A general strategy exploiting m^5C duplex-remodelling effect for selective detection of RNA and DNA m^5C methyltransferase activity in cells

Tianming Yang, Joanne J.A. Low and Esther C.Y. Woon

Department of Pharmacy, National University of Singapore, 18 Science Drive 4, 117543 Singapore

Received July 20, 2019; Revised October 07, 2019; Editorial Decision October 21, 2019; Accepted October 30, 2019

ABSTRACT

RNA:5-methylcytosine (m^5C) methyltransferases are currently the focus of intense research following a series of high-profile reports documenting their physiological links to several diseases. However, no methods exist which permit the specific analysis of RNA:m^5C methyltransferases in cells. Herein, we described how a combination of biophysical studies led us to identify distinct duplex-remodelling effects of m^5C on RNA and DNA duplexes. Specifically, m^5C induces a C3'-endo to C2'-endo sugar-pucker switch in CpG RNA duplex but triggers a B-to-Z transformation in CpG DNA duplex. Inspired by these different ‘structural signatures’, we developed a m^5C-sensitive probe which fluoresces spontaneously in response to m^5C-induced sugar-pucker switch, hence useful for sensing RNA:m^5C methyltransferase activity. Through the use of this probe, we achieved real-time imaging and flow cytometry analysis of NOP2/Sun RNA methyltransferase 2 (NSUN2) activity in HeLa cells. We further applied the probe to the cell-based screening of NSUN2 inhibitors. The developed strategy could also be adapted for the detection of DNA:m^5C methyltransferases. This was demonstrated by the development of DNA m^5C-probe which permits the screening of DNA methyltransferase 3A inhibitors. To our knowledge, this study represents not only the first examples of m^5C-responsive probes, but also a new strategy for discriminating RNA and DNA m^5C methyltransferase activity in cells.

INTRODUCTION

The DNA and RNA of all living organisms, as well as that of viruses, mitochondria and chloroplasts, undergo a wide range of modifications (1–2). These modifications not only expand the structural diversity of nucleic acids, but also provide an epigenetic mechanism to fine-tune their biological functions (3–4). To date, at least 160 naturally-occurring chemical modifications have been identified, amongst which 5-methylcytosine (m^5C) is currently one of the most intensively studied epigenetic modifications. The m^5C modification is prevalent in DNA and multiple RNA classes (5–6). Its biological functions are best understood in DNA, where it is involved in the regulation of gene expression, genome reprogramming, organismal development and cellular differentiation (5,7,8). In contrast, there has been comparatively less studies on the biological roles of m^5C in RNA. Indeed the significance of m^5C modification in mRNA was not fully appreciated until recently, following the landmark discovery of widespread m^5C sites in the transcriptomes of diverse organisms (9–16), suggesting that the m^5C modification is far more pervasive in human mRNA than previously realised. This has reignited intense interest in the study of this epitranscriptomic mark. At present, the exact biological function of m^5C modification in mRNA remains elusive although it has been linked to many cellular processes, such as nuclear export regulation (14), modulation of protein translation (17–19), and stress response (20).

The m^5C methylation landscape is regulated by a complex array of m^5C methyltransferases (MTases) and m^5C demethylases, which specifically add and remove a 5-methyl group from cytosine base, respectively (6,21–23). In humans, C-5 cytosine methylation of DNA is catalysed by at least three DNA:m^5C MTases (i.e. DNMT1, 3A and 3B), whereas m^5C methylation of RNA is catalysed by the NOL1/NOP2/Sun domain (NSUN) RNA methyltransferase family, which includes NSUN1-7, as well as the DNA MTase homologue TRDMT1 (formerly DNMT2). Notably, both the DNA and RNA m^5C MTases employ a common S-adenosyl-L-methionine (SAM) dependent mechanism for methyl group transfer, although different active-site cysteine residues are likely involved in the catalysis (24).

In recent years, it has become increasingly clear that aberrant expression of m^5C MTases may underlie the pathogenesis of several diseases. For instance, DNMT3A, NSUN1, NSUN2 and NSUN4 were found to be overexpressed in a number of human cancers, including breast, prostate, cervi-
cal, and colorectal cancers (25–29). NSUN3 is linked to mitochondrial diseases (30,31), whilst NSUN5 and NSUN7 are likely implicated in Williams-Beuren syndrome (a neurodevelopmental disorder) (32) and male infertility (33), respectively. Given the strong physiological links, there is currently immense interest in studying the biological roles of these enzymes and their potentials as therapeutic targets. However, at present, no methods exist which permit the di-

These enzymes and their potentials as therapeutic targets. Given the strong physiological links, there is currently immense interest in studying the biological roles of these enzymes and their potentials as therapeutic targets. However, at present, no methods exist which permit the di-

rence in methylation between DNA and RNA MTase activities. This severely impedes our understanding of these medically-important enzymes.

It is known that several nucleic acid base methylations can directly impact the secondary structure of DNA and RNA. For instance, we (34–36), and others (37–40), have recently shown that the presence of a single N⁶-methyadenosine (mA) or N⁴-methyladenosine (m⁴A) modification can trigger a duplex-hairpin transformation in certain RNA sequence contexts. It is also clear from a number of NMR and X-ray crystallographic studies that m⁴C modification is able to induce a major B-Z structural change in some DNA sequences. This phenomenon was first reported by Behe and Felsenfeld (41), and subsequently by others (42–44), for a number of DNA duplexes with alternating CpG sequences, such as d(m⁵CGCGm⁵CG), poly d(m⁵CG), and poly d(Gm⁵C)m₆. The structural influence of m⁵C on RNA duplexes, however, is significantly less-studied. One early study on a poly(CG) RNA sequence suggests that m⁵C methylation likely elicit a conformational change from an A-type duplex to an atypical duplex structure (45), whereas a separate study using a different sequence context indicates no significant structural alteration (46). To the best of our knowledge, these two reports provide the only experimental data available in this regard and, to date, it remains unclear whether m⁵C modification has any major impact on the secondary structures of RNA duplexes.

Herein we demonstrated, through a combination of NMR, circular dichroism (CD), and thermodynamic studies, that m⁵C has different conformational effects on DNA and RNA duplexes, even for duplexes with identical CpG sequences. In the case of CpG RNA duplex, it induces a considerable distortion of the phosphate backbone and a C3'-endo to C2'-endo sugar pucker switch in the terminal residues whereas, in the case of CpG DNA duplex, m⁵C triggered a remarkable B-to-Z structural transformation. m⁵C therefore produces distinctly different 'structural signatures' on CpG RNA and DNA duplexes under the same physiologically-relevant conditions.

Inspired by these interesting findings, we envisioned that the different duplex-remodelling effects of m⁵C could provide a basis for the selective detection of RNA and DNA m⁵C MTase activity in cells. This was demonstrated by the development of a novel m⁵C-sensitive nucleic acid probe which, by design, is capable of switching its terminal sugar pucker conformation spontaneously in response to m⁵C methylation (Figure 1). When coupled with an environment-sensitive fluorophore, it provides a powerful visual tool for sensing RNA:m⁵C methylation changes in cells. Prior to this work, we are not aware of any assay methods which are based on m⁵C-induced conformational change.

As we shall demonstrate, the m⁵C-probe is highly-selective and could specifically target NSUN2 over other RNA:m⁵C MTases (including structurally-related subfamily members NSUN3, NSUN5A, NSUN6) and DNA:m⁵C MTases (including DNMT1 and DNMT3A). Through the m⁵C-probe approach, we achieved live cell imaging and flow cytometry analyses of NSUN2 activity in HeLa cells. We also successfully applied the probe to the high-throughput cell-based screening of NSUN2 inhibitors. The discovery of such highly selective probes is rarely achieved and may provide insights on the functions of NSUN2 in m⁵C-regulated processes. Although this study focus on the sensing of RNA:m⁵C MTase activity, the m⁵C-switchable probe strategy outlined here could also be adapted for the study of DNA:m⁵C MTase. This was demonstrated by the successful development of DNA m⁵C-probe which is useful for the in vitro screening of DNMT3A inhibitors.

To the best of our knowledge, this study represents not only the first examples of m⁵C-responsive fluorescent probes, but also a new strategy for the discrimination of RNA and DNA m⁵C MTases activity in cells, which hopefully will inspire the development of epigenetic biosensors and diagnostics tools.

