Organic & Biomolecular Chemistry

Alteration of target cleavage patterns and off-target reduction of antisense oligonucleotides incorporating 2-*N*-carbamoyl- or (2-pyridyl)guanine

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-ART-04-2023-000574.R1
Article Type:	Paper
Date Submitted by the Author:	26-May-2023
Complete List of Authors:	Kanagawa, Takayuki; tokyo insititute of technology, Life Science Koyama, Aya; tokyo insititute of technology, Life Science Masaki, Yoshiaki; Tokyo Institute of Technology, Department of Life Science and Technology Seio, Kohji; tokyo insititute of technology, Life Science

ARTICLE

Alteration of target cleavage patterns and off-target reduction of antisense oligonucleotides incorporating 2-*N*-carbamoyl- or (2-pyridyl)guanine

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Takayuki Kanagawa,[a] Aya Koyoma,[a] Yoshiaki Masaki,*[ab] Kohji Seio*[a]

* Corresponding author: seio.k.aa@m.titech.ac.jp; ymasaki@bio.titech.ac.jp

Antisense oligonucleotides (ASOs) are therapeutic modalities that are successfully used as pharmaceuticals. However, there remains a concern that treatment with ASOs may cleave mismatched RNAs other than the target gene, leading to numerous alterations in gene expression. Therefore, improving the selectivity of ASOs is of paramount importance. Our group has focused on the fact that guanine forms stable mismatched base pairs and has developed guanine derivatives with modifications at the 2-amino group, which potentially change the mismatch recognition ability of guanine and the interaction between ASO and RNase H. In this study, we evaluated the properties of ASOs containing two guanine derivatives, 2-*N*-carbamoyl-guanine and 2-*N*-(2-pyridyl)guanine. We conducted ultraviolet (UV) melting experiments, RNase H cleavage assays, *in vitro* knockdown assays, and off-target transcriptome analyses using DNA microarrays. Our results indicate that the target cleavage pattern of RNase H was altered by the modification with guanine. Furthermore, global transcript alteration was suppressed in ASO incorporating 2-*N*-(2-pyridyl)guanine, despite a decrease in the thermal mismatch discrimination ability. These findings suggest that chemical modifications of the guanine 2-amino group have the potential to suppress hybridization-dependent off-target effects and improve ASO selectivity.

[a] *Department of Life Science and Technology, Tokyo Institute of Technology 4259-J2-16 Nagatsuta, Midori, Yokohama, Kanagawa, 226-8501, Japan.*
[b] *PRESTO, JST, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*
Kohji Seio
E-mail: seio.k.aa@m.titech.ac.jp
Yoshiaki Masaki
E-mail: ymasaki@bio.titech.ac.jp
†Supplementary Information is available: See DOI: 10.1039/x0xx00000x

Introduction

Gapmer antisense oligonucleotides (ASO) are short, single-stranded oligonucleotides that hybridize to complementary target mRNA and promote RNA cleavage by RNase H.[1] They are promising as a means of expanding the range of therapeutic options because they modulate, at the mRNA level, the expression of proteins that were previously undruggable with small-molecule drugs.[2]

Several gapmers have already been launched commercially[2,3] and the expectation is growing. However, safety issues remain to be addressed[4,5], although the mechanisms underlying these side effects have been extensively studied.[6–10] The degradation of nontarget RNAs by ASO plays an important role in one of these mechanisms. ASOs bind to RNAs with near-complementary sequences, resulting in their degradation.[11–13] This phenomenon, referred to as hybridization-dependent off-target effect, has been implicated as the cause of toxicity such as hepatotoxicity.[9,10] To address these concerns, the development of chemical modifications that alter the target selectivity is important. Previous studies have demonstrated the feasibility of altering the pattern of cleavage with RNase H through the use of nucleotides, including modified sugars[14–16] or backbone[17–20] nucleotides, in the gap region of the gapmer. In addition, it is also reported that the incorporation of 2-thiothymine or 5-triazolylphenyuracil[21] in place of canonical thymine changes the pattern of RNase H-mediated cleavage. This suggests that it is important to develop new nucleobases that can change cleavage patterns. Nucleobases other than thymine are important for expanding the usefulness of modified bases in ASOs with various sequences. When hydrogen-bonding patterns between individual nucleobases are considered, guanine forms relatively stable mismatched base pairs with thymine, adenine, and guanine.[22] This may be a potential cause of unintended off-target cleavage. One way to address the guanine mismatch base pairing is through nucleobase modification, but because modification of the base moiety may inhibit Watson-Crick base pairing and reduce antisense activity, it is necessary to design new compounds.

In a previous study, several guanine derivatives were synthesized and incorporated into oligonucleotides, and their properties were evaluated, mainly regarding tRNA[23], oxidatively damaged bases,[24,25] and G-quadruplexes.[26,27] Several studies have focused on the derivatives that suppress CpG-mediated immune stimulation[28] or reduce ASO hepatotoxicity.[29] However, the impact of guanine modification on RNase H recognition and off-target hybridization has not been well studied. Our group has also developed nucleotide derivatives bearing a guanine with a 2-amino group, 2-N-carbamoylguanine (cmG),[30] and 2-N-(2-pyridiyl)guanine (pyG)[27] (Figure 1A). Although these guanine derivatives have modifications at the 2-amino group necessary for Watson-Crick base pairs, they can still form base pairs, probably because the substituent rotates opposite to the base pairing faces, as shown in Figure 1D. Interestingly, theoretical calculations suggest that the more stable conformation was the closed type (Figure 1D), which could not form a Watson-Crick base pair, but the carbamoyl or pyridyl residue rotated and adopted a less stable open-type conformation when three hydrogen bonds were formed with cytosine. The effect of these modifications on the thermal stability of DNA duplexes has also been investigated, showing that the cmG modification slightly improved base selectivity, whereas the pyG modification lowered selectivity. However, the application to ASOs has not been studied.[18] These results motivated us to verify how the modified guanines affect the hybridization properties of ASOs and RNase H cleavage patterns, and whether these differences led to unintended hybridization-dependent off-target properties. In this study, we synthesized ASOs incorporating these modified guanines and analyzed their hybridization properties.

Experimental section

Synthesis of ASO

ASOs targeting hMALAT1 were designed and synthesized using standard solid-phase phosphonamidite chemistry with the DNA synthesizer nS8-II (GeneDesign, Suita, Japan). After cleavage and deprotection with NH_4OH and purification with Sep-Pak Plus C18 cartridge (Waters), further reversed-phase HPLC purification was conducted using a SHIMADZU LC-6AD system and a Waters XBridgeC18 5 μm 10 × 250 mm column with 8 mM TEA and 0.1 M hexafluoroisopropyl alcohol (HFIP)/MeOH (5% to 45% gradient). The results of MS measurements using MALDI-TOF MS (ltrafleXtreme, Bruker Daltonics) supported the target oligonucleotide (Figures S1-S5). Other ASOs were purchased from Nihon Techno Service (Ushiku, Japan). Complementary RNAs were purchased from Eurofins Genomics (Tokyo, Japan).

UV-Melting experiment

ASO and complementary RNA were dissolved in sodium phosphate buffer (0.1 M NaCl and 0.1 mM EDTA, pH 7). The final concentration of the duplex was 2 μM. The mixture was denatured at 95 °C for 3 min and then cooled to room temperature for annealing. The mixture was placed in quartz cells, cooled to 25 °C, and then heated to 95 °C at a rate of 0.5 °C/min. The absorption at 260 nm was recorded as a UV-melting curve in a JASCOV-730 spectrometer and then smoothed using the Svitzky-Golay method. The measurement was conducted in three independent experiments, and the temperature mean that led to the maximum of the first deviation of each melting curve was used for each T_m value.

