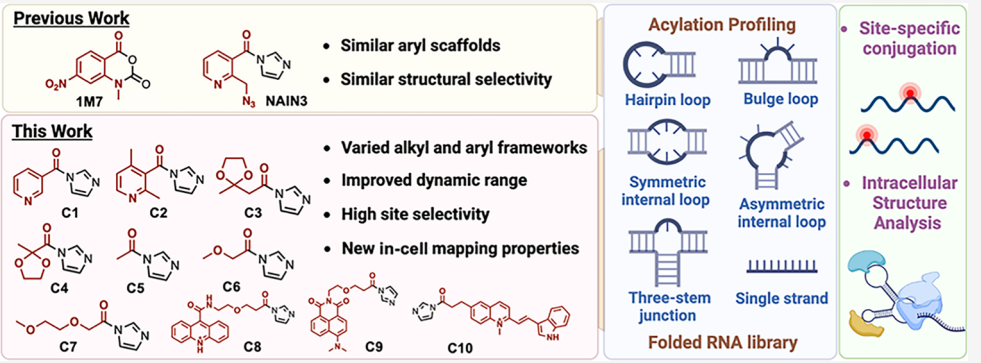
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Diverse Reagent Scaffolds Provide Differential Selectivity of 2′-OH Acylation in RNA

Lu Xiao, Linglan Fang, Sayantan Chatterjee, and Eric T. Kool*

Cite This: J. Am. Chem. Soc. 2023, 145, 143−151
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: RNA 2′-OH acylation is widely used both for mapping structure and for conjugating RNA, generally relying on selective reactions with unpaired nucleotides over paired ones. Common reagents for this acylation have been chiefly restricted to two similar aryl scaffolds, leaving open the question of how more broadly varied reagent structure might affect selectivity. Here, we prepared a set of 10 structurally diverse acylimidazole reagents and employed deep sequencing to profile their reactivity and selectivity in an RNA library of systematically varied structure. We show that structure-directed reactivity profiles vary significantly with the reagent scaffold, and we document new acylating agents that have altered selectivity profiles, including reagents that show elevated selectivity within loops, as well as compounds with reduced off-target reactivity in loop closing base pairs. Interestingly, we also show that the simplest reagent (acetylimidazole) is cell permeable and is small enough to map RNA structure in the presence of protein contacts that block other reagents. Finally, we describe reagents that show elevated selectivity within small loops, with applications in site-selective labeling. The results provide new tools for improved conjugation and mapping of RNA.

INTRODUCTION

Acylation of RNA is a rapidly growing field for studying RNA-related cellular processes[1] and for development of RNA-associated biotechniques.[2] It has been widely used both in mapping folded RNA structures and interactions,[3,4] as well as in labeling, caging, and general functionalization of RNAs.[5,6] The primary target of this acylation is the 2′-OH group, which occurs at essentially every position of an RNA strand, thus enabling the probing of folding along virtually the entire sequence. In addition, with proper reagent design, high-yield conjugation of transcripts can be performed in either the general sense (via stochastic polyacylation)[6,7] or at specific sites.[8,9]

With the existing RNA-acylating reagents studied to date, it is generally understood that reactivity of the 2′-OH group to acylation is greater at unpaired nucleotides than it is in the double-stranded structure as the right-handed A-form helix shields 2′-OH groups sterically from access by reagents.[10] The differences in trace-level acylation of unpaired versus paired nucleotides are employed widely in mapping of the folded secondary structure, by assigning nucleotides of higher reactivity to non-paired regions, and segments of low reactivity to paired duplexes via selective 2′ hydroxyl acylation analyzed by the primer extension (SHAPE) methodology originated by Weeks.[3] A number of studies have further been carried out to better understand, from the RNA standpoint, how different folded domains compare in their reactivity to an acylating reagent for accurate structure prediction.[11,12] In addition, this selectivity for unpaired groups has recently been exploited in high-yield local RNA conjugation reactions, by use of complementary DNAs that induce reactive loops while protecting other nearby RNA sequences.[8,9]

Overall, both chief applications of RNA acylation—mapping and conjugation—rely on the selectivity of an acylating reagent