**MATERIALS AND METHODS**

**Preparation of m⁵C-probes**

All oligonucleotide probes investigated in this study were synthesized using standard β-cyanoethyl phosphoramidite chemistry on an automated DNA/RNA synthesiser (Applied Biosystems 394). All synthesiser reagents were purchased from Glen Research; 2′-O-methyl 6-phenylpyrrolocytidine phosphoramidite (47,48) and pyrene deoxynucleoside phosphoramidite (49) were synthesised according to literature procedure. For details of synthesis, see Supplementary Data. All oligonucleotides have a purity of at least 95% (Supplementary Tables S1-S3).

**Studies on the duplex-remodelling effects of m⁵C using NMR, CD and UV-based thermodynamic analyses**

All ¹H NMR experiments were recorded on a Bruker DRX-500 spectrometer operating at 500.23 MHz and the data processed using the standard Bruker processing software TopSpin. Estimated accuracy for protons is within 0.02 ppm. The oligos were dissolved in either D₂O (99.98%) or D₂O solvent mix (pH 7.4) containing 150 mM NaCl and 20 mM MgCl₂. Final concentrations were 3 mM for the unmethylated duplexes (1a, 7a) and 2 mM for the m⁵C-methylated duplexes (1b, 7b). The proton chemical shifts were referenced relative to residual H₂O signal (4.75 ppm at 25°C). The ¹H kinetic spectra of duplexes in 9:1 buffer/D₂O solvent mix (pH 7.4) were acquired at 4°C and 10°C with suppression of water signal by gradient-tailored excitation (WATERGATE) sequence, as previously described (34,50); at these temperatures, the oligos exist in the duplex state, as verified by UV-melting and CD analyses. Two-dimensional
Figure 1. The m$^5$C-switchable probe strategy. (A) The m$^5$C-probe 8a contains a 5'-terminal fluorescent nucleotide PC (2'-O-methyl 6-phenylpyrrolocytidine), which lights up spontaneously in response to m$^5$C-induced terminal sugar pucker switch. (B) When the probe is unmethylated, PC is able to base-pair with guanine in the complementary strand and stack strongly with its adjacent base. This results in efficient quenching of PC fluorescence through photoinduced electron transfer. m$^5$C methylation of the probe by RNA m$^5$C MTase (e.g. NSUN2), however, is expected to trigger a C2'-endo to C3'-endo sugar pucker switch in PC and, since the sugar ring pucker defines the glycosidic bond angle, such a change in sugar puckering will also convert the orientation of PC base from axial to equatorial. This, in turn, disrupts its base-pairing and base-stacking interactions, leading to fluorescence activation. (C) Schematic representation of 2'-OMe RNA probes and their methylated counterparts. The composition of the probes was confirmed through MALDI-TOF mass spectrometric analysis (see Supplementary Table S2). The structure of 'locked 6-phenylpyrrolocytidine' (LC) is shown. The C2'–C4' covalent link is geometrically incompatible with C2'-endo pucker mode, it therefore locks the nucleotide into the C3'-endo conformation.

$^1$H NMR spectra of duplexes in D$_2$O and buffer/D$_2$O were acquired in the phase-sensitive mode using previously described method (51). Two-dimensional NOESY spectra of duplexes in D$_2$O (52–54) were recorded at 25°C with a spectral width of 4000 Hz using 1000 t$_2$ complex points, 1000 t$_1$ (real) points, pulse repetition delay of 2.5 s and mixing times (r$_m$) of 80, 150 and 300 ms. Two-dimensional NOESY spectra of duplexes in 9:1 buffer/D$_2$O were acquired at 10°C with water suppression by WATERGATE sequence (34,50). Spectra were recorded using 2048 t$_2$ complex points, 256 t$_1$ (real) points, pulse repetition delay of 1.7 s and mixing time of 200 ms. Proton-decoupled $^{31}$P NMR spectra of duplexes in 9:1 buffer/D$_2$O were acquired at 10°C on a Bruker AV300 spectrometer equipped with an autotune 5 mm QNP probe. Spectra were recorded at 121.4 MHz, with a spectral width of 10 kHz. The $^{31}$P chemical shifts were externally referenced to 85% phosphoric acid (H$_3$PO$_4$). Estimated accuracy for phosphorus is within 0.03 ppm. For CD spectroscopy and UV-based thermal denaturation experiments, see Supplementary Data.

Fluorescence analyses of m$^5$C-probes

Fluorescence analysis was performed using a Cary Eclipse fluorescence spectrophotometer, as previously reported with modifications (36). The fluorescence emission spectra (λ$_{em}$ = 400–600 nm) of the probes were recorded in a quartz cuvette at 37°C, at a total stand concentration 5 μM in a buffer of 10 mM sodium phosphate buffer (pH 7.4) containing 150 mM NaCl and 20 mM MgCl$_2$. An average of five
scans was recorded. Background fluorescence spectra were acquired after the probe has been incubated for 30 min at 37°C. For determination of fluorescence quantum yields of the m5C-probes, see Supplementary Data.

**In vitro m5C-probe methyltransferase assay**

RNA methyltransferase assay: the assay was performed at 37°C in a Corning Costar 96-well flat bottom black plate. Reaction consisted of enzyme (0.5 μM), methyl donor S-adenosyl-L-methionine (SAM; 200 μM), m5C-probe 9a (substrate; 5 μM) or locked-probe 10a (control; 5 μM) in a 50 mM HEPES buffer (pH 7.4) containing 150 mM NaCl and 20 mM MgCl2. The final reaction volume is 25 μl. The time course of fluorescence activation was recorded immediately after the addition of enzyme with a Tecan ultra microplate reader (λex 360 nm; λem 465 nm). An average of five scans was recorded. DNA methyltransferase assay: the assay was performed as described above. Reaction consisted of enzyme (NSUN2 or DNMT3A; 0.5 μM), methyl donor S-adenosyl-L-methionine (SAM; 200 μM), DNA m5C-probe 11a (substrate; 5 μM) in a 50 mM HEPES buffer (pH 7.4) containing 150 mM NaCl and 20 mM MgCl2. The fluorescence of the reaction was monitored at λem 400 nm (λex 340 nm).

**Steady-state kinetic analyses of the methylation of m5C-probe, tRNA and ssRNA by NSUN2**

The substrates investigated include m5C-probe 9a, human tRNAmet(CAA) (5′-CCAGACUCAGUUCUGG-3′), and CpG-rich ssRNA (5′-CGCGCGCGCGCGCG-3′). The kinetic constants (Km, kcat) of the substrate were determined using an NSUN2 concentration of 0.1 μM and substrate concentrations of 1, 2, 3, 5, 8 and 10 μM. The amount of m5C-methylated product formed at different substrate concentrations was determined based on the relative intensities of the substrate and the product peaks observed in MALDI-TOF mass spectra. All reactions were performed at 37°C in triplicate.

**Fluorescence microscopy**

The HeLa cells were grown in 35 mm glass bottom culture dishes (Thermo Scientific). The cells were transfected with either m5C-probe 9a (10 μM) or control probe 10a (10 μM; control) using Lipofectamine 2000 (Invitrogen) following the manufacturer’s protocols. Fluorescence images were then acquired at 37°C in the presence of 5% CO2 using a Nikon BioStation IM-Q live cell imaging system equipped with 20× and 63× oil immersion objective lenses. The fluorescence images were captured using the DAPI filter setting (i.e. λex 340–380 nm; λem 435–485 nm) and the fluorescence intensity quantified using Biostation IM multichannel software by collecting mean gray values for regions exhibiting fluorescence. All experiments were performed in triplicates and the mean fluorescence intensities were normalised to that of the control. The corresponding transmitted light image for each fluorescence image was also acquired.

**Single-cell flow cytometry analysis**

The HeLa cells were transfected with m5C-probe 9a (10 μM) or control probe 10a (10 μM) using Lipofectamine 2000. They were then collected by trypsinization at 37°C with 5% CO2, washed with phosphate buffered saline (PBS) media, pH 7.4, resuspended in PBS media, and analysed using a BD LSFRFortessa™ flow cytometer (BD Bioscience). Fluorescence was measured using an excitation laser of 355 nm, and a 450/50 bandpass emission filter. Acquisition was stopped when 20,000 events per sample were acquired. The fluorescence data were then analysed using BD FACS software. Violin plot was produced using the Origin software.

