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 Hiroa d and Yoshihito Uenoa,b,c

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.

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.5 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 phosphoramidate 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).3 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 aryloboronic 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 N1 position of thymine and the aryloboronic acid, we used an N7-protected thymine, N7-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.8,9 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.
N^3-Benzoylthymine was coupled with the arylboronic acid derivative 5 in the presence of Cu(OAc)_2 and pyridine 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 phosphorylated 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 Cu(OAc)_2 and pyridine in DMF to form an N^9-arylpurine 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. 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.
Table 1 Sequences of duplexes and ONs, and \( T_m \)s and thermodynamic parameters of the duplexes

| Duplex | ON  | Sequencesa | \( T_m \) (°C) | \( \Delta T_m \) (°C) | \( \Delta H^o \) (kcal mol\(^{-1}\)) | \( \Delta S^o \) (cal mol\(^{-1}\) K\(^{-1}\)) | \( \Delta G^o_{10} \) (kcal mol\(^{-1}\)) |
|--------|-----|------------|----------------|------------------|-------------------|-----------------|------------------|
| Duplex 1 | DNA 1 | 5'-d(AAAGTTCACTACTCAA)-3' | 52.3 | -98.4 | -275.2 | -13.1 |
| RNA 1 | 3'-r(UUUCAAGUGAUGGUU)-5' |
| Duplex 2 | ON 1 | 5'-d(AAAAGTTCACTACTCAAA)-3' | 55.6 | +3.3 | -94.7 | -259.8 | -14.1 |
| RNA 1 | 3'-r(UUUCAAGUGAUGGUU)-5' |
| Duplex 3 | DNA 2 | 5'-d(AACAGTTCACTACTAA)-3' | 58.0 | -144.1 | -406.6 | -18.0 |
| RNA 2 | 3'-r(UUCAAGUGAUGAGUU)-5' |
| Duplex 4 | ON 2 | 5'-d(AACAGTTCACTACTCAA)-3' | 57.9 | -0.1 | -113.6 | -315.0 | -16.0 |
| RNA 2 | 3'-r(UUCAAGUGAUGAGUU)-5' |
| Duplex 5 | DNA 3 | 5'-d(AAGGTTCACTACTCA)-3' | 57.5 | -132.5 | -373.5 | -16.8 |
| RNA 3 | 3'-r(UUCUCAGUGAGUUC)-5' |
| Duplex 6 | ON 3 | 5'-d(AAAAGTTCACTACTCAAA)-3' | 56.4 | -1.1 | -112.3 | -313.0 | -15.3 |
| RNA 3 | 3'-r(UUCUCAGUGAGUUC)-5' |
| Duplex 7 | DNA 4 | 5'-d(TAGGTTCACTACTA)-3' | 54.8 | -122.0 | -343.9 | -15.4 |
| RNA 4 | 3'-r(UUCAAGUGAUGAGU)-5' |
| Duplex 8 | ON 4 | 5'-d(TATGTTCACTACTCAAA)-3' | 56.0 | +1.2 | -117.6 | -329.6 | -15.5 |
| RNA 4 | 3'-r(UUCAAGUGAUGAGU)-5' |
| Duplex 9 | DNA 5 | 5'-d(AAAATCACTACTAC)-3' | 36.8 | -89.6 | -260.9 | -8.7 |
| RNA 5 | 3'-r(UUUGAGUAGUAC)-5' |
| Duplex 10 | ON 5 | 5'-d(AAAATCACTACTACAA)-3' | 35.0 | -1.8 | -66.0 | -186.1 | -8.4 |
| RNA 5 | 3'-r(UUUGAGUAGUAC)-5' |
| Duplex 11 | DNA 6 | 5'-d(ACATCTAATAC)-3' | 47.7 | -108.6 | -310.5 | -12.4 |
| RNA 6 | 3'-r(UUUGAGUAGUAC)-5' |
| Duplex 12 | ON 6 | 5'-d(ACAAATCACTACTCAAA)-3' | 41.8 | -5.9 | -99.9 | -288.8 | -10.4 |
| RNA 6 | 3'-r(UUUGAGUAGUAC)-5' |
| DNA 7 | 5'-d(AAACCGCAGCGGCGA)-3' |
| ON 7 | 5'-d(AACCGCAGCGGCGA)-3' |
| RNA 7 | 5'-r(UUUGAGUAGUAGU)-3' |
| RNA 8 | 5'-r(UUUGAGUAGUAC)-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 \( T_m \) 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. \( \ln(C_r/4) \), where \( C_r \) 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 (\( \Delta T_m = -1.1 \sim +3.3 \, ^\circ \text{C}, \Delta \Delta G_{110} = +2.0 \sim -1.0 \, \text{kcal mol}^{-1} \)). 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 \( \Delta T_m \) and \( \Delta \Delta G_{110} \) values between duplexes 1 and 2 were +3.3 \, ^\circ \text{C} \) and \(-1.0 \, \text{kcal mol}^{-1} \), respectively, while those between duplexes 9 and 10 were \(-1.8 \, ^\circ \text{C} \) and +0.3 \, \text{kcal mol}^{-1} \), respectively. Similarly, the \( \Delta T_m \) and \( \Delta \Delta G_{110} \) values between duplexes 3 and 4 were \(-0.1 \, ^\circ \text{C} \) and +2.0 \, \text{kcal mol}^{-1} \), respectively, while those between duplexes 11 and 12 were \(-5.9 \, ^\circ \text{C} \) and +2.0 \, \text{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 \( \Delta T_m \) \( (T_m \) of duplex containing mismatched base pair \(-T_m \) of complementary duplex) values for the unmodified duplexes were \(-0.6 \sim -12.3 \, ^\circ \text{C} \) \), whereas those for the duplexes containing the analogs were \(-0.3 \sim -5.2 \, ^\circ \text{C} \). 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 ℃. Concentration of duplexes: 4 μM.
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 BGNA–DNA chimeras are more resistant to hydrolysis by a snake venom phosphodiesterase (a 3′-exonuclease) than unmodified DNA.\textsuperscript{57} To be able to function as drugs, AONs must be stable in blood, which contains various enzymes, including nuclease. Therefore, we investigated the stability of BGNA–DNA 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 BGNA–DNA chimera (ON 7) was approximately 20-fold more stable in the buffer containing 30% BS than the unmodified DNA 7.
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. 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. 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-d₆ (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 ³¹H NMR in DMSO-d₆) as an internal reference with coupling constants (J) in Hz. The abbreviations s, d, t, q, and p signify singlet, doublet, quartet, and quintet, respectively.