Received: August 24, 2022
Published: December 21, 2022

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151

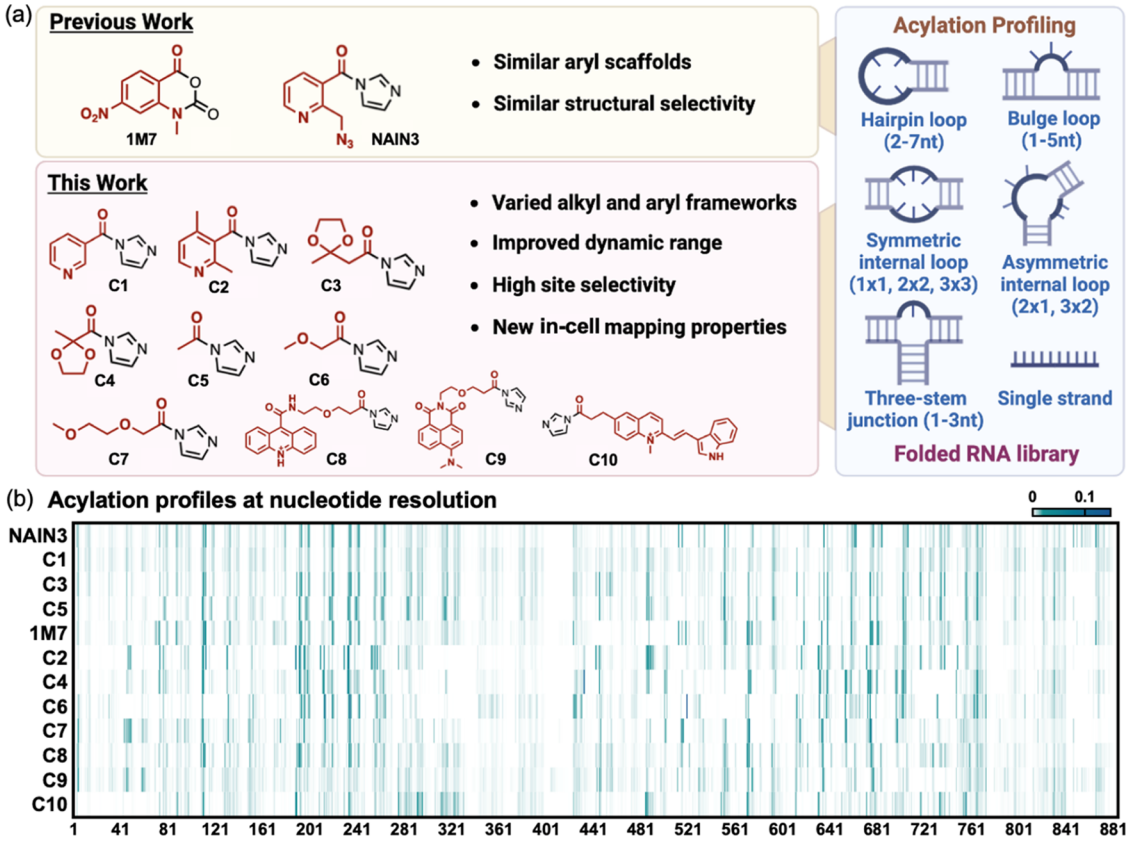


Figure 1. Reagent structures and differential reactivity patterns. (a) Schematic of acylation profiling of structure-varied reagents with a folded RNA library, revealing scaffolds with new acylation properties. Reagent structures (C1−C10, NAIN3, and 1M7) and RNA elements (hairpin, bulge loops, internal loops, three-way junctions, and single-stranded RNAs) studied in this work are shown. (b) Acylation profiles of structurally different reagents for each nucleotide in the folded RNA library (884 nucleotide positions in all), showing variations in positional reactivity. Band intensities represent the reactivity of the reagent at each nucleotide.

generically for unpaired structures over base-paired ones. However, only a few acylating probes have been explored quantitatively for their structure selectivity till date. Indeed, virtually all RNA structural mapping studies have relied on only two aryl scaffolds, namely, the isatoic anhydrides (1M7 and NAIM) and nicotinic acylimidazoles (NAI and NAIN3), and the two classes show similar (although not identical) selectivity,[12−15] likely due to their comparable sizes and aryl frameworks (Figure 1a). In addition, these structures have reactivity features that could potentially be improved upon; for example, current reagents show a bias for reaction at 5′-ends of loops over the 3′-region, and they often give elevated off-target reaction at base-paired positions (closing base pairs) adjacent to loops.[12] Arguably, for structure mapping, it would be helpful to identify reagents with better loop/stem dynamic range, less positional bias in loops, and low off-target reactivity at base paired nucleotides. Ideally, reagents would be cell permeable as well, for reacting with RNAs in cells. One recent example is 2-aminopyridine-3-carboxylic acid imidazolide (2A3), which shows increased reactivity and sensitivity, especially in bacterial RNAs.[15]