RNase H cleavage assay

ASO and complementary 5' FAM-labelled RNA were dissolved in a buffer containing 50 mM Tris-HCl, 75 mM KCl, and 125 μM EDTA (final concentration of ASO and RNA of 100 nM) and annealed at 60 °C for 2 min and cooled to 37 °C. Next, 5 μL of RNase H (0.005 U/μL, *Escherichia coli*, TaKaRa Bio , Otsu, Japan) in 50 mM Tris-HCl, 75 mM KCl, 20 mM $MgCl_2$, and 5 mM DTT was added to each 100 μL of ASO/RNA-containing mixture and incubated at 37 °C for 5 min. Then, 10 μL of the

reaction mixture was collected and added to 10 μL of a solution containing 10 M urea, 50 mM EDTA・2Na, and 0.1 wt% Bromophenol Blue. Finally, 8 μL of samples were applied to 20% denaturing polyacrylamide gel and electrophoresed at 60 °C, 60 W for 30 min. As a control, distilled water was added instead of ASO. In addition, ladders of 13, 11, and 9 mer FAM-RNA (100 nM) that had shortened hRluc and hMALAT1 RNAs from the 5' side were used as size markers. The separated bands were quantified using Typhoon™ FLA, and the ratio of the degradation product was calculated using ImageJ software (ver. 1.53). Minor values thought to be background noise (< 5% of the total RNA in the lane) were ignored.

Reporter assay

HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal inactivated bovine serum (FBS), 100 U/mL penicillin-100 μg/mL streptomycin at 37 °C under 5% CO_2 until sub-confluence. The cells were seeded at 2×10^4 cells/100 μL/well in 96-well plates with DMEM containing 10% FBS. On the following day, the cells were washed twice with Opti-MEM and the medium was replaced with 100 μL/well Opti-MEM. Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA), each ASO, and plasmids pGL4.13 (luc2 / SV40; 0.1 ng/well; Promega, Madison, WI, USA) and pGL4.73 (hRluc / SV40; 0.1 ng/well; Promega) were mixed with Opti-MEM. The mixture was transfected into the cell culture medium and incubated at 37 °C under 5% CO_2. The cells were harvested the next day, washed once with 1× PLB (Dual-Luciferase Reporter Assay System, Promega) at 40 μL/well, and analyzed using the Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's instructions. The luminescence signals of Renilla luciferase were measured by adding Stop & Go Reagent. Finally, luciferase activity was calculated as the ratio of the Renilla luciferase signal to that of firefly (hRluc / luc2) in each well, and further normalization was conducted using the value obtained from the Lipofectamine+/ASO- controls. Each treatment was performed in triplicate.

Target knockdown assay

HeLa cells were grown in DMEM supplemented with 10% FBS, 100 U/mL penicillin-100 μg/mL streptomycin 37 °C under 5% CO_2 until sub-confluence. The cells were seeded at 2×10^4 cells/100 μL/well in 96-well plates with DMEM containing 10% FBS. On the following day, the cells were washed twice with PBS and the medium was replaced with 100 μL/well Opti-MEM. The mixture of MALAT1 ASO, Lipofectamine 2000 (0.5 μL/well; Invitrogen), and Opti-MEM were added to the cell culture and incubated for 24 h at 37 °C under 5% CO_2. The cells were then harvested, and heat denaturation, DNase treatment, and RT-PCR were performed using ReverTra Ace qPCR RT Master Mix with gDNA Reverse (TOYOBO, Osaka, Japan). Finally, qPCR was performed using TaqMan probes. The forward primer was 5'-GCTTGG CTTCTTCTGGACTCA-3' and the reverse primer was 5'-TCGCGAGCTTCACCATGA-3. ' The results were normalized to GAPDH expression levels. Each treatment was performed in triplicate.

Microarray analysis and *in silico* off-target analysis

Microarray analysis was performed using an Agilent SurePrint G3 Human GE v3 8× 60 K system (Agilent, Santa Clara, CA, USA). HeLa cells were incubated with Lipofectamine 2000 (Invitrogen) for 24 h. The intervention consisted of 10 nM of hRluc- and hMALAT1-targeting ASOs. Each treatment was performed in triplicate. The data were processed using R 4.1 and Python 3.7 and pre-processed with quantile normalization following median polish summarization for each gene. P-values were calculated using Limma[31] and corrected using the Benjamini-Hochberg procedure. The latter values were used as adjusted p-values. Differentially expressed genes (DEG) were defined as those with adjusted p-values < 0.05. Downregulated DEG were defined as adjusted p-value < 0.05 and log2 fold-change < –1.0. However, because the gene expression variation was small in the hMALAT1-targeting ASO, another cut-off of log2 fold-change < –0.5 was also used to capture the trends in the differences of the modifications. The number of off-target genes was determined using the GGGenome (https://gggenome.dbcls.jp/). "Human pre-spliced RNA, RefSeq curated on hg38.p12, D3G 21.01 (released Jan 2021)" and "Human spliced RNA, RefSeq curated on hg38.p12, D3G 21.01 (released Jan 2021)" were used as the target sequence of human unspliced or spliced RNAs and extract unique gene symbols. In addition, distance was defined as the number of mismatches + inclusions + deletions analyzed using the GGGenome. Then, potential hybridization-dependent downregulated DEG were defined as adjusted p-value < 0.05 log2fc < –1, distance ≤ 2 for hRluc and adjusted p-value < 0.05 log2fc < –0.5, distance ≤ 2 for hMALAT1.

Comparison of genes downregulated by ASO5 and ASO7

A comparison of the hybridization-dependent off-target effects of **ASO5** (G) and **ASO7** (pyG) focused on the number of potentially hybridization-dependent downregulated DEG. The number of genes was corrected using auxiliary lines drawn such that the knockdown (KD) effects of each on-target hMALAT1 were equally proportional. Genes with absolute fold-changes greater than 0 were counted and are shown at the top left and bottom right of Figure 5A. The mismatch + inclusion + deletion for these genes was analyzed using GGGenome and the results are shown in Tables S2 and S3. In these tables, "=" means the base of the ribonucleotide residue at this position formed a Watson-Crick base pair. "X" means mispairing other than a Watson-Crick base pair, "I" means that the base at this position has no counter base. "D" means that there is no nucleotide residue at this position. Example: If ASO is 5'- CATatgca-[pyG]-ataaTGT-3' and RNA is 3'-GTATAC-CCTATTACA-5', and the notation is ======DX========.

Finally, the position of mismatch dependency of the pyG effect was analyzed by calculating the sum of the "X," "D," and "I" numbers in each of 5' and 3' regions relative to the modification position.

Result and discussion

Design and synthesis of ASOs

We designed and synthesized 16mer ASOs containing deoxyguanosine derivatives, cmG and pyG, with modifications at the 2-amino group. The targets of the ASOs were the mRNA of Renilla luciferase (hRluc), whose codon usage was optimized for expression in human cells, and hMALAT1 RNA, a long non-coding RNA. The sequences and chemical structures of the modifications are shown in Figure 1. Underlined uppercase letters indicate LNA[32] (Figure 1A, Figure 1B), uppercase letters indicate RNA, and lowercase letters indicate DNA. Each gapmer ASO is composed of six LNA residues and a phosphorothioate backbone. Phosphorothioate[33] is a structure commonly used in approved oligonucleotide drugs and is known to improve the *in vivo* stability of nucleic acids. LNA is also known to significantly improve the duplex stability of nucleic acids. While both are promising chemical modifications, the toxicities possibly associated with these modifications have been reported,[33, 34] so modification of the properties of ASOs incorporating these structures are important. The sequence of the ASO targeting the mRNA of hRluc was 5'-CCAgtttcc-x-catGAT-3' where x is deoxyguanosine (**ASO1**), cmG (**ASO2**), pyG (**ASO3**) or deletion (**ASO4**). Similarly, human hMALAT1-targeting ASO was 5'-CATatgca-y-ataaTGT-3' (**ASO5-8**). Single pyG and cmG residues were incorporated into the 10th position from the 5' end of the hRluc ASOs and the 9th position of the hMALAT1 ASOs, respectively. We also prepared RNAs hybridized with ASOs, as shown in Figure 1C. The sequence of **RNA1-4** was AUCAUG-Z-GGAAACUGG, and that of **RNA5-8** was 5'-ACAUUAU-Z-UGCAUAUG -3', where Z is C, U, A, or G (Figure 1C). Previous studies have shown that cmG and pyG adopt two conformations (Figure 1D): a more stable closed type conformation that can form intramolecular hydrogen bonds and an open-type conformation required to form a Watson-Crick base pair with cytosine. Therefore, its base recognition ability may differ from that of unmodified guanine, and the chemical and biological properties of ASO containing cmG or pyG were examined.