**Cell-based m5C-probe methyltransferase assay**

The HeLa cells were pre-incubation with various concentrations of non-specific MTase inhibitors, including sinedfungin, S-adenosyl-L-homocysteine (SAH), adenosine, and homocysteine at 37°C for 30 min. Nine different concentrations of inhibitors were used (0.1, 0.3, 1, 3, 10, 50, 100, 250, 1000 μM) and the final DMSO concentration was kept at less than 1% (v/v) of the assay mix. These concentrations of inhibitors were shown to be non-toxic to HeLa cells in MTT cytotoxicity assays. This was followed by the delivery of either the m5C-probe 9a (10 μM) or control probe 10a (10 μM) into the cells via Lipofectamine 2000. Aliquots were then collected for flow cytometry measurements (λex 355 nm; λem 425–475 nm) at the indicated time points. The mean fluorescence intensity of at least 20,000 live cells was determined. IC50 determination: the IC50 values were then calculated from the variation in fluorescence at different inhibitor concentrations using nonlinear regression, with normalised dose-response fit on GraphPad Prism 6.0™. The assay was performed in triplicate for each inhibitor concentration. Z’ factor determination: the Z’ factor for our m5C-probe assay was calculated using previously described method (55). For details, see Supplementary Data.

**Cell culture, gene silencing and gene overexpression**

HeLa cells (ATCC) were cultured in Dulbecco’s modified Eagle’s medium (DMEM: Life Technologies) supplemented with 10% fetal bovine serum (FBS) and antibiotics (100 units/ml penicillin, 100 μg/ml streptomycin). Cells were grown at 37°C in a 5% CO2 humidified incubator. To silence the expression of specific gene transiently, the cells were first plated in a 24-well plate (0.5 × 105 cells/well). The next day, the cells were transfected with either siNSUN2, siNSUN6, siTRDMT1, siDNMT3A or siControl using Lipofectamine RNAiMAX (Invitrogen), according to manufacturer’s protocols. The sequences used for siRNAs and siControls are provided in the Supplementary Data. The protein levels were analysed by western blot using specific antibodies. Overexpression of specific gene. The construction of pcDNA3-Flag vectors expressing NSUN2, TRDMT1 and DNMT3A was previously described (56–58). Full length NSUN2, TRDMT1 and DNMT3A genes were amplified by PCR using human HeLa cells cDNA, then subcloned into the EcoRI-BamHI sites of mammalian pcDNA3-Flag vector (Invitrogen) to obtain pcDNA3-Flag-NSUN2, pcDNA3-Flag-TRDMT1, and pcDNA3-Flag-DNMT3A.
plasmids, respectively. The plasmids were then transfected into HeLa cells using the calcium chloride transfection method. For construction of pcDNA3-Flag-NSUN6 plasmid (59), the cDNA encoding NSUN6 was PCR-amplified using Phusion DNA polymerase and subcloned into the KpnI-XhoI sites of pcDNA3.1-Flag with CloneExpress II One step Cloning Kit (Vazyme). The primer sequences used for vector construction are given in Supplementary Data.

Western blot analysis

The HeLa cells were first lysed using radioimmunoprecipitation assay (RIPA) buffer (i.e. 25 mM Tris–HCl (pH 7.6), 150 mM NaCl, 0.1% SDS, 1% sodium deoxycholate, 1% NP-40 and 1× protease inhibitors), followed by immunoblotting using standard protocol. In brief, the extract was fractionated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), transferred onto nitrocellulose filter membranes (Whatman), then incubation with the respective antibodies. This was followed by incubation with either horseradish peroxidase (HRP)-conjugated antirabbit IgG antibody (Bio-Rad) or HRP-conjugated antimouse IgG antibody (Bio-Rad) for 2 h at room temperature. Enhanced chemiluminescence substrates (Luminata Crescendo, EMD Millipore) were then applied and the signals exposed to autoradiography film. The immunoblots were then quantified by densitometric analyses using ImageJ software. All antibodies used were purchased from commercial sources (for details, see Supplementary Data).

RESULTS AND DISCUSSION

m5C induces a local distortion of the phosphate backbone and a C3′-endo to C2′-endo sugar pucker switch in CpG RNA duplex

To date, there has not been a systematic study directly comparing the structural impact of m5C methylation on CpG RNA duplexes and CpG DNA duplexes. Such an analysis might provide insights into the respective conformational and/or mechanistic roles of m5C in DNA and RNA, hence is of special interest. To investigate this question in multiple sequence contexts, we prepared a set of 6-mer and 12-mer RNA duplexes containing either alternating CpG sequences (1a–2a), or C3′-endo/C2′-endo (3a–6a), as well as their corresponding m5C-methylated counterparts (1b–6b; Table 1). To facilitate comparison, the CpG DNA duplexes 7a and 7b were also generated. For details of their chemical synthesis, see Supplementary Data.

We began by analysing the secondary structure of CpG RNA hexamer r(CG-C-GCG)2, 1a and its methylated equivalent r(CG-m5C-GCG)2, 1b using 1H and 31P NMR. Assignment of the proton resonances was accomplished by 2D NOESY experiments and by comparison with reported NMR data for 1a (60). The results are summarised in Figure 2 and Table 2. Consistent with previous NMR studies (60), the imino proton spectrum of 1a exhibited three NH signals at 10°C, which is expected for a palindromic duplex with six base-pairs (Figure 2A). Upon m5C methylation (1b), we observed a downfield shift in G2 resonance by ~0.35 ppm and a disappearance of the terminal G6 resonance. The latter is likely due to ‘fraying’ of the terminal base-pairs as the G6 signal could again be detected at a lower temperature of 4°C (Supplementary Figure S1). This result suggests that m5C modification likely promotes the disruption of base-pairing interactions at duplex terminal ends. As anticipated, there is little or no change in the imino resonance of G4 which is involved in base-pairing with m5C, confirming that C5-cytosine methylation does not hinder C:G Watson–Crick base-pairing interactions. Nevertheless, we did observe considerable variations in 31P resonances between the methylated and unmethylated duplexes, indicating a significant distortion of the phosphate backbone (Figure 2B).

Prompted by this finding, we proceeded to examine whether changes in phosphate backbone also alter the sugar ring geometry. To this end, we determined the coupling constant between sugar proton H1′ and H2′ (JH1′−H2′), which is highly sensitive to the dihedral angle and, therefore, provides a direct indication of the sugar pucker conformation. In general, a JH1′−H2′ value <2 Hz is suggestive of a C3′-endo sugar pucker (North type), whereas JH1′−H2′ value >8 Hz is characteristic of a C2′-endo sugar puck (South type) (61). The JH1′−H2′ values for 1a and 1b are summarised in Table 2 and the representative spectra are shown in Supplementary Figure S2. In the absence of cytosine methylation, all ribose rings in 1a exhibit a C3′-endo conformation (JH1′−H2′ ~1.2 Hz), which is the preferred sugar puckering mode for A-form RNA duplex. Upon m5C methylation (1b), however, the 5′-terminal cytosine (C1) undergoes a sugar pucker switch from C3′-endo to C2′-endo, as evidenced by a large increase in JH1′−H2′ from 1.1 Hz to 8.2 Hz. This is accompanied by a switch in sugar puckering of the 3′-terminal guanosine (G6) to an intermediate conformation between C3′-endo and C2′-endo (JH1′−H2′ = 6.0 Hz). No changes in sugar puckering mode are observed for the four internal residues G2 to C5, which continue to adopt a C3′-endo orientation. Our data support the findings of an earlier NMR study (45) and, together, they demonstrate that m5C can alter the sugar pucker conformation of the 5′- and 3′-terminal residues in CpG RNA duplexes.

To confirm our results, we further analysed the intensity of intranucleotide NOEs between the base protons (H6 in cytosine; H8 in guanine) and the sugar protons (H2′ and H3′), which correlates with the distance between the protons (Figures 2C, D and Table 2) (62,63). Consistent with our JH1′−H2′ data, the C1 residue of 1b displays a strong NOE between its base proton H6 and sugar proton H2′ (~2.2 Å) and a weak NOE between H6 and H3′ (~3.9 Å), which is indicative of a C2′-endo conformation, and the G6 residue of 1b shows approximately similar NOE intensities for H8-H2′ (~2.7 Å) and H8-H3′ (~3.3 Å), suggesting an intermediate C2′-endo/C3′-endo conformation. In comparison, both the C1 and G6 residues of unmethylated 1a exhibit a C3′-endo sugar pucker, thus m5C methylation indeed produces a switch in the terminal sugar pucker orientation.