Although general selectivity data exist for only two classes of the reagent structure, RNA acylating activity has recently been reported to occur more broadly with a wider range of structurally varied reagents having both aryl and alkyl scaffolds.[13,16] Till date, very little is known about how reagent structure and reactivity affect acylation selectivity among different RNA structural elements. Given that a reagent's reactive carbonyl group must closely approach the bulky RNA substrate for successful reaction, it seems likely that reagent shape, its steric bulk, its reactivity, and other physicochemical features could have substantial influences on selectivity in acylating RNAs. It further seems probable that different RNA secondary structures also play a direct role in this selectivity as the steric and dynamic accessibility of a given reacting 2′-OH varies widely depending on its local structural constraints.[10] Investigating such reagent/RNA interactions is significant because the results may identify specific reagent/structure combinations that give unusual selectivity and reactivity, which can have practical applications in developing new methods for conjugating and labeling RNAs. Moreover, it can define and rank reagent selectivity for basic mapping applications, quantifying dynamic ranges, and assisting users in reagent choice for improving the analysis of RNA structure and interactions. Finally, it may help elucidate structural and reactivity features in reagents that govern their selectivity, pointing to future reagent design strategies for both conjugation and mapping purposes.

To address these issues, here we have employed deep sequencing to quantitatively profile a set of structurally diverse acylimidazole agents (compounds C1−C10, in comparison to NAIN3 and 1M7, Figure 1a) in reaction with a systematically varied library of RNA secondary structure elements (Figure 1 and S1). Most of the reagents have not been described before, and they vary substantially in steric bulk near the reactive carbonyl. The results identify new reagents with superior

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151

selectivity profiles. They also elucidate differences and similarities among the range of reagent structures, suggesting structural elements that can be useful in RNA labeling and mapping. We further identify new compounds that have elevated selectivity for site-directed conjugation and cell-permeable reagents that have altered selectivity profiles for structure mapping in solution and in cells.

RESULTS AND DISCUSSION

Acylation Profiling with a Systematically Varied RNA Library. To study the role of the reagent structure on reactivity and selectivity with RNAs, we prepared a panel of 10 acylimidazole reagents (C1−C10) with diversified substitution near the reactive carbonyl group (Figure 1a). Two common structure-mapping reagents, 1M7[17] and NAIN3,[18] were added as benchmarks. The acylating reagents vary substantially in their physicochemical features, including alkyl versus aryl scaffolds as well as increasing steric crowding near the reactive acyl group. We also included compounds (C8−C10) that contain polycyclic moieties reported to have affinity for RNA.[19,20] All compounds were shown independently by mass spectrometry to selectively react with an RNA oligonucleotide over a DNA oligomer having the same sequence, thus indicating reaction at 2′-OH groups rather than exocyclic amines (see Figure S2).

We investigated the reaction of these reagents with a sequence-generic RNA library that has been employed recently for evaluation of loop reactivities with NAIN3 and 1M7.[12] The library contains 43 secondary structures and sequence elements with homogeneous loops (polyadenosine or polyuridine sequence) including hairpins, bulge loops, internal loops, and three-way junctions in varied loop sizes, as well as single strands (Figure 1a). The handling and analysis of reactions between each reagent and the generic RNA library followed previously established methods[12] (Figure S1). Acylation was analyzed at each nucleotide by quantifying reverse transcriptase (RT) stops after library preparation and deep sequencing. Average number of reads for each RNA/reagent combination was ∼7 million, providing ample data for analysis. To achieve single-hit kinetics for each reagent with the RNA library, we first optimized the reaction conditions of compounds C1−C10 with a short model RNA (Figure S3a) to adjust for differential reactivity. A ranking of general reactivity of the reagents is provided in Figure S3b. Under the optimized conditions, we then carried out the acylation profiling (see positional map comparisons in Figure 1b).