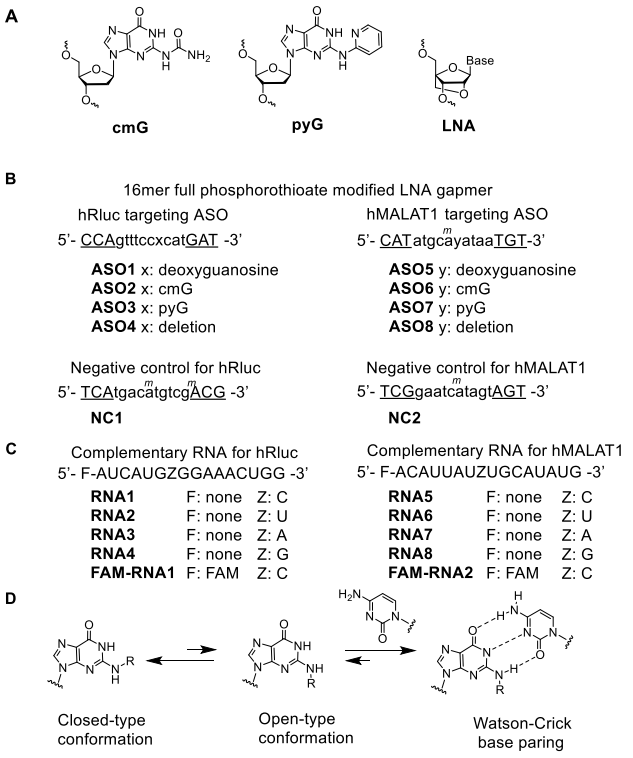


Fig. 1 (A) Chemical structure of base-modified nucleoside cmG, pyG and sugar structure of LNA (B) Sequence of the ASOs used in this study (C) Sequence of RNAs used in this study. Characters used in (B, C) denote the following nucleosides N: LNA, n: DNA, x: deoxyguanosine, deoxyguanosine derivatives or deletion, N: RNA and ^{m}C: 5-methyl-dC. (D) Rotation of substitution at position 2 in cmG and pyG.

Binding affinity of ASOs towards complementary and mismatched RNAs

First, the stability of the fully matched ASO/RNA duplex was evaluated using UV-melting (**RNA1** in Table 1). For the full-match hRluc duplexes (**RNA1**/**ASO1-4**), compared to **ASO1** whose T_m value was 67.5 °C, that of **ASO2** containing a cmG was 68.5 °C. Similarly, the T_m value was reduced to 63.7 °C by the introduction of the bulkier pyG (**ASO3**). We also evaluated the effect of removing guanine at position x as a negative control. As a result, the stability of **ASO4**, which formed no base pairs, significantly decreased to 55.3 °C. Thus, cmG and pyG may form base pairs with opposite C residues. For **ASO5-8** targeting hMALAT1 (**RNA5** in Table 2), the T_m of the duplex with **RNA5** was 51.3, 52.0, 46.8, and 30.5 °C, which were lower than those of the **ASO1-4**/**RNA1** duplex. Next, we compared the duplexes of ASOs and mismatched targets (Tables 1 and 2). When **RNA2**, **3**, **and 4** containing a mismatched base were hybridized with **ASO1**, the T_m values were 58.5, 56.6, and 58.4 °C, respectively. The differences from the perfectly matched duplex **ASO1**/**RNA1** were –9.0, –10.9, and –9.1 °C respectively. Thus, it was confirmed that the G-U mismatch formed in **ASO1**/ **RNA2** was the most stable. For modification with cmG (**ASO2**), the T_m of the duplexes with **RNA2**, **3**, and **4** were 58.6, 56.3, and 59.6 °C, respectively, which were lower than those of the duplex with **RNA3** by –9.9, –12.2, –8.9 °C, respectively. This suggests that modification with cmG is well tolerated in terms of mismatch recognition. Especially, the G-U and G-A mismatch discrimination was slightly improved from –9.0 to –9.9 °C and –10.9 to –12.2 °C, respectively. In contrast, pyG modification (**ASO3**) decreased the base recognition ability, giving only smaller differences in duplex stability compared to the perfectly matched duplex. For example, the T_m value of the duplex of **RNA2** and **ASO3** was 62.2 °C, which was –1.5 °C lower than that of the full-match duplex with **RNA1**. In addition, it was not surprising that the guanine-deleted **ASO4** completely lost its mismatch recognition ability, as shown in the last row of Table 1. Thus, when the target RNA contained a mismatched base in

front of the modified guanine, the mismatch recognition ability remained, but the ΔT_m was smaller in pyG. A similar trend was observed for the mismatch discrimination ability of **ASO5-8**. For example, the difference in T_m value of the **RNA5**/**ASO5** and **RNA6**/**ASO5** duplexes was –12.5 °C, whereas the difference in T_m value of the pyG-modified **RNA6**/**ASO5** and **RNA6**/**ASO7** duplexes was –3.5 °C. Interestingly, the GU mismatch recognition ability was increased by cmG modification compared with guanine, and a similar tendency was observed for both **ASO6**/**RNA6** and **ASO2**/**RNA2**. These results are consistent with those of previous studies on DNA duplexes.[30]

Analyses of RNA fragments cleaved by RNase H

Next, we analyzed the RNA fragments generated by RNase H digestion using gel electrophoresis to reveal the alteration in RNase H preference owing to guanine modification (Figure 2). RNase H is known to show nucleotide sequence preference[35,36], which suggests direct or indirect recognition of the chemical structure of nucleobases. Thus, we expected the cleavage position to change by modifying the x and y positions of the ASOs.

RNase H cleaves the target RNA at multiple sites. Thus, the potential cleavage sites were defined as **a**-**g** in order of proximity to the modification position (Figure 2A and 2B). Based on the X-ray structure of the RNase H catalytic domain[37] (PDB: 2QK9) reported by Nowtony et al., RNase H has a key interaction site with the DNA strand (Figure 2C). The interacting site was named the phosphate-binding pocket formed by Arg179, Thr191, and Asn240, which recognized the 3'-phosphate of a nucleotide residue. The cleaved phosphodiester bond is the 3'-phosphate group of the ribonucleoside residue in the RNA strand that forms a base pair with the 3'-deoxynucleotide residue at the 5'-upstream of the above deoxynucleotide captured in the phosphate-binding pocket. Figure 2D shows the DNA in 2QK9 in more detail. Assuming that this DNA were the modified gapmers in this study, the letters a', b,' c', d', and e' in the figure indicate the position of the modified guanine when RNA was cleaved at sites a, b, c, d, and e, respectively. According to the crystal structure (Figure 2D), we expected that cmG and pyG might not be acceptable at positions a, ' b, ' and c' because the minor groove side of nucleobases is recognized at these positions.