Notably, however, these structural perturbations appear to have little or no effect on the overall secondary structure of the RNA duplex. This is apparent from CD analysis where the spectra of 1a and 1b exhibit characteristics which are typical of a right-handed A-form duplex, and they superimpose almost perfectly with each other (Figure 3A). Hence, m5C modification likely only affects the local
Table 1. Sequences of RNA and DNA duplexes investigated in this study. The $^3J_{H'}-H_2$ values for the 5'- and 3'-terminal residues of each sequence are indicated in red and green, respectively. m5C-induced terminal sugar pucker switch was observed in 6-mer and 12-mer CpG duplexes (1b and 2b) but not in out-of-alternation CpG duplex 3b and random duplexes 4b–6b, suggesting that the structural effect of m5C may be unique to alternating CpG RNA sequences. m5C methylation of 7a i.e. the DNA equivalent of 1a caused a C2'-endo to C3'-endo sugar pucker switch in all guanosine residues, whilst all cytosine residues preserved their C2'-endo orientation, consistent with a B-Z structural conversion.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Sequences$^a$</th>
<th>Conformations</th>
<th>$^3J_{H'}-H_2$ (Hz)$^b$</th>
<th>Sugar pucker mode</th>
<th>$^3J_{H'}-H_2$ (Hz)$^b$</th>
<th>Sugar pucker mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>5'-r(CG C G CG)-3' 3'-r(GC G C GC)-5'</td>
<td>A-RNA</td>
<td>1.1</td>
<td>C3'-endo</td>
<td>2.0</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>1b</td>
<td>5'-r(CG C G CG) m5C 3'-r(GC G C GC)-5'</td>
<td>A-RNA</td>
<td>8.2</td>
<td>C2'-endo</td>
<td>6.0</td>
<td>C2'-endo</td>
</tr>
<tr>
<td>2a</td>
<td>5'-r(CG C G GCCG)-3' 3'-r(GC G CGCG C G C)-5'</td>
<td>A-RNA</td>
<td>1.2</td>
<td>C3'-endo</td>
<td>1.1</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>2b</td>
<td>5'-r(CG m5C C G GCCG)-3' 3'-r(GC C GCGCG C G C)-5'</td>
<td>A-RNA</td>
<td>8.3</td>
<td>C2'-endo</td>
<td>7.9</td>
<td>C2'-endo</td>
</tr>
<tr>
<td>3a</td>
<td>5'-r(CG C G CG)-3' 3'-r(GC G CGC)-5'</td>
<td>A-RNA</td>
<td>1.0</td>
<td>C3'-endo</td>
<td>1.0</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>3b</td>
<td>5'-r(CG m5C C G CG)-3' 3'-r(GC G CGCGC)-5'</td>
<td>A-RNA</td>
<td>1.0</td>
<td>C3'-endo</td>
<td>1.0</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>4a</td>
<td>5'-r(U C G C G AU)-3' 3'-r(A G C C UA)-5'</td>
<td>A-RNA</td>
<td>1.0</td>
<td>C3'-endo</td>
<td>1.1</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>4b</td>
<td>5'-r(U m5C C G AU)-3' 3'-r(A G C C UA)-5'</td>
<td>A-RNA</td>
<td>1.1</td>
<td>C3'-endo</td>
<td>1.2</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>5a</td>
<td>5'-r(G U C G G A)-3' 3'-r(C A G C C U)-5'</td>
<td>A-RNA</td>
<td>1.0</td>
<td>C3'-endo</td>
<td>1.2</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>5b</td>
<td>5'-r(G U m5C C G A)-3' 3'-r(C A C C U)-5'</td>
<td>A-RNA</td>
<td>1.2</td>
<td>C3'-endo</td>
<td>1.1</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>6a</td>
<td>5'-r(G C UGACUA G UA)-3' 3'-r(G C ACUGAU C AU)-5'</td>
<td>A-RNA</td>
<td>1.1</td>
<td>C3'-endo</td>
<td>1.0</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>6b</td>
<td>5'-r(G m5C C UGACUA G UA)-3' 3'-r(G C ACUGAU C AU)-5'</td>
<td>A-RNA</td>
<td>1.2</td>
<td>C3'-endo</td>
<td>1.2</td>
<td>C3'-endo</td>
</tr>
<tr>
<td>7a</td>
<td>5'-d(CG C G CG)-3' 3'-d(GC G C GC)-5'</td>
<td>B-DNA</td>
<td>8.1</td>
<td>C2'-endo</td>
<td>8.0</td>
<td>C2'-endo</td>
</tr>
<tr>
<td>7b</td>
<td>5'-m5C d(CG C G CG)-3' 3'-d(GC G C GC)-5'</td>
<td>Z-DNA</td>
<td>8.4</td>
<td>C2'-endo</td>
<td>1.2</td>
<td>C3'-endo</td>
</tr>
</tbody>
</table>

$^a$The composition of the oligos was confirmed through MALDI-TOF mass spectrometric analysis (see Supplementary Table S1).

$^b$Observed at 25°C in D$_2$O.
Figure 2. \( m^3C \) methylation directly alters the phosphate backbone in CpG RNA duplex, leading to a C3'-endo to C2'-endo sugar pucker switch of the terminal residues under physiologically-relevant conditions. (A) Representative \( ^1H \) imino NMR spectra of r(CG-C-GCG)\(_2\) 1a (top) and its methylated counterpart r(CG-\( m^3C \)-GCG)\(_2\) 1b (bottom) in 9:1 buffer/D\(_2\)O solvent mix acquired at 10°C. The buffer used was 10 mM sodium phosphate buffer (pH 7.4) containing 150 mM NaCl and 20 mM MgCl\(_2\). The G6 imino resonance disappeared upon \( m^3C \) methylation, presumably due to 'fraying' of the terminal base-pairs. (B) Proton-decoupled \( ^31P \) NMR spectra of 1a (top) and 1b (bottom) in 9:1 buffer/D\(_2\)O solvent mix acquired at 10°C. There is considerable variation in \( ^31P \) resonances between the methylated and unmethylated duplexes, suggesting a significant distortion of the phosphate backbone. Representative 2D NOESY spectra of the H2'/H3'-aromatic region of (C) 1a and (D) 1b in D\(_2\)O at 25°C (pH 7.4; \( \tau_m = 150 \) ms) showing the H6/8-H2' NOE connectivity pathways (blue dashed-line) and the H6/8-H3' NOE connectivity pathways (black solid line). The schematic illustrations (left) showed the likely sugar pucker mode for terminal residues C1 and G6, and their calculated intranucleotide base-sugar distances (labelled in red; determined based on the intensity of NOEs between the base and sugar protons).
Table 2. 1H NMR chemical shifts (ppm) for CpG RNA duplex 1a (top) and its m5C-methylated counterpart 1b (bottom) observed at 25°C in D2O with WATERGATE suppression (τ = 150 ms). Assignment of the proton resonances was accomplished by 2D NOESY experiments, and by comparison with reported NMR data for 1a (60). The representative NMR spectra for the determination of JH1′−H2′ coupling constants are given in Supplementary Figure S2. For calculation of sugar-base proton distances, see Supplementary Data.

