 7.0) and 0.1 M NaCl. The sample temperature was increased by 0.5 °C min^{-1} . The thermodynamic parameters of the duplexes on duplex formation were determined by calculations based on the slope of a $1/T_m$ vs. $\ln(C_T/4)$ plot, where C_T (1, 3, 6, 12, 15, 21, 30, and 60 μM) is the total concentration of single strands.

Partial hydrolysis of ONs in a buffer containing bovine serum

Each ON (3000 pmol) labeled with fluorescein at the 5'-end was incubated with 30% bovine serum in a buffer (150 μL) comprised of 0.1 M Tris-HCl (pH 8.0) and 20 mM MgCl_2 at 37 °C. Aliquots (5 μL) were taken at 0, 0.5, 1, 5, 30, 60, 180, 360 min and mixed with the loading buffer (15 μL) containing Tris-borate-EDTA (TBE) buffer and 20% glycerin on ice. Each sample was analyzed by 20% denaturing PAGE at room temperature at 20 mA for 2 h. The gel was visualized by use of a Luminescent Image analyzer LAS-4000 (Fujifilm).

DNA constructs

A DNA fragment of *E. coli mal E* gene encoding maltose binding protein (MBP) and a fragment encoding catalytic domain of human RNase H1 (residues 136–286, H-domain) were subcloned into the multiple cloning site of pET15b, resulting in pET-MBP-hRNase H1 (H-domain) vector for expressing MBP-hRNase H1 (H-domain) fusion protein. Between the two genes, a specific cleavage site for human Rhinovirus 3C (HRV3C) protease was



engineered. All the DNA fragments were prepared by the standard PCR method.

Protein expression and purification

The pET-MBP-hRNase H1 (H-domain) vector was transformed into *E. coli* BL21(DE3). The cells were cultured in Luria-Bertani medium at the 37 °C until OD₆₀₀ reached to approximately 0.4. Final 0.2 mM of isopropyl-β-D-galactoside was added to induce the recombinant protein expression, and the temperature was shifted to 20 °C. After 20 hours, the cells were harvested. The bacterial cells were suspended to the lysis buffer (pH 7.0) containing 40 mM sodium phosphate, 1 M NaCl, 0.5 mM EDTA and 2 mM dithiothreitol (DTT). Then, the cells were lysed by sonication, and the supernatant was collected by centrifugation. First, a cleared supernatant was passed through DEAE-Sepharose® (GE Healthcare Life Sciences Corp, Piscataway, NJ, USA), and then the fusion protein was captured by Amylose resin (New England Biolabs, Ipswich, MA, USA). From the fusion protein, MBP tag was removed by on-column digestion by adding HRV3C protease at room temperature. The digested hRNase H1 (H-domain) was purified by cation exchange chromatography by using SP-Sepharose® (GE Healthcare Life Sciences) and further purified by gel-filtration by Superdex® 75 pg (GE Healthcare Life Sciences). Finally, the enzyme was dialyzed against 20 mM sodium phosphate buffer (pH 7.0) containing 100 mM NaCl and 5 mM DTT.

Hydrolysis of RNA with human recombinant RNase H1

The solution containing ON (600 pmol) and RNA (3000 pmol) labeled with fluorescein at the 5'-end in a buffer comprised of 50 mM Tris-HCl (pH. 8.0), 75 mM KCl, 3 mM MgCl₂ and 10 mM dithiothreitol was heated at 100 °C for 5 min. The reaction mixtures were then cooled gradually to an appropriate temperature. Then human RNase H1 (2 pmol) was added to the

solution, and the mixture was incubated at 37 °C. Aliquots (5 μL) were taken at 0, 1, 5, 15, 30, 60, 120 and 240 min, and mixed with formamide (15 μL) on ice. Each sample was analyzed by 20% denaturing PAGE at room temperature at 20 mA for 2 h. The gel was visualized by use of a Luminescent Image analyzer LAS-4000 (Fujifilm).

Notes and references

- (a) P. C. Zamecnik and M. L. Stephenson, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 280–284; (b) M. L. Stephenson and P. C. Zamecnik, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 285–288.
- (a) M. Muthiah, *Antisense Nucleic Acid Drug Dev.*, 2002, **12**, 103–128; (b) T. Aboul-Fadl, *Curr. Med. Chem.*, 2005, **12**, 2193–2214.
- P. M. D. Moreno and A. P. Pêgo, *Front. Chem.*, 2014, **2**, 87, DOI: 10.3389/chem.2014.00087.
- Editorial, *Nat. Biotechnol.*, 2016, **34**, 1078, DOI: 10.1038/nbt.3733.
- (a) Y. Ueno, T. Kato, K. Sato, Y. Ito, M. Yoshida, T. Inoue, A. Shibata, M. Ebihara and Y. Kitade, *J. Org. Chem.*, 2005, **70**, 7925–7935; (b) Y. Ueno, A. Kawamura, K. Takasu, S. Komatsuzaki, T. Kato, S. Kuboe, Y. Kitamura and Y. Kitade, *Org. Biomol. Chem.*, 2009, **7**, 2761–2769; (c) N. Niwa, K. Ueda and Y. Ueno, *Eur. J. Org. Chem.*, 2016, 2435–2443.
- Y. Yue, Z.-G. Zheng, B. Wu, C.-Q. Xia and X.-Q. Yu, *Eur. J. Org. Chem.*, 2005, 5154–5157.
- M. F. Jacobsen, M. M. Knudsen and K. V. Gothelf, *J. Org. Chem.*, 2006, **71**, 9183–9190.
- M. Nowotny, S. A. Gaidamakov, R. Ghirlando, S. M. Cerritelli, R. J. Crouch and W. Yang, *Mol. Cell*, 2007, **28**, 264–276.