We digested **FAM-RNA1** and **FAM-RNA2** with RNase H in the presence of **ASO1-4** and **ASO5-8**, respectively (Figure 2E and 2F). To observe the initial degradation, the reaction was stopped after 5 min and the products were quantified as band density to To estimate the length of the RNA fragments, 5' FAM-RNA shortened to 13, 11, and 9 mers, which corresponded to the fragments cleaved at positions **g**, **e**, and **c** of hRluc and positions **f**, **d**, and **b** of hMALAT1, were applied to lanes 1 and 7 as size markers. Intact **FAM-RNA1** and **FAM-RNA2** were applied in lanes 2 and 8, respectively, and the digested products of **FAM-RNA1** and **FAM-RNA2** were applied in lanes 3-6, and 9-12, respectively. The products in lanes 3-6 were those obtained in the presence of **ASO1-4**, and those in lanes 9-12 were obtained in the presence of **ASO5-8**. The panels below the gel images represent heat maps of the cleavage patterns. In lane 3, 41% of full-length RNA remained, and 48% of the 13mer RNA was generated by cleavage at position **g**. In addition, 12% of the fragments cleaved at position **f** was present in between the 13mer and 11mer bands. Many minor bands with intensities less than 5% were also detected but not quantified. Thus, **FAM-RNA1**

Table 1 Thermal stability of hRluc duplexes with guanine modifications

hRluc-targeting ASO
ASO: 5′- CCAgtttccxcatGAT -3′
RNA: 3′- GGUCAAAGGZGUACUA -5′

	T_m (°C) ($T_{m\ mismatch} - T_{m\ perfect\ match}$ (°C))			
	RNA1 (Z = C)	**RNA2** (Z = U)	**RNA3** (Z = A)	**RNA4** (Z = G)
ASO1 (x = G)	67.5	58.5 (–9.0)	56.6 (–10.9)	58.4 (–9.1)
ASO2 (x = cmG)	68.5	58.6 (–9.9)	56.3 (–12.2)	59.6 (–8.9)
ASO3 (x = pyG)	63.7	62.2 (–1.5)	61.0 (–2.7)	61.2 (–2.5)
ASO4 (x = deletion)	55.3	55.0 (–0.3)	55.0 (–0.3)	57.7 (2.4)

Column of Z = C, thermal stability of perfectly matched duplex of hRluc-targeting ASO with complementary RNA; Z = U, A and G: target RNA containing a mismatched base in front of the guanine or guanine modification. Mismatch discrimination was calculated as differential thermal stability of duplexes containing a mismatched base pair in front of guanine modification compared to that of perfectly matched duplexes. The measurements were repeated at least three times and the standard deviations were < 1.0.

Table 2 Thermal stability of hMALAT1 duplexes with guanine modifications

hMALAT1-targeting ASO
ASO: 5′- CATatgcayataaTGT-3′
RNA: 3′- GUAUACGUZUAUUACA-5′

	T_m (°C) ($T_{m\ mismatch} - T_{m\ perfect\ match}$ (°C))			
	RNA5 (Z = C)	**RNA6** (Z = U)	**RNA7** (Z = A)	**RNA8** (Z = G)
ASO5 (y = G)	51.3	38.8 (–12.5)	38.5 (–12.8)	39.5 (–11.8)
ASO6 (y = cmG)	52.0	39.1 (–12.9)	39.1 (–12.9)	41.9 (–10.1)
ASO7 (y = pyG)	46.8	43.3 (–3.5)	43.0 (–3.8)	42.6 (–4.2)
ASO8 (y = deletion)	30.5	33.4 (2.9)	38.1 (7.6)	36.0 (5.5)

Column of Z = C: thermal stability of perfectly matched duplex of hMALAT1-targeting ASO with complementary RNA; Z = U, A and G: target RNA containing a mismatched base in front of the guanine or guanine modification. Mismatch discrimination was calculated as differential thermal stability of duplexes containing a mismatched base pair in front of the guanine modification compared to that of perfectly matched duplexes. The measurements were repeated at least three times and the standard error was < 1.0.

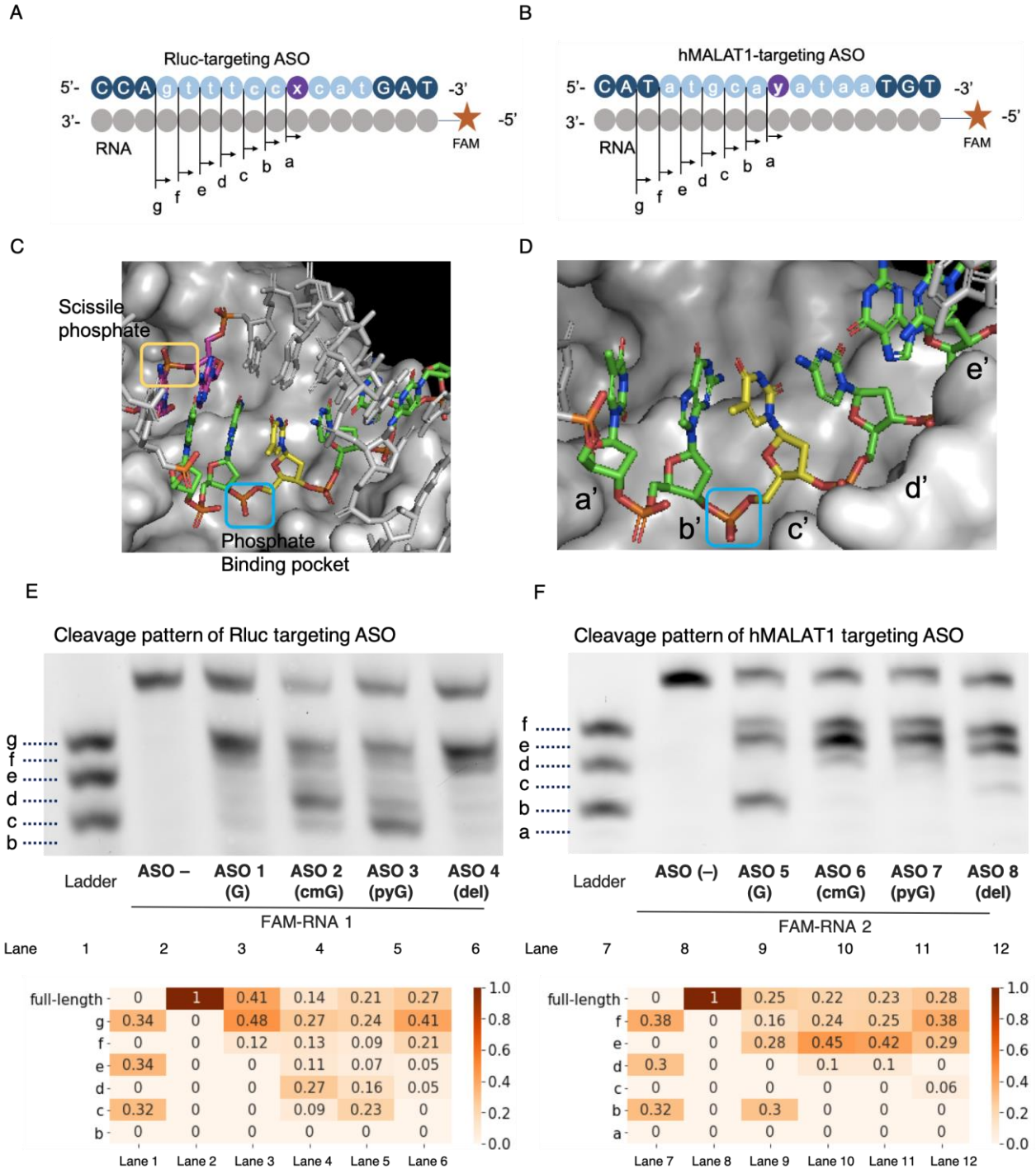


Fig. 2 RNase H footprinting and its quantification. (A, B) Possible cleavage site and positions in duplexes between complementary RNA and A: hRluc B: hMALAT1 ASO. (C) Crystal structure of RNase H and DNA/RNA duplex. (D) Image of (C) in greater detail and focusing on the modification position of the ASO strand and RNase H. (E, F) RNA cleavage pattern digested with RNase H and the band quantification treatment of E: hRluc-targeting ASO F: hMALAT1-targeting ASO.

was mainly cleaved at position **g** after treatment with **ASO1** and RNase H. Similarly, in lane 4 (x: cmG), 14% of full-length RNA remained, and the fragments cleaved at positions **g**: 27%, **f**: 13%, **e**: 11%, **d**: 27%, and **c**: 9% were observed. In lane 5 (x: pyG), 21% of full-length RNA remained, and products cleaved at positions **g**: 24%, **f**: 9%, **e**: 7%, **d**: 16%, and **c**: 23% were observed. In lane 6 (x: del), 27% of full-length RNA remained, with products cleaved at positions **g**: 41%, **f**: 21%, **e**: 5%, and **d**: 5%. Thus, it was observed that the remaining full-length RNA decreased cmG (14%), pyG (21%), and del (27%) modifications compared to unmodified G (41%), as shown in lanes 4, 5, and 6, respectively.