<table>
<thead>
<tr>
<th>Residues</th>
<th>H5</th>
<th>H6/H8</th>
<th>H1′</th>
<th>H2′</th>
<th>H3′</th>
<th>H4′</th>
<th>H5′</th>
<th>H6/H8/H2′</th>
<th>H6/H8/H3′</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>6.00</td>
<td>8.03</td>
<td>5.53</td>
<td>4.55</td>
<td>5.40</td>
<td>4.23</td>
<td>4.03</td>
<td>3.93</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>–</td>
<td>7.83</td>
<td>5.79</td>
<td>4.57</td>
<td>4.73</td>
<td>4.52</td>
<td>4.52</td>
<td>4.20</td>
<td>13.15</td>
</tr>
<tr>
<td>C3</td>
<td>5.31</td>
<td>7.73</td>
<td>5.55</td>
<td>4.52</td>
<td>4.59</td>
<td>4.46</td>
<td>4.58</td>
<td>4.15</td>
<td>1.16</td>
</tr>
<tr>
<td>G4</td>
<td>–</td>
<td>7.56</td>
<td>5.75</td>
<td>4.52</td>
<td>4.56</td>
<td>4.48</td>
<td>4.51</td>
<td>4.13</td>
<td>12.97</td>
</tr>
<tr>
<td>C5</td>
<td>5.25</td>
<td>7.58</td>
<td>5.52</td>
<td>4.35</td>
<td>4.47</td>
<td>4.40</td>
<td>4.52</td>
<td>4.07</td>
<td></td>
</tr>
<tr>
<td>G6</td>
<td>–</td>
<td>7.61</td>
<td>5.65</td>
<td>4.11</td>
<td>4.29</td>
<td>4.24</td>
<td>4.45</td>
<td>4.05</td>
<td>12.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residues</th>
<th>H5/S-Me</th>
<th>H6/H8</th>
<th>H1′</th>
<th>H2′</th>
<th>H3′</th>
<th>H4′</th>
<th>H5′</th>
<th>H6/H8/H2′</th>
<th>H6/H8/H3′</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>6.04</td>
<td>8.01</td>
<td>5.51</td>
<td>4.54</td>
<td>4.54</td>
<td>4.37</td>
<td>4.09</td>
<td>4.09</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>–</td>
<td>7.80</td>
<td>5.74</td>
<td>4.59</td>
<td>4.72</td>
<td>4.55</td>
<td>4.58</td>
<td>4.23</td>
<td>13.50</td>
</tr>
<tr>
<td>m5C3 (1.53)</td>
<td>7.09</td>
<td>5.47</td>
<td>4.50</td>
<td>4.63</td>
<td>4.52</td>
<td>4.63</td>
<td>4.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>–</td>
<td>7.56</td>
<td>5.79</td>
<td>4.50</td>
<td>4.59</td>
<td>4.55</td>
<td>4.57</td>
<td>4.17</td>
<td>12.93</td>
</tr>
<tr>
<td>C5</td>
<td>5.26</td>
<td>7.60</td>
<td>5.28</td>
<td>4.33</td>
<td>4.46</td>
<td>4.42</td>
<td>4.55</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>G6</td>
<td>–</td>
<td>7.54</td>
<td>5.16</td>
<td>4.06</td>
<td>4.28</td>
<td>4.27</td>
<td>4.51</td>
<td>4.07</td>
<td>6.00 c</td>
</tr>
</tbody>
</table>

m5C triggers a B-to-Z DNA transformation in CpG DNA duplex

We next investigated the structural influence of m5C on CpG DNA duplex 7a, i.e. the DNA equivalent of 1a under the same physiologically-relevant conditions. A preliminary analysis of the JH1′−H2′ values and sugar-base proton distances of 7a and 7b revealed a switch in sugar puckering of all guanosine residues from C2′-endo to C3′-endo mode, whilst all cytosine residues preserve their C2′-endo orientation, which is highly suggestive of a B-to-Z structural transformation (Table 3; the representative 2D NOESY spectrum of 7b is shown in Supplementary Figure S3). As further evidence, CD analysis showed a characteristic inversion of the CD spectrum upon m5C methylation, confirming that m5C-induced terminal sugar puckers switch is likely unique to alternating CpG RNA sequences

m5C triggers a B-to-Z DNA transformation in CpG DNA duplex

We next investigated the structural influence of m5C on CpG DNA duplex 7a, i.e. the DNA equivalent of 1a under the same physiologically-relevant conditions. A preliminary analysis of the JH1′−H2′ values and sugar-base proton distances of 7a and 7b revealed a switch in sugar puckering of all guanosine residues from C2′-endo to C3′-endo mode, whilst all cytosine residues preserve their C2′-endo orientation, which is highly suggestive of a B-to-Z structural transformation (Table 3; the representative 2D NOESY spectrum of 7b is shown in Supplementary Figure S3). As further evidence, CD analysis showed a characteristic inversion of the CD spectrum upon m5C methylation, confirming that m5C methylation affects terminal sugar pucker conformation in alternating CpG duplexes 1a and 2a. The JH1′−H2′ values are summarised in Supplementary Table S4. For 6-mer CpG duplexes, terminal sugar switch occurred only when the third cytosine residue (C3) was methylated (i.e. 1b), and not when the first (1c) or second (1d) cytosine residue was methylated. Similarly for 12-mer CpG duplexes, m5C-induced sugar switch was observed only when methylation was on C3 (2b), and not on C1 (2c) or C5 (2d), thus the structural effect of m5C is highly-dependent on the position of m5C methylation.

The sequence requirements and mechanisms by which m5C induces terminal sugar pucker switch are unclear at present and require further investigations. Nevertheless, the present study, the first to investigate the structural influence of m5C on multiple RNA sequence contexts, suggests that this conformational effect might be unique to alternating CpG RNA sequences.
m^5C does indeed trigger a major B-Z structural transformation in CpG DNA duplex 7a (Figure 3B). This result is notable because it demonstrates that m^5C produces different structural outcomes on RNA and DNA duplexes.

The observed B-Z conversion is likely attributed to the ability of m^5C to stabilize the Z-form DNA over its B-form structure (42,64). Indeed, UV-based thermodynamic analysis of 7a showed that m^5C methylation increases the stability of its Z-duplex by ∼2.5 kcal/mol relative to its B-duplex (compare ΔG°310 of 7a with 7b; Supplementary Table S3 and Supplementary Figure S4). Moreover, m^5C-induced B-Z helicity change is primarily an entropy-driven process, as judged by the more favourable entropic contributions (ΔΔS°7a→7b = 31.6 cal/mol/K) which counteracts the enthalpy cost (ΔΔH°7a→7b = 7.3 kcal/mol). Further thermodynamic analysis of 7b in the absence of MgCl2 (a known Z-conformation inducer) revealed a free energy of B-to-Z transition (ΔΔG°+1) of ∼1.5 kcal/mol (compare ΔG°310 of 7b with 7b*); since 7b does not undergo B-Z conversion in the absence of MgCl2, any free energy differences observed at these two salt conditions, i.e. between 7b and 7b* is largely ascribed to conformational transition (Supplementary Table S3 and Supplementary Figure S5).

m^5C does not promote A-Z transition in CpG RNA duplex

To date, the effects of m^5C methylation on the stability of Z-form RNA is unclear. We are also not aware of any prior studies investigating whether m^5C could promote an analogous A-Z transition in RNA duplex. Therefore, we performed further CD experiments to monitor the conformation of CpG RNA duplex 1a and its methylated equivalent 1b at a range of MgCl2 concentrations (1–300 mM). The results showed that both 1a and 1b maintain an A-type double helical structure at all MgCl2 concentrations investigated (Supplementary Figure S6). In fact, the methylated RNA 1b could not be induced to form Z-RNA even at MgCl2 concentration as high as 300 mM. Consistent with this observation, the presence of m^5C does not significantly improve the stability of 1a (ΔΔG°310 ∼0.8 kcal/mol; Supplementary Table S3). Thus m^5C methylation favours B-Z conversion in CpG DNA duplex, but not A-Z transition in CpG RNA duplex under the same physiological salt and pH conditions.

Overall, our results demonstrated that m^5C methylation can directly influence the conformation of RNA and DNA duplexes. In the case of CpG RNA hexamer, m^5C modification induces a local distortion of the phosphate backbone and a C3′-endo to C2′-endo terminal sugar pucker switch whereas, in CpG DNA hexamer, m^5C triggers a transformation from a right-handed B-DNA to a left-handed Z-DNA. We appreciate that detailed crystallographic studies are required to fully characterise these structural changes, nevertheless, our available data clearly demonstrated that m^5C modification produces distinctly different ‘structural signatures’ on RNA and DNA duplexes, even for duplexes with identical CpG sequence. The significance of this finding is unclear at present, however it might infer a possible involvement of m^5C-induced conformational change in biological regulatory processes.

Design principle of m^5C-switchable probes

Inspired by these interesting findings, we envisaged that the different structure-remodelling effects of m^5C on RNA and DNA duplexes could be exploited for the design of ‘methyltion-switchable probe’ useful for studying m^5C methylation in DNA and RNA. Notably, at present, no methods exist which permit the direct analysis of m^5C MTase activity in living cells. There are also no reports of assay methods that are able to distinguish between the activities of DNA m^5C MTases and RNA m^5C MTases.