Given the general assumption that all reagents prefer unpaired over paired nucleotides, surprising differences in the reactivity patterns are seen, sometimes even for subtle structural modifications. For example, pyridyl reagent C1 is structurally very closely related to NAIN3 (which is different only by an azidomethyl group, Figure 1a), and although the patterns are similar, notable reactivity differences are seen at some RNA positions. However, the addition of two ortho methyl groups to C1, giving the sterically more bulky C2, results in one of the most divergent reactivity patterns in the entire data set. Pairwise comparisons of the patterns reveal quantitative differences and similarities among the reagent set (Figure S4). For example, the RNA structure-directed reactivity patterns of C1 vs. NAIN3 or C1 vs. C3 showed relatively high Pearson correlations (r = 0.71 and 0.80, respectively), while NAIN3 versus C6 or C7 revealed dramatically different reactivity patterns, with Pearson correlation values (r) of 0.22 and 0.20, respectively. The significantly distinct acylation profiles encouraged us to further interrogate how reagent structure affects the reactivity and selectivity among varied RNA-folded motifs.

Variations in the Loop/Stem Dynamic Range. For the previously known RNA acylating reagents, the literature has established a general reaction preference for 2′-OH groups of unpaired nucleotides over those in base-paired positions.[10] Given this preference, one can provide a measure of selectivity by quantifying the dynamic range of reactivity, averaging reactivity with loops over duplex stems (L/S) (Figure 2). We find that dynamic range varies from less than L/S = 2 to over 5 (Figure 2b). The lowest dynamic range was seen for compound 10, containing a reported RNA-binding motif,[20] which might be explained by ligand-directed enhancement of the reaction in the duplex regions of the RNAs. The top four reagents with the highest loop-stem selectivity (L/S > 4) are NAIN3, C1, C5, and 1M7. Gratifyingly, the most common employed structure mapping reagents rank high. Our results with C1 are in general agreement with a recent study that found favorable loop/stem selectivity for the compound in ribosomal RNAs.[15] The new results indicate that new compound C5 could potentially also be highly effective as the SHAPE reagent for RNA structure mapping (see below).

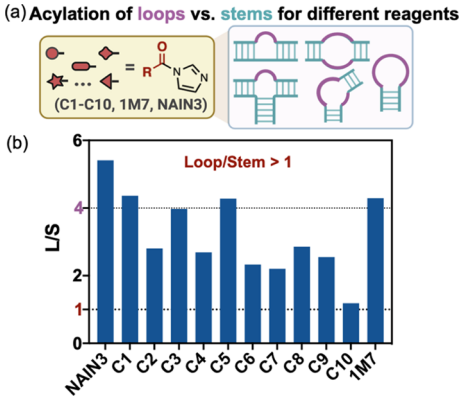


Figure 2. Reagent variations in dynamic range as measured by averaged loop/stem reactivity ratios. (a) Schematic of different reagents carrying out acylation in loops (purple) and stems (green) in library RNAs. (b) Bar graph of RNA loop-stem reactivity ratios (L/S) of acylating reagents.

Acylation Features of Reagents among Unpaired RNA Motifs. Next, we explored how reagent structures affect the acylation patterns among varied loop motifs (Figure 3a), with the aims of (a) identifying reagents that may avoid some existing loop biases and (b) to identify loop/reagent combinations that might be highly selective for applications in site-specific conjugation. Low positional selectivity would be desirable for mapping, where all unpaired nucleotides might show similar levels of the reaction. Conversely, high positional selectivity would be a favorable property for site-specific conjugation.

We first calculated the average reactivity of reagents for each type of unpaired RNA motif (hairpin, bulge loop, internal loop, three-way junction, and single strand) (Figure 3b). We found that loops are far more reactive than single strands for nearly all acylating reagents irrespective of structure, which is consistent with previous observations with mapping reagents.[12] However, the acylation selectivity among RNA loop motifs