On the other hand, for hMALAT1-targeting ASOs, 25% of full-length RNA remained and was cleaved at positions **f**: 16%, **e**: 28%, and **b**: 30% in lane 9 (y: G); 22% of full-length RNA remained and was cleaved at positions **f**: 24%, **e**: 45% and **d**: 10% in lane 10 (y: cmG); 23% of full-length RNA remained and was cleaved at positions **f**: 25%, **e**: 42%, and **d**: 10% in lane 11 (y: pyG); and 28% of full-length RNA remained and was cleaved at positions **f**: 38%, **e**: 29%, **c**: 6%, in lane 12 (y: del). Thus, it was observed that the remaining full-length RNA was comparable among cmG (22%), pyG (23%), and del (28%), when compared to unmodified G (25%), as shown in lanes 10, 11, 12, and 9, respectively. Hence, the effect of the modification on the overall cleavage activity sequence-dependent.

In addition to the overall cleavage activity, the cleavage patterns were also altered. For hMALAT1-targeting ASOs (Figure 2F), the cleavage product **b** disappeared in both modified ASOs. As mentioned above, the cleavage product **b** corresponds to the modification at position **b'**, which is expected to be inhibited owing to steric hindrance between the substituents and RNase H. Surprisingly, for the hRluc-targeting ASO (Figure 2E), the cleavage product **c** increased in the modified ASOs compared to the unmodified ASO. The cleavage product **c** corresponds to the modification at position **c'** and was expected to be inhibited by steric hindrance. One possible reason is phenylalanine F213, which is the closest amino acid around position **c'**. The substituents may be attracted to phenylalanine residue interactions, allowing the cleavage of product **c** of **FAM-RNA1**. We also studied the effects of multiple modification on the target cleavage pattern (Figure S10). For this purpose, we synthized hMALAT1 targetting ASO incorporating two pyG (**ASO9**: 5'-CATaty^{m}cayataaTGT-3', N: LNA, y: pyG). As shown in Figure S10, the two-points modification resulted in selective cleavage of the **FAM-RNA2** at position f. However, more than half of the full-length **FAM-RNA2** remained uncleaved. Thus, as long as the hMALAT-1 targeting ASO used in this study were concerned, there was a trade-off between activity and selectivity.

In vitro antisense activities

Next, we evaluated the on-target knockdown (KD) activity (Figure 3). The KD assay of hRluc ASO was performed by evaluating the luminescence of the reporter gene, and those of hMALAT1 ASOs were evaluated by qRT-PCR. Each ASO was transfected into HeLa cells. These results show that dose-dependent KD was observed in all cases. For example, for **ASO1** at 0.5 nM, the expression of hRluc was 56% of the control (ASO-), and the expression decreased as the concentration of **ASO1** increased to 2.5 to 10 nM. Similarly, for **ASO2** and **ASO3** at 0.5 nM, the expression of hRluc was 65% and 78%, respectively, and the expression decreased as the concentration of **ASO1** increased from 2.5 to 10 nM. **ASO4** hRluc did not decrease at 0.5 nM, but downregulation was detected over 2.5 nM in a dose-dependent manner. In addition, in **ASO5, ASO6,** and **ASO7**, hMALAT1 expression decreased as the concentration of each ASO increased from 1 nM to 2.5 nM. Compared to 1 nM, where hMALAT1 KD was clearly observed, treatment with **ASO5, ASO6,** and **ASO7** decreased the expression of hMALAT1 by 26%, 55%, and 47%, respectively. Therefore, the targets hRluc and hMALAT1 were significantly downregulated in each ASO except for the guanine deletion ASO, and modification of pyG and cmG maintained KD activity in each case. It is possible that the carbamoyl group was cleaved in the cell. However, the results of the microarray analysis described below show that global transcript alterations differed between cmG and G, suggesting that the carbamoyl group remained undecomposed. In addition, although the RNase H digestion assay (Figure 2) showed similar amounts of undigested RNA with G, cmG, pyG, and deletion, the ASO with deletion significantly lost KD activity. The cause of this difference is the difference in the concentrations of ASO and RNase H between the cell-free system and *in vitro* experiments.