As a proof-of-principle study, we developed a novel m^5C-probe which is able to fluoresce spontaneously in response to m^5C-induced terminal sugar pucker switch, hence useful for sensing RNA:m^5C MTase activity. The probe (8a; Figure 1) is modelled after the CpG RNA duplex 1a but contains two important modifications. First, in order to visualise m^5C-induced sugar pucker switch, the 5′-cytosine residues on the forward strand was replaced with the fluorescent cytosine analogue 6-phenylpyrrolocytosine (pC) (47,48). pC is well-suited as reporter for our probe because it is highly fluorescent (quantum yield (ΦF) 0.31) and its emission is extremely sensitive to microenvironment changes (65–67). In particular, previous photophysical studies showed that 5C emits brightly when in a single-stranded environment, but becomes significantly quenched when hybridised with its complementary strand (65–67). Second, to render the m^5C-probe sufficiently stable for live-cell applications, the RNA backbone was replaced with a 2′-O-methyl backbone, which is reasonably resistant to cellular nuclease degradation.

The analytical principle of the m^5C-probe is illustrated in Figures 1A and B. By design, when the probe is unmethylated, the pC fluorophore is able to base-pair with guanine and stack strongly with its adjacent base. This results in efficient quenching of pC fluorescence through photoinduced electron transfer (PET). m^5C methylation of the probe by RNA:m^5C MTases, however, is expected to trigger a spontaneous C3′-endo to C2′-endo sugar pucker switch in pC. Since the sugar ring pucker defines the glycosidic bond angle, a change in sugar puckering mode will also convert the orientation of pC base from axial to equatorial, thereby disrupting its base-pairing and base-stacking interactions, leading to fluorescence activation.

The m^5C-probe has the sensitivity to detect a single m^5C mark change

To verify our probe design, we first compared the fluorescence emission of the probes at different m^5C methylation levels (λex 360 nm; λem 465 nm; Supplementary Figure S7). Consistent with our detection strategy, the starting unmethylated probe 8a displayed weak fluorescence, with a relatively low fluorescence quantum yield (ΦF) of ∼0.03, whilst its methylated counterpart 8b yielded a 2-fold increase in fluorescence intensity (ΦF ∼0.06; Supplementary Figure S7). Similarly the unmethylated probe 8c, which was designed to simultaneously exploit sugar pucker switch at both 5′-ends, also gave negligible fluorescence response (ΦF ∼0.04; Supplementary Figure S7). This result is notable because most DNA/RNA duplexes are known
Table 3. 1H NMR chemical shifts δH (ppm) for Cpg DNA duplex 7a (top) and its m5C-methylated counterpart 7b (bottom) observed at 25°C in D2O with WATERGATE suppression. The representative 2D NOESY spectrum is given in Supplementary Figure S3. For calculation of sugar-base proton distances, see Supplementary Data

<table>
<thead>
<tr>
<th>Residues</th>
<th>H5</th>
<th>H6/H8</th>
<th>H1′</th>
<th>H2′</th>
<th>H2⁎</th>
<th>H3′</th>
<th>H4′</th>
<th>H5′</th>
<th>H5⁎</th>
<th>Imino*</th>
<th>JH1–H2 (Hz)</th>
<th>Sugar pucker conformation</th>
<th>Base-sugar proton distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>5.85</td>
<td>7.64</td>
<td>5.93</td>
<td>7.52</td>
<td>2.39</td>
<td>4.68</td>
<td>4.10</td>
<td>3.71</td>
<td>3.69</td>
<td></td>
<td>8.1</td>
<td>C2′-endo (S)</td>
<td>2.6</td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td>8.00</td>
<td>6.00</td>
<td>2.70</td>
<td>2.73</td>
<td>4.98</td>
<td>4.36</td>
<td>4.21</td>
<td>4.05</td>
<td></td>
<td>8.6</td>
<td>C2′-endo (S)</td>
<td>2.3</td>
</tr>
<tr>
<td>C3</td>
<td>5.67</td>
<td>7.15</td>
<td>5.73</td>
<td>2.00</td>
<td>2.40</td>
<td>4.82</td>
<td>4.20</td>
<td>4.16</td>
<td>4.15</td>
<td></td>
<td>8.5</td>
<td>C2′-endo (S)</td>
<td>2.1</td>
</tr>
<tr>
<td>G4</td>
<td>7.97</td>
<td>5.92</td>
<td>6.08</td>
<td>2.76</td>
<td>4.95</td>
<td>4.37</td>
<td>4.11</td>
<td>4.04</td>
<td>4.11</td>
<td></td>
<td>8.5</td>
<td>C2′-endo (S)</td>
<td>2.2</td>
</tr>
<tr>
<td>C5</td>
<td>5.44</td>
<td>7.31</td>
<td>5.76</td>
<td>1.90</td>
<td>2.34</td>
<td>4.41</td>
<td>4.16</td>
<td>4.20</td>
<td>4.19</td>
<td></td>
<td>8.3</td>
<td>C2′-endo (S)</td>
<td>2.0</td>
</tr>
<tr>
<td>G6</td>
<td></td>
<td>7.92</td>
<td>6.17</td>
<td>2.61</td>
<td>2.37</td>
<td>4.68</td>
<td>4.18</td>
<td>4.12</td>
<td>4.11</td>
<td></td>
<td>8.0</td>
<td>C2′-endo (S)</td>
<td>2.3</td>
</tr>
</tbody>
</table>

m5C-methylated Cpg DNA duplex: d(CG-m5CGGCG)7b

<table>
<thead>
<tr>
<th>Residues</th>
<th>H5/S-Me</th>
<th>H6/H8</th>
<th>H1′</th>
<th>H2′</th>
<th>H2⁎</th>
<th>H3′</th>
<th>H4′</th>
<th>H5′</th>
<th>H5⁎</th>
<th>Imino*</th>
<th>JH1–H2 (Hz)</th>
<th>Sugar pucker conformation</th>
<th>Base-sugar proton distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>5.66</td>
<td>7.30</td>
<td>5.66</td>
<td>1.58</td>
<td>2.40</td>
<td>4.58</td>
<td>3.82</td>
<td>3.50</td>
<td>3.53</td>
<td></td>
<td>8.4</td>
<td>C2′-endo (S)</td>
<td>2.2</td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td>7.82</td>
<td>6.24</td>
<td>3.19</td>
<td>2.50</td>
<td>4.75</td>
<td>4.17</td>
<td>4.01</td>
<td>3.85</td>
<td></td>
<td>13.24</td>
<td>C2′-endo (S)</td>
<td>3.6</td>
</tr>
<tr>
<td>m5C3 (109)</td>
<td></td>
<td>7.35</td>
<td>5.75</td>
<td>1.06</td>
<td>2.55</td>
<td>4.79</td>
<td>3.76</td>
<td>3.73</td>
<td>3.65</td>
<td></td>
<td>8.1</td>
<td>C2′-endo (S)</td>
<td>2.1</td>
</tr>
<tr>
<td>C4</td>
<td>7.64</td>
<td>6.29</td>
<td>2.78</td>
<td>2.57</td>
<td>4.93</td>
<td>4.18</td>
<td>3.96</td>
<td>3.99</td>
<td>3.99</td>
<td></td>
<td>13.17</td>
<td>C2′-endo (S)</td>
<td>3.7</td>
</tr>
<tr>
<td>G5</td>
<td>5.23</td>
<td>7.07</td>
<td>5.67</td>
<td>1.59</td>
<td>2.58</td>
<td>4.80</td>
<td>3.77</td>
<td>3.91</td>
<td>3.92</td>
<td></td>
<td>8.3</td>
<td>C2′-endo (S)</td>
<td>2.1</td>
</tr>
<tr>
<td>G6</td>
<td></td>
<td>7.87</td>
<td>6.30</td>
<td>2.82</td>
<td>2.96</td>
<td>4.94</td>
<td>4.20</td>
<td>3.98</td>
<td>3.85</td>
<td></td>
<td>1.2</td>
<td>C2′-endo (S)</td>
<td>3.6</td>
</tr>
</tbody>
</table>

aObserved at 10°C in 9:1 buffer/D2O solvent mix. The buffer used was 10 mM sodium phosphate buffer (pH 7.4) containing 150 mM NaCl and 20 mM MgCl2.

bChemical shift for 5-methyl group on m5C.