(5)-N²-Benzoyl-1-[4-[[1,2-bis(tet-butyldimethylsilyloxy)ethyl]phenyl]thymine (6). To a mixture of (5)-[4-[[1,2-bis(tet-butyldimethylsilyloxy)ethyl]phenyl]thymine (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, 5H), 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 = 56.4, 56.5, 56.6, 56.7, 61.3, 62.0, 71.9, 119.8, 122.0, 122.4, 127.1, 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.

(5)-1-[4-(1,2-Dihydroxyselyle)phenyl]thymine (3). Compound 6 (0.51 g, 0.86 mmol) was dissolved in THF (5 mL). TBAP (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-d₆): 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-d₆): 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-(2-[4-(1,1-Dimethylsilyloxy)-1-hydroxyethyl]phenyl)-2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
2-
(saturated). The organic layer was washed with brine, dried (Na₂SO₄), and concentrated. The residue was purified by column chromatography (SiO₂, 5–16% MeOH in CHCl₃) to give 13 (0.47 g, 0.72 mmol, 98%): ¹H NMR (600 MHz, CDCl₃): 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). ¹³C NMR (151 MHz, CDCl₃): 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 C₁₆₃H₁₉₈N₆₅O₁₁₇P₁₆ [M + H]⁺, found 645.2858.

[(S)-9-4-1-(2-Cyanoethyl)diisopropyloaminoephosphonoyloxy]-2-(4,4’-dimethoxytrityloxy)ethyl]phenyl)-N²[(N,N-dimethylaminomethylene)guanine (14). Compound 13 (0.43 g, 0.67 mmol) was dissolved in CH₂Cl₂ (4 mL) containing N,N-diisopropylethylamine (0.35 mL, 2.02 mmol, 3 eq.). Chloro(2-cyanoethyl)(N,N-diisopropylphosphine (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 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, 5–9% MeOH in CHCl₃) to give 14 (0.54 g, 0.64 mmol, 96%): ¹³P NMR (162 MHz, CDCl₃): δ = 149.3, 149.8. HRMS (ESI): calcd for C₁₄₉H₁₄₉N₆₅O₉₆P₁₆ [M + H]⁺, found 645.3903.