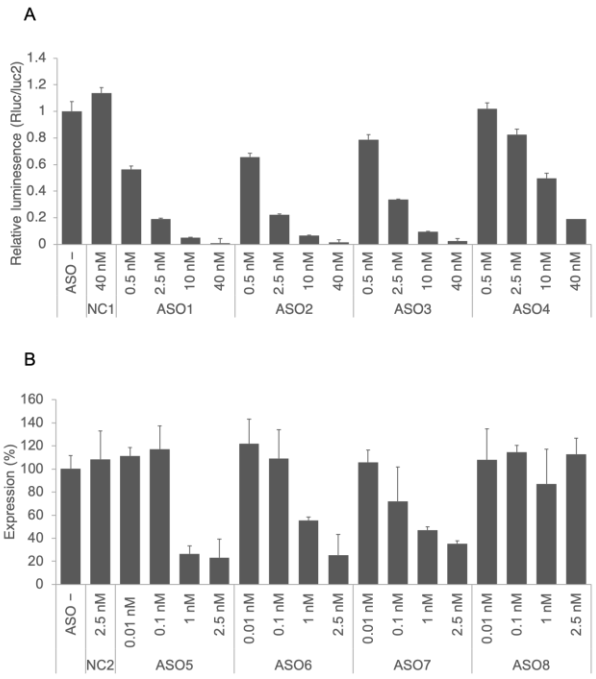


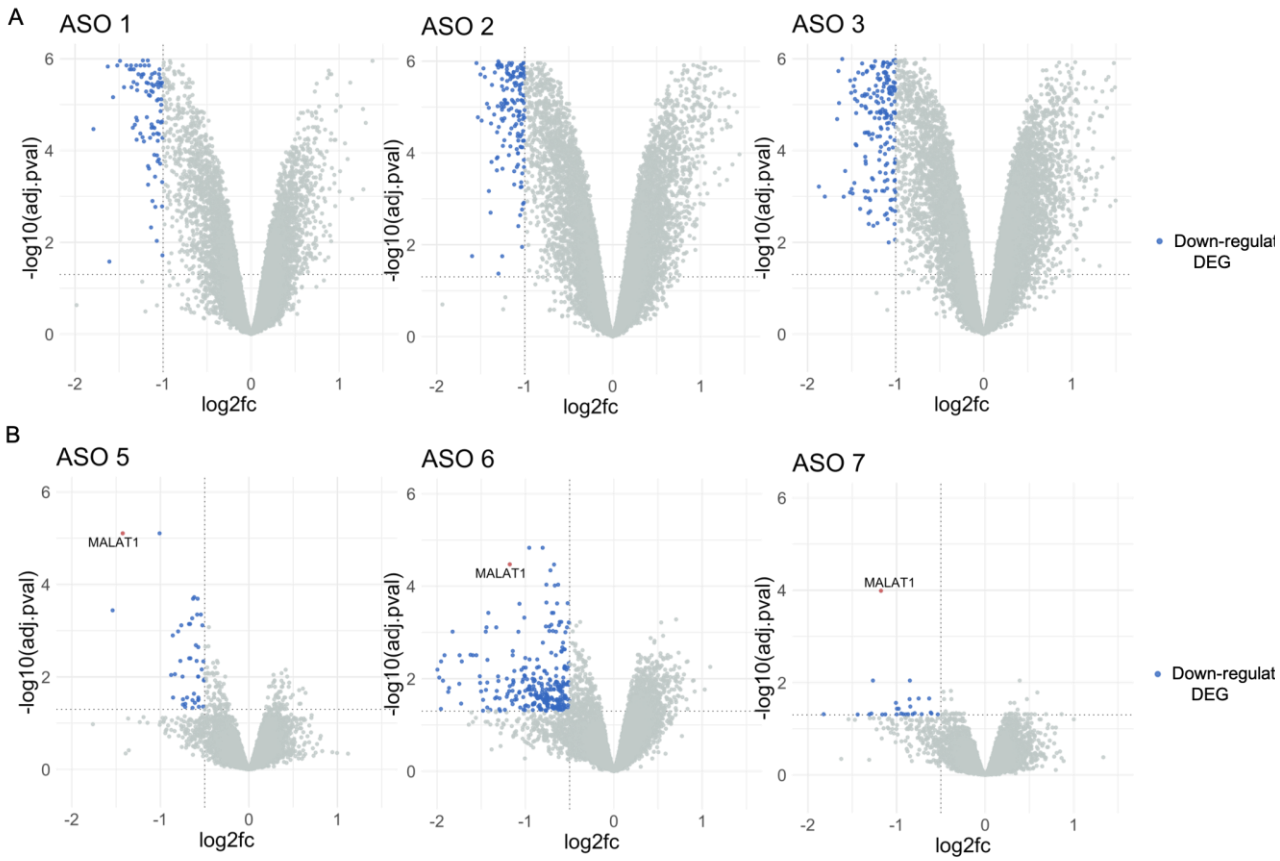
Fig. 3 Effects of modification of guanine analogues on antisense activity. (A) Relative luminescence of hRluc when **ASO1-4** was transfected into HeLa cells. (B) Quantification of endogenous hMALAT1 expression after **ASO5-8** transfection into HeLa cells. Experiments were performed independently three times and mean values are shown.

Global transcript alterations analyzed with DNA microarray

To study global transcript alterations, we treated HeLa cells with 10 nM ASOs, conducted DNA microarray experiments, and analyzed the differences in the presence or absence of ASOs (see experimental section **2.6**) (Figure 4). Data were visualized using volcano plots (Figure 4A and 4B). In addition, we summarized the data as shown in Figure 4C and 4D based on the adjusted p-value (see experimental section), log2 fold-change, and the 'distance,' which we defined as the number of mismatch +

inclusion + deletion counted for the pairs of each ASO and the target gene.

For the hMALAT1-targeting **ASO5-7** (Figure 4D), we used another threshold of adjusted p-value < 0.05, log2 fold-change <



C

	Number of genes		
	ASO 1 G	ASO 2 cmG	ASO 3 pyG
Downregulated DEG (Adj.pvalue < 0.05 & log2fc < –1)	98	181	266
Potential hybridization dependent downregulated DEG (Adj.pvalue < 0.05 & log2fc < –1 & distance ≦ 2)	33	68	102

D

	Number of genes		
	ASO 5 G	ASO 6 cmG	ASO 7 pyG
Downregulated DEG (Adj.pvalue < 0.05 & log2fc < – 0.5)	45	285	27
Potential hybridization dependent downregulated DEG (Adj.pvalue < 0.05 & log2fc < – 0.5 & distance ≦ 2)	26	46	4

Fig. 4 Global transcript alteration detected by microarray. (A, B) Volcano plots A; hRluc-targeting ASO, B; hMALAT1-targeting ASO. (C, D) Number of variable genes defined by each threshold; downregulated differentially expressed genes (downregulated DEG) and potential hybridization-dependent downregulated DEG; hRluc-targeting ASO: downregulated DEG: adjusted p-value < 0.05 and log2fc < -1 / potential hybridization-dependent downregulated DEG: adjusted p-value < 0.05, log2fc < -1 and distance ≤ 2. hMALAT1-targeting ASO; downregulated DEG: adjusted p-value < 0.05 and log2fc < -0.5 / potential hybridization-dependent downregulated DEG: adjusted p-value < 0.05, log2fc < -0.5 and distance ≤ 2.

First, we analyzed the data for **ASO1-3**. We defined downregulated DEG by applying the criteria adjusted p-value < 0.05 and log2 fold-change < –1.0, which corresponds to more than 50% downregulation and is shown in the first row of Figure 4C. These genes are shown as blue dots in Figure 4A. Among these downregulated DEG, the potential hybridization-dependent downregulated DEG was also defined as a threshold of adjusted p-value <0.05, log2fc <–1.0, and distance ≤ 2 in the second row of Figure 4C.

–0.5, for downregulated DEG as shown in the first row of Figure 4D because, if log2 fold-change < –1.0 were used, the number of the genes classified in this class was less than 10 for **ASO5** and **ASO7**. These numbers are too small for comparison purposes. Accordingly, we also used another cut-off of log2 fold-change < –0.5 for the classification of potentially hybridization-dependent downregulated DEG listed in the second row of Figure 4D. In addition, the sequence based on the *in silico* analysis confirmed that the hRluc-targeting ASO did not have any perfectly matched

human genes, and had a smaller number of mismatched genes (distance ≤ 2; 616 genes (Table S1)) than that of the hMALAT1-targeting ASO (distance ≤ 2; 2476 genes).

Before comparing these data, the on-target KD activities of **ASO5-7** were evaluated in microarray experiments. The target gene, hMALAT1, is indicated by red points in Figure 4B. As a result, these ASOs suppressed the target mRNA levels to 37% (**ASO5**), 44% (**ASO6**), and 44% (**ASO7**) (Figure S11). The same analyses could not be performed for hRluc because the hRluc ASO did not have an on-target human gene, as described above.

Next, we compared the number of downregulated DEG by each chemical modification. For the hRluc-targeting ASO, the number of downregulated DEG was 98 for **ASO1** (G), 181 for **ASO2** (cmG), and 266 for **ASO3** (pyG). Thus, the number of downregulated DEG in hRluc cells increased by cmG and pyG modifications. In contrast, the number of downregulated DEG in hMALAT1 was 45 for **ASO5** (G), 285 for **ASO6** (cmG), and 27 for **ASO7** (pyG). In these cases, the number increased by cmG, but decreased by pyG. The results of the potential hybridization-dependent downregulated DEG showed the same alteration tendency as the downregulated DEG mentioned above. Notably, the downregulated DEG of **ASO6** (Figure 4D), which was 285, was much larger than that of **ASO5** and **ASO7**. We assumed that **ASO6** might contain many downregulated genes independent of hybridization.

In the hRluc experiments (Figure 4C), the number of potential hybridization-dependent downregulated DEG for **ASO2** (cmG) and **ASO3** (pyG), which were 68 and 102, respectively, was larger than 33 for **ASO1** (G). In hMALAT1 (Figure 4D), the number of potential hybridization-dependent downregulated DEG in **ASO6** (cmG) which was 46, was larger than that in **ASO5** (G) (26). In contrast, that of **ASO7** (4) was smaller than that of **ASO5**. The results of **ASO2** and **ASO6** indicated that ASO containing cmG induced more global gene expression alterations independent of the ASO sequence. In contrast, the results for **ASO3** and **ASO7** showed that the effect of pyG was sequence-dependent. Therefore, we focused on the pyG modification for further analysis.

Comparison of log2 fold-changes of genes downregulated by ASO5 and ASO7

To investigate in more detail the hybridization-dependent downregulation by pyG modification, we compared the fold-change of the genes extracted using the criteria log2 fold-change < 0, adjusted p-value < 0.05, distance ≤ 2, between **ASO5** and **ASO7** (Figure 5A and 5B). In this comparison, to consider the different on-target KD activities of **ASO5** and **ASO7**, log2 fold-changes were plotted with the line shown in Figure 5B, which connects the point of hMALAT1 to the origin. In this plot, the genes below this line were considered to be less downregulated by **ASO7** (pyG) than **ASO5** (G). We analyzed 65 genes that were downregulated by **ASO5** or **ASO7** (Figure 5A). Among them, 50 showed greater downregulation in **ASO5,** 8 showed greater downregulation in **ASO7**, and 7 genes remained almost unchanged (log2 difference < 0.05). Even when considering on-target effects, ASO with pyG suppressed the downregulation of off-target genes.