Note: 4°C.

to undergo transient opening of the terminal base-pairs in solution. Conceivable, this might cause auto-activation of 1C fluorescence, which will severely limit the applications of these probes. Nevertheless, our fluorescence data from starting probes 8a and 8c suggests that emission arising from such transient end-fraying event is relatively insignificant. Indeed, with the improved probe 8c, we are able to clearly detect a 2.8-fold increase in fluorescence intensity upon m5C methylation (8d: ΔΦE ∼ 0.11; Supplementary Figure S7), which is remarkable in light of only a single m5C mark change. Thus our m5C-probe has the sensitivity to detect subtle methylation changes. Moreover, the introduction of a second m5C modification, i.e. 8e led to approximately a doubling of fluorescence emission (5.3-fold), hence the fluorescence intensity of the probe increases proportionately with m5C methylation level.

Interestingly, the fluorescence output of the probe could be further increased by incorporating additional 1C fluorophores, as demonstrated by probe 9a (Figure 4A) where the concurrent replacement of cytosine 1 and 5 in both strands with 1C led to significant improvements in fluorescence light-up response (ΔΦE (9b) 3.6-fold; (9c) 7.2-fold). Despite the introduction of 2′-O-Me backbone and four relatively bulky 1C residues, NMR data indicate that both the 5′- and 3′-terminal residues of probe 9a continue to adopt a C3′-endo pucker conformation (JH1–H2 ∼ 1–2 Hz; Supplementary Table S4), and they spontaneously assume a C2′-endo pucker orientation (JH1–H2 ∼ 8 Hz) upon m5C methylation (9c), thus these modifications do not interfere with fluorescence activation of the probe. There is also minimal disturbance to the overall duplex structure of probe 9a relative to parent probes 8a and 1a, as verified by CD analysis (Supplementary Figure S8). Amongst the probes investigated, 9a gives the greatest fluorescence response, it was therefore selected as representative 1C-probe for subsequent evaluations.

The m5C-probe is highly-selective for NSUN2 over other RNA/DNA:m5C MTases

We next examined whether m5C-probe 9a could function as fluorogenic substrate for RNA:m5C MTase by testing it against NSUN2, which is of particular interest in light of emerging evidence linking NSUN2 with m5C methylation of human mRNA (14) and its association with human cancers (25–29). In a typical assay, 5 µM of 9a was incubated with NSUN2 (0.5 µM) and methyl donor S-adenosyl-L-methionine (SAM; 200 µM) under physiologically-relevant conditions (at 37°C in 50 mM HEPES buffer containing 150 mM NaCl and 20 mM MgCl2, pH 7.4). The formation of methylated probe was measured by an increase in fluorescence signal at 465 nm (λex, 360 nm). Our result showed that the probe could be readily methylated by NSUN2, yielding ~8-fold increase in fluorescence signal after 3 h (Figure 4B).

To rule out fluorescence activation due to other non-specific mechanisms, we repeated the assay in the absence of NSUN2 or methyl donor SAM; in both cases, there was no detectable fluorescence light-up response (Supplementary Figure S9). Furthermore, pre-incubation of NSUN2 with 10 µM of sinfungin (a non-specific DNA/RNA MTase inhibitor) (25,68,69) resulted in a significant reduction in fluorescence intensity, confirming that fluorescence activation was specifically mediated by NSUN2 MTase activity (Figure 4B). Notably, the fluorescence intensity increases linearly with the concentration of methylated probe, thus providing a direct read-out of NSUN2 MTase activity (Figure 4C).

We next verified whether m5C-induced sugar pucker switch in 1C was indeed responsible for fluorescence turn-on. For this purpose, we synthesised ‘locked-probe’ 10a (Figure 1C) wherein the 1C fluorophore was conformationally restricted in a C3′-endo sugar pucker orientation via a methylene bridge connecting O2’ with C4’ (1C; Figure 1C). As a result of this modification, the locked-probe will
no longer be able to undergo sugar pucker switch upon m5C methylation. Consistent with our hypothesis, 10a gave negligible fluorescence light-up response when exposed to DNA:m5C MTases (Supplementary Figures S9 and S10). Steady-state kinetic analysis using MALDI-TOF MS assay indicated that the m5C-probe is a relatively good substrate for NSUN2 (kcat/Km = 0.3 min⁻¹ μM⁻¹), and is only ~2-fold less efficiently methylated compared with its native tRNA substrate, tRNA44(CAA), and CpG-rich ssRNA (Figures 4D and E).

The m5C-probe is applicable for live-cell imaging, single-cell flow cytometry analysis and cell-based inhibitor screening

On the basis of these promising data, we proceeded to investigate whether the m5C-probe strategy could provide direct visualisation of NSUN2 activity in living cells. Accordingly, HeLa cells, which constitutively express NSUN2, were transfected with either 9a (10 μM) or the locked-probe 10a (control) via Lipofectamine 2000, and then imaged using fluorescence microscopy (λexc 340–380 nm, λem 435–485 nm). Cells treated with probe 9a gave off bright blue fluorescence after 1 h; this is ~7-fold higher intensity compared with control (Figures 5A–C).

Further characterisation of the probe by time course fluorescence analysis showed that the probe remained strongly emissive in HeLa cell lysate for at least 3 h, suggesting good photostability (Supplementary Figure S11A). Moreover, there was no obvious increase in probe fluorescence in cell lysate which had been spiked with 1 mM simeprevir, suggesting no significant fluorescence arising from NSUN2-independent mechanisms (Supplementary Figure S11A). The lack of fluorescence increase further suggests that the probe is reasonably resistant to cellular degradation, which would otherwise release unstacked fluorophore. Consistent with this result, a repeat of the cell lysate experiment using locked-probe 10a gave very little or no increase in fluorescence in HeLa cell lysate, both in the presence and absence of 1 mM simeprevir (Supplementary Figure S11B). These data suggest that probe degradation by cellular nu...
Figure 5. The m\textsuperscript{5}C-probe provides real-time visualisation of cellular NSUN2 MTase activity. Live cell imaging of HeLa cells following treatment with m\textsuperscript{5}C-probe 9a (10 \textmu M) using (A) bright field and (B) fluorescence microscopy. (C) Fluorescence image of HeLa cells treated with locked-probe 10a (10 \textmu M) obtained from a separate experiment than the correlated images in panels A and B. It is displayed using the same brightness and contrast settings as that in panel B. Scale bar, 20 \textmu m. (D) Violin plots showing the flow cytometry analysis of HeLa cells transfected with probe 9a (green), and with NSUN2 knockdown (orange) or NSUN2 overexpression (red) (\textlambda\textsubscript{ex} 355 nm; \textlambda\textsubscript{em} 425–475 nm). The boxes in violin plots show the median (white dot), 25–75 percentile (black box) and 5–95 percentile values (black bar). (E) The fluorescence intensity of m\textsuperscript{5}C-probe changes with the expression levels of NSUN2 in HeLa cells, but is unresponsive towards NSUN6, TRDMT1 and DNMT3A. A minimum of 20 000 live cells were analysed for determination of mean fluorescence intensity (MFI). Data are expressed as mean ± SD of three biological replicates. **P < 0.01; n.s. = not significant.

Cleavases and other nucleic acid modifying enzymes do not interfere with the performance of the m\textsuperscript{5}C-probe 9a significantly, at least not within the time frame of our cell-based assay (typically 1 h). Furthermore, the probe is relatively non-cytotoxic, as evidenced by MTT toxicity assay and cell morphological assessment, where > 80% of the cells remained viable after being exposed to 40 \textmu M probe for 24 h (Supplementary Figure S11C).

Besides live-cell imaging, the m\textsuperscript{5}C-probe is also suitable for use in single-cell flow cytometry analysis (\textlambda\textsubscript{ex} 355 nm; \textlambda\textsubscript{em} 425–475 nm). In particular, NSUN2 knockdown in HeLa cells caused a marked reduction in probe fluorescence, whereas overexpression of NSUN2 led to significant increase in fluorescence signal (Figures 5D and E). Remarkably, the fluorescence intensity was unchanged with NSUN6, TRDMT1, and DNMT3A knockdown and overexpression, thus, consistent with our in vitro data, the probe is highly selective for NSUN2 over other m\textsuperscript{5}C MTases. Prior to this work, we are not aware of any assay method that is able to discriminate between DNA and RNA m\textsuperscript{5}C MTase activities in cells.