Oligonucleotide synthesis

The synthesis was carried out with a DNA/RNA synthesizer by the phosphoramidite method. Deprotection of DNAs was performed in concentrated NH₄OH at 55 °C for 16 h. Deprotection of RNAs was performed in concentrated NH₄OH/EtOH (3 : 1, v/v) at 55 °C for 4 h. 2’-O-TBDMS groups were removed by Et₃N-N3-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 C18 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-hydroxycipionic acid (3-HPA) and diammonium hydrogen citrate in H₂O was used as the matrix. Data of synthetic ONs: DNA 1: m/z = 5145.40 (calcd for C₁₉₂H₂₂₉N₇₆O₉₁P₁₇ [M – H]⁻ 5743.88); ON 1: m/z = 5245.06 (calcd for C₁₉₁H₂₂₉N₇₆O₉₁P₁₇ [M – H]⁻ 5245.57); ON 2: m/z = 5198.15 (calcd for C₁₇₉H₂₀₈N₅₉O₅₃P₁₆ [M – H]⁻ 5197.53); ON 3: m/z = 5277.04 (calcd for C₁₈₁H₂₀₈N₅₉O₅₃P₁₆ [M – H]⁻ 5277.57); ON 4: m/z = 5228.18 (calcd for C₁₄₁H₂₁₆N₉₃O₉₆P₁₆ [M – H]⁻ 5227.55); ON 5: m/z = 4018.86 (calcd for C₁₄₁H₂₁₆N₉₃O₉₆P₁₆ [M – H]⁻ 4018.79); ON 6: m/z = 3970.97 (calcd for C₁₄₀H₂₅₈N₉₅O₉₂P₁₂ [M – H]⁻ 3970.75); ON 7: m/z = 5844.20 (calcd for C₁₄₁H₂₁₆N₉₃O₉₆P₁₆ [M – H]⁻ 5844.03); RNA 1: m/z = 5389.64 (calcd for C₁₆₁H₂₆₉N₅₉O₅₃P₁₆ [M – H]⁻ 5389.19); RNA 2: m/z = 5467.74 (calcd for C₁₆₃H₂₆₉N₅₉O₅₃P₁₆ [M – H]⁻ 5467.27); RNA 3: m/z = 5387.00 (calcd for C₁₆₁H₂₆₉N₅₉O₅₃P₁₆ [M – H]⁻ 5387.22); RNA 4: m/z = 5436.10 (calcd for C₁₆₃H₂₆₉N₅₉O₅₃P₁₆ [M – H]⁻ 5435.27); RNA 5: m/z = 4080.14 (calcd for C₁₂₂H₁₄₁N₄₁O₉₃P₁₂ [M – H]⁻ 4080.39); RNA 6: m/z = 4158.26 (calcd for C₁₂₄H₁₅₀N₄₉O₉₃P₁₂ [M – H]⁻ 4158.47); RNA 7: m/z = 5927.30 (calcd for C₁₈₈H₂₂₁N₆₀₃O₅₁P₁₇ [M – H]⁻ 5927.65); RNA 8: m/z = 4618.50 (calcd for C₁₄₁H₁₇₃N₄₉O₁₀₄P₁₃ [M – 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/ln(Th/4) plot, where Th (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 MgCl₂ 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**

1. (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.
2. (a) M. Muthiah, *Antisense Nucleic Acid Drug Dev.*, 2002, 12, 103–128; (b) T. Aboul-Fadl, *Curr. Med. Chem.*, 2005, 12, 2193–2214.
3. P. M. D. Moreno and A. P. Pégo, *Front. Chem.*, 2014, 2, 87, DOI: 10.3389/chem.2014.00087.
4. Editorial, *Nat. Biotechnol.*, 2016, 34, 1078, DOI: 10.1038/nbt.3733.
5. (a) Y. Ueno, T. Kato, K. Sato, 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.
6. Y. Yue, Z.-G. Zheng, B. Wu, C.-Q. Xia and X.-Q. Yu, *Eur. J. Org. Chem.*, 2005, 5154–5157.
7. M. F. Jacobsen, M. M. Knudsen and K. V. Gothelf, *J. Org. Chem.*, 2006, 71, 9183–9190.
8. M. Nowotny, S. A. Gaidamakov, R. Ghirlando, S. M. Cerritelli, R. J. Crouch and W. Yang, *Mol. Cell.*, 2007, 28, 264–276.