The sequences of the 50 and 8 genes most downregulated in

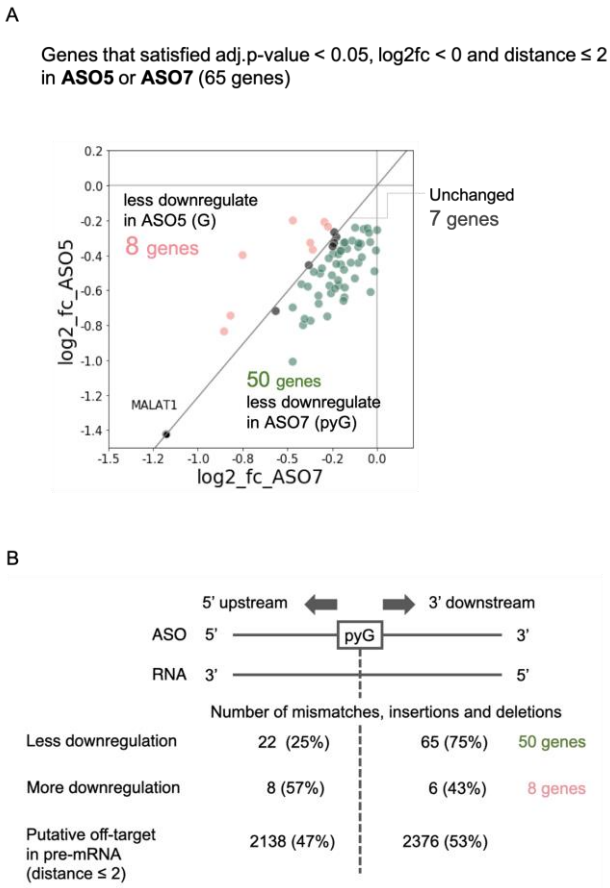


Fig. 5 Fold-change comparison of all hybridization-dependent downregulated genes (adjusted p-value < 0.05, log2fc < 0, distance ≤ 2) by ASO5 (G) and ASO7 (pyG) for each gene. (A) A plot of fold-change of ASO5 and ASO7. The line passing the point of MALAT1 and the origin is drawn to compensate the on-target efficacies of both ASO. (B) Positions of mismatches relative to pyG. Total number of mispairing, deletion, and inclusion are separately displayed for the 5' upstream and 3' downstream regions of pyG.

ASO5 and **ASO7**, respectively, were extracted and each nucleotide in these genes was classified in three types of mismatches: mispairing ("X"), inclusion of a base in RNA ("I"), or deletion of a base in RNA ("D") (see experimental part **2.7**) (Tables S2 and S3). When the sequences of these 50 genes were aligned, a trend was observed in which they formed mismatches with **ASO7** in the region 3' downstream to pyG (Table S3). This trend was not observed in the 8 genes (Table S2). It should be noted that putative off-targets in pre-mRNAs whose distance ≤ 2 extracted from "Human pre-spliced RNA, RefSeq curated on hg38.p12, D3G 21.01 (released Jan 2021)" did not show such a trend either. To quantify the trend, the number of "X," "I," or "D" in the 3' downstream and the 5' upstream regions of pyG were summarized (Figure 5B). In the 50 genes, there were 65 mismatches in the 3' downstream region of pyG, whereas 22 were observed in the 5' upstream region. Thus, the genes having mismatches in the 3' downstream region of pyG tended to be weakly downregulated. In contrast, in the 8 more downregulated

genes and the pre-mRNA (distance ≤ 2), the mismatches in the 3' downstream of pyG were 6 and 2376, and in the 5' upstream were 8 and 2138, respectively. From a structural point of view, the 2-amino group of guanine was stacked in the 3' downstream nucleobase, suggesting that mismatches may be more structurally distorted by the pyG moiety and resulting in suppression of downregulation.

Conclusions

In conclusion, it was shown that modified guanines incorporated at position 2, cmG, and pyG, had different physicochemical and biological properties from those of guanine. As shown in Tables 1 and 2, cmG increased the GU base recognition ability in duplex melting experiments. In addition, as shown in Figure 2, cmG and pyG altered RNase H cleavage patterns. Finally, pyG modification suppressed the off-target effects of the hMALAT1-targeting ASO. In the development of antisense oligonucleotide therapy, ASO sequences are designed such that the number of off-target genes predicted *in silico* is minimized to maximize safety.[13,38] Despite these efforts, safety problems may arise during the final stages of development. In such cases, it would be useful to alter off-target profiles using modified nucleic acid analogs. For such applications, pyG is likely to be more useful than cmG as far as the present data are concerned. The reduction in the number of off-target genes by pyG was sequence-dependent; therefore, it cannot be universally applied to all ASOs. However, there was a correlation between the cleavage pattern of the gel and the microarray results. For example, for hRluc, the selectivity of RNase H cleavage decreased and off-target gene variation was increased by the modification. In contrast, the selectivity increased and the off-target variation decreased in the hMALAT1-targeting ASO. Thus, a prior experiment on RNase H cleavage patterns, such as that shown in Figure 2, may help predict the sequence dependence resulting from the modification. To further evaluate the properties of ASOs incorporating pyG, the T_m value with various off-target RNAs, 3-dimensional structure of the ASO-target duplexes, *in vivo* activity, and toxicity should be investigated in the future. In addition, the interaction between ASO incorporating pyG and various intracellular and extracellular proteins should also be clarified, together with the *in vivo* side effects and ADME.[39] Nevertheless, because we revealed the potential of modifying position 2 of guanine for the suppression of the hybridization-dependent off-target effect, it is promising for the future development of ASO.

Data availability

The data supporting the findings of this study are available within the article and in the Supplementary Information.† The microarray data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo/. The accession number is GSE228747 (Rluc: GSE228741 and hMALAT1: GSE228740)

Author Contributions

Conceptualization, T. K, Y. M and K. S.; methodology, T. K, A. Y, Y. M and K. S.; investigation, T. K and A. Y; data curation T. K; writing–original draft, T. K.; writing–review & editing, A. Y, Y. M and K. S.; funding acquisition, Y. M and K. S.; resources, Y. M and K. S

Conflicts of interest

Y. M. and K. S. are unpaid members of Fastide, Inc

Acknowledgments

This study was supported by the Program for Building Regional Innovation Ecosystems of the Ministry of Education, Culture, Sports, Science and Technology (MEXT). This study was also partly supported by KAKENHI (20K21245 and 20H02857) from the Japan Society for the Promotion of Science and the Research Program on Hepatitis (21fk0210089h0001) from the Japan Academy for Medical Research and Development. This work was also partly supported by JST PRESTO "Genome programming" project (grant No. JPMJPR19K9), Japan.