Finally, the developed m\textsuperscript{5}C-probe could also be used to directly measure NSUN2 inhibition in cells. Treatment of HeLa cells with increasing concentrations of known MTase inhibitors, sinefungin or S-adenosyl-L-homocysteine (SAH), for 30 min prior to the addition of probe 9a (10 \textmu M) led to a concentration-dependent decrease in fluorescence signal (Figures 6A–C). The determined IC\textsubscript{50} values of sinefungin (8.9 \textmu M) and SAH (3.1 \textmu M) against NSUN2 are comparable with their inhibitory activities against other RNA MTases (IC\textsubscript{50} values typically range between 0.1 \textmu M to 2 \textmu M (25,68); Figure 6D). Importantly, the negative controls, adenosine and homocysteine, showed poor inhibition towards NSUN2, as anticipated (IC\textsubscript{50}> 200 \textmu M and >400 \textmu M, respectively). Thus our assay is able to identify known NSUN2 inhibitors from amongst structurally-related analogues. We further performed Z’ factor analysis (55) to evaluate the reliabil-
Figure 6. Application of the m^5C-probe assay in cell-based screening of NSUN2 inhibitors. (A) IC_{50} values of non-specific MTase inhibitors against NSUN2. Results are from m^5C-probe assay. (B) Analysis of NSUN2 inhibition in live HeLa cells by real-time flow cytometry (λ_{ex} 355 nm; λ_{em} 425–475 nm). Cells were pre-incubated with sinefungin (0, 10 and 1000 μM; colour coded) for 30 min prior to treatment with m^5C-probe 9a (10 μM) or locked-probe 10a (10 μM; control). (C) The flow cytometry profile at 1 h showed a concentration-dependent decrease in mean fluorescence intensity (MFI) in cells treated with 9a. (D) Inhibition of NSUN2 by selected MTase inhibitors. The m^5C-probe assay could distinguish known inhibitors from structurally-related negative controls. Data are expressed as mean ± SD of three replicates. (E) Screening validation of the m^5C-probe assay in a 96-well format. The assay has a Z’ factor of 0.73, which demonstrates excellent reproducibility.

ity of our assay for high-throughput screening of inhibitor (for experimental details, see Supplementary Data). Our assay produces an average Z’ factor of 0.73 in a 96-well plate format, indicating that it has the required statistical reproducibility for cell-based screening of NSUN2 inhibitors (Figure 6E).

The m^5C-switchable probe strategy could be adapted for the study of DNA:m^5C MTase activity

In light of present finding that m^5C methylation can induce a spontaneous B-Z conformational change in CpG DNA duplex 7a (Figure 3B and Supplementary Figure S6), our m^5C-probe strategy may, in principle, also be adapted for the study of DNA:m^5C MTase activity. To examine this possibility, we prepared DNA probe 11a, which is an analogue of 7a containing a 5’-overhang pyrene deoxyribose (P_s) in the forward strand (Figure 7). The design of this probe exploits differences in the ability of P_s to participate in endstacking interactions in the B-DNA and Z-DNA forms. It is clear, from a number of studies, that whereas the B-DNA (diameter of helix ∼20 Å) is able to adopt a continuous base-stacking arrangement, the more compact Z-DNA (diameter of helix ∼18 Å) only permits discontinuous series of four-base stack (70). Because of these inherent differences in base-stacking pattern, the unpaired P_s will only be able to end-stack on adjacent G:C base pair when it is in a B-DNA environment, and not in the Z-DNA environment. Accordingly, conversion of probe 11a from B-DNA to Z-DNA form upon m^5C methylation is expected to cause a destacking of P_s fluorophores and, consequently, fluorescence activation.

Consistent with our probe detection strategy, m^5C methylation readily triggered a B-Z transformation in probe 11a, with concomitant increase in fluorescence emission (ΔΦ_F ~ 5.5-fold; λ_{ex} 340 nm; λ_{em} 380 nm and 400 nm; Figures 8A and B). Notably, there was no fluorescence activation in the absence of MgCl_2 where 11a did not undergo B-to-Z conversion, thus the observed fluorescence response was primarily mediated by a B-Z conversion of the probe (Figure 8B). Moreover, our thermodynamic analysis supports the notion that B-Z transition triggers a destacking of the...
CONCLUSION

Overall, we use a combination of NMR, CD and thermodynamic analyses to investigate the structural impact of m\(^5\)C methylation on RNA and DNA duplexes. Our results revealed that, although m\(^5\)C does not hinder canonical Watson–Crick base-pairing interactions, this modification can, in fact, elicits significant conformational change in certain RNA and DNA sequence contexts. We further demonstrated that m\(^5\)C produces distinctly different ‘structural signatures’ on RNA and DNA duplexes, including duplexes with identical CpG sequences. In the case of CpG RNA duplex (as exemplified by 1a), m\(^5\)C methylation induces a local distortion of the phosphate backbone and a C\(^3\)-endo to C\(^2\)-endo sugar pucker switch in the terminal residues. However, in the case of CpG DNA duplex (as exemplified by 7a), m\(^5\)C triggers a remarkable B-Z structural transformation. The significance of this result is unclear at present nevertheless, given that majority of the m\(^5\)C modifications in DNA and mRNA are distributed within the CpG consensus motif, sequence information alone is insufficient to discriminate m\(^5\)C marks on RNA and DNA CpG sites. Therefore, one question which arises from this finding is whether the duplex-remodelling property of m\(^5\)C might provide a mechanism for its specific recognition by RNA/DNA:m\(^5\)C binding proteins.

On the basis of this interesting finding, we further provided proof-of-principle that m\(^5\)C-induced conformational
change can be exploited for direct detection of m^5^C MTase activity in cells. This was demonstrated by the development of the first m^5^C-responsive probe 9a, which switches sugar pucker conformation spontaneously according to its m^5^C methylation status, hence useful for sensing RNA:m^5^C MTase activity.

The m^5^C-probe is simple, inexpensive and highly selective for NSUN2 over other RNA:m^5^C MTases (including NSUN3, NSUN5A, NSUN6 and TRDMT1) and DNA:m^5^C MTases (including DNMT1 and DNMT3A). Through the use of this probe, we achieved fluorescence imaging and real-time flow cytometry analysis of NSUN2 activity in live HeLa cells. We further demonstrated the utility of the probe in cell-based screening of NSUN2 inhibitors. Prior to this study, we are not aware of any methods that permit direct sensing of RNA:m^5^C MTase activity in cells. There are also no reports of assays that selectively target NSUN2 over other RNA/DNA:m^5^C MTases. The discovery of such highly selective probes is rarely achieved and may prove valuable in advancing our knowledge of NSUN2 in m^5^C-regulated processes. Importantly, our m^5^C-probe approach could also be adapted for the analysis of DNA:m^5^C MTase activity. This was demonstrated by the development of DNA:m^5^C-probe 11a, which is useful for in vitro screening of DNMT3A inhibitors.

We appreciate that m^5^C-induced terminal sugar pucker switch is likely an interesting anomaly that is unique to alternating CpG RNA duplexes, since this phenomenon was not observed in majority of other sequences investigated in this study. Therefore, one limitation of the proposed approach is that it is not applicable to RNA:m^5^C methyltransferases that do not methylate CpG substrates. Nevertheless, given the importance of NSUN2, both as a key regulator of m^5^C marks in mRNA and a potential therapeutic target, we envisaged that the proposed NSUN2 assay strategy would be of broad scientific interest.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We would like to thank the Chemical, Molecular and Materials Analysis Centre at the National University of Singapore for their support in NMR measurements and helpful discussions.

FUNDING
Singapore Ministry of Health’s National Medical Research Council [NMRC/BNIG/2008/2013]; Singapore Ministry of Education [AcRF Tier 1 Grant R148-000-231-114 and R148-000-238-114]. Funding for open access charge: Singapore Ministry of Education [AcRF Tier 1 Grant R148-000-238-114].

Conflict of interest statement. None declared.
REFERENCES

d(m5C-G-C-G-m5C-G): mechanism of the Z-B-coil transitions. Biochemistry, 23, 1362–1371.