References

1 B. P. Monia, E. A. Lesnik, C. Gonzalez, W. F. Lima, D. McGee, C. J. Guinosso, A. M. Kawasaki, P. D. Cook and S. M. Freier, *J. Biol. Chem.*, 1993, **268**, 14514–14522.
2 K. Dhuri, C. Bechtold, E. Quijano, H. Pham, A. Gupta, A. Vikram and R. Bahal, *Biochem. Pharmacol.*, 2021, **9**, 2004.
3 S. T. Crooke, X. hai Liang, R. M. Crooke, B. F. Baker and R. S. Geary, *Biochem. Pharmacol.*, 2021, 189.
4 K. S. Frazier, *Toxicol Pathol*, 2015, **43**, 78–89.
5 F. Alhamadani, K. Zhang, R. Parikh, H. Wu, T. P. Rasmussen, R. Bahal, X. Zhong and J. E. Manautou, *Drug Metab. Dispos.*, 2022, **51**, 1–45.
6 X. Chi, P. Gatti and T. Papoian, *Drug Discov. Today*, 2017, 22, 823–833.
7 W. Shen, C. L. de Hoyos, M. T. Migawa, T. A. Vickers, H. Sun, A. Low, T. A. Bell, M. Rahdar, S. Mukhopadhyay, C. E. Hart, M. Bell, S. Riney, S. F. Murray, S. Greenlee, R. M. Crooke, X. hai Liang, P. P. Seth and S. T. Crooke, *Nat. Biotechnol*,, 2019, **37**, 640–650.
8 L. Zhang, T. A. Vickers, H. Sun, X. H. Liang and S. T. Crooke, *Nucleic Acids Res.*, 2021, **49**, 2721–2739.
9 T. Kasuya, S. I. Hori, A. Watanabe, M. Nakajima, Y. Gahara, M. Rokushima, T. Yanagimoto and A. Kugimiya, *Sci. Rep.*, 2016, **6**, 1–12.
10 S. A. Burel, C. E. Hart, P. Cauntay, J. Hsiao, T. Machemer, M. Katz, A. Watt, H.-H. Bui, H. Younis, M. Sabripour, S. M. Freier, G. Hung, A. Dan, T. P. Prakash, P. P. Seth, E. E. Swayze, C. F. Bennett, S. T. Crooke and S. P. Henry, *Nucleic Acids Res.*, 2015, **44**, 2093–2109.
11 T. Yoshida, Y. Naito, H. Yasuhara, K. Sasaki, H. Kawaji, J. Kawai, M. Naito, H. Okuda, S. Obika and T. Inoue, *Genes to Cells*, 2019, **24**, 827–835.
12 P. H. Hagedorn, M. Pontoppidan, T. S. Bisgaard, M. Berrera, A. Dieckmann, M. Ebeling, M. R. Møller, H. Hudlebusch, M. L. Jensen, H. F. Hansen, T. Koch and M. Lindow, *Nucleic Acids Res.*, 2018, **46**, 5366–5380.
13 P. J. Kamola, J. D. A. Kitson, G. Turner, K. Maratou, S. Eriksson, A. Panjwani, L. C. Warnock, G. A. D. Guilloux, K. Moores, E. L. Koppe, W. E. Wixted, P. A. Wilson, N. J. Gooderham, T. W. Gant, K. L. Clark, S. A. Hughes, M. R.

Edbrooke and J. D. Parry, *Nucleic Acids Res.*, 2015, **43**, 8638–8650.
14 M. E. Østergaard, J. Nichols, T. A. Dwight, W. Lima, M. E. Jung, E. E. Swayze and P. P. Seth, *Mol. Ther. Nucleic Acids*, 2017, **7**, 20–30.
15 D. Honcharenko, J. Barman, O. P. Varghese and J. Chattopadhyaya, *Biochemistry*, 2007, **46**, 5635–5646.
16 M. B. Danielsen, C. Lou, J. Lisowiec-Wachnicka, A. Pasternak, P. T. Jørgensen and J. Wengel, *Chem. Eur. J.*, 2020, **26**, 1368–1379.
17 M. E. Østergaard, C. L. de Hoyos, W. B. Wan, W. Shen, A. Low, A. Berdeja, G. Vasquez, S. Murray, M. T. Migawa, X.-H. Liang, E. E. Swayze, S. T. Crooke and P. P. Seth, *Nucleic Acids Res.*, 2020, **48**, 1691–1700.
18 Y. Masaki, A. Tabira, S. Hattori, S. Wakatsuki and K. Seio, *Org. Biomol. Chem.*, 2022, **20**, 8917–8924.
19 B. A. Anderson, G. C. Freestone, A. Low, C. L. De-Hoyos, W. J. Drury, M. E. Østergaard, M. T. Migawa, M. Fazio, W. B. Wan, A. Berdeja, E. Scandalis, S. A. Burel, T. A. Vickers, S. T. Crooke, E. E. Swayze, X. Liang and P. P. Seth, *Nucleic Acids Res.*, 2021, **49**, 9026–9041.
20 L. Zhang, X. H. Liang, C. L. de Hoyos, M. Migawa, J. G. Nichols, G. Freestone, J. Tian, P. P. Seth and S. T. Crooke, *Nucleic Acid Ther.*, 2022, **32**, 401–411.
21 M. E. Østergaard, P. Kumar, J. Nichols, A. Watt, P. K. Sharma, P. Nielsen and P. P. Seth, *Nucleic Acid Ther.*, 2015, **25**, 266–274.
22 A. Granzhan, N. Kotera and M. P. Teulade-Fichou, *Chem. Soc. Rev.*, 2014, **43**, 3630–3665.
23 J. E. Jackman and J. D. Alfonzo, *WIREs RNA*, 2013, **4**, 35–48.
24 H. Zang, Q. Fang, A. E. Pegg and F. P. Guengerich, *J. Biol. Chem.*, 2005, **280**, 30873–30881.
25 A. Mazurek, M. Berardini and R. Fishel, *J. Biol. Chem.*, 2002, **277**, 8260–8266.
26 C. J. Lech, J. K. Cheow Lim, J. M. Wen Lim, S. Amrane, B. Heddi and A. T. Phan, *Biophys. J.*, 2011, **101**, 1987–1998.
27 T. Inde, Y. Masaki, A. Maruyama, Y. Ito, N. Makio, Y. Miyatake, T. Tomori, M. Sekine and K. Seio, *Org. Biomol. Chem*,, 2017, **15**, 8371–8383.
28 E. R. Kandimalla, D. Yu, Q. Zhao and S. Agrawal, *Bioorg. Med. Chem.*, 2001, **9**, 807–813.
29 T. Yoshida, K. Morihiro, Y. Naito, A. Mikami, Y. Kasahara, T. Inoue and S. Obika, *Nucleic Acids Res.*, 2022, **50**, 7224–7234.
30 T. Sasami, Y. Odawara, A. Ohkubo, M. Sekine and K. Seio, *Tetrahedron Lett.*, 2007, **48**, 5325–5329.
31 M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi and G. K. Smyth, *Nucleic Acids Res.*, 2015, **43**, 1–13.
32 M. Petersen and J. Wengel, *Trends Biotechnol.*, 2003, **21**, 74–81.
33 S. T. Crooke, T. A. Vickers and X. H. Liang, *Nucleic Acids Res*, 2021, 48, 5235–5253.
34 P. H. Hagedorn, R. Persson, E. D. Funder, N. Albæk, S. L. Diemer, D. J. Hansen, M. R. Møller, N. Papargyri, H. Christiansen, B. R. Hansen, H. F. Hansen, M. A. Jensen and T. Koch, *Drug Discov Today*, 2018, 23, 101–114.
35 Ł. J. Kiełpiński, P. H. Hagedorn, M. Lindow and J. Vinther, *Nucleic Acids Res.*, 2017, **45**, 12932–12944.
36 O. V. Matveeva, A. D. Tsodikov, M. Giddings, S. M. Freier, J. R. Wyatt, A. N. Spiridonov, S. A. Shabalina, R. F. Gesteland and J. F. Atkins, *Nucleic Acids Res.*, 2000, **28**, 2862–2865.
37 M. Nowotny, S. A. Gaidamakov, R. Ghirlando, S. M. Cerritelli, R. J. Crouch and W. Yang, *Mol. Cell*, 2007, **28**, 264–276.
38 M. Lindow, H.-P. Vornlocher, D. Riley, D. J. Kornbrust, J. Burchard, L. O. Whiteley, J. Kamens, J. D. Thompson, S. Nochur, H. Younis, S. Bartz, J. Parry, N. Ferrari, S. P. Henry and A. A. Levin, *Nat. Biotechnol.*, 2012, **30**, 920–923.
39 S. T. Crooke, X. H. Liang, B. F. Baker and R. M. Crooke, *J. Biol. Chem.*, 2021, 296.