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151

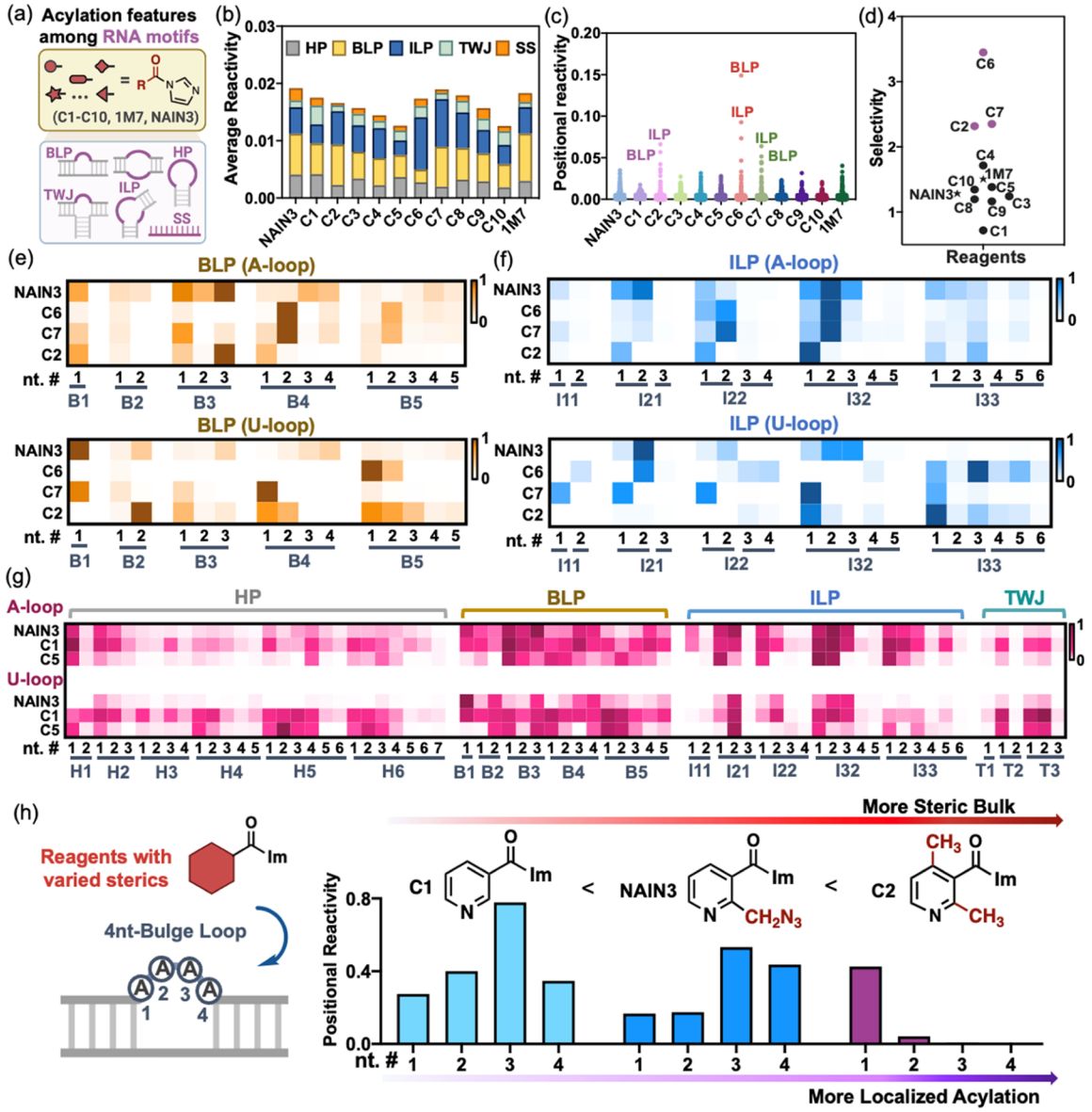


Figure 3. Localized acylation features of different reagents among unpaired RNA motifs including hairpin (HP), bulge loop (BLP), internal loop (ILP), three-way junction (TWJ), and single strands (SS). (a) Schematic of RNA motifs. (b) Average reactivity comparison of RNA motifs across different acylating reagents (C1−C10, 1M7, and NAIN3). (c) Waterfall graph of positional reactivity of all library RNA loops for each acylating reagent; note labeled individual cases with exceptional positional selectivity; (d) acylation selectivity profile of reagents among RNA loop nucleotides. Selectivity was evaluated by the relative standard deviation (RSD) of loop positional reactivity of each acylating reagent. Reagents with larger RSD have higher localization of reactivity. (e) Positional reactivity comparison of selected reagents (C6, C7, C2, and NAIN3) showing elevated selectivity in all bulge loops (B1−B5: 1−5 nt bulge loops). (f) Positional reactivity comparison of NAIN3 with reagents (C6, C7, and C2) revealing elevated selectivity in all internal loops (I11, I21, I22, I32, I33: 1 × 1 nt, 2 × 1 nt, 2 × 2 nt, 3 × 2 nt, 3 × 3 nt internal loop). (g) Positional reactivity comparison of NAIN3 with reagents C1 and C5, showing more balanced acylation in certain RNA loops (H2−H7: 2−7 nt hairpin; T1-T3: 1−3 nt three-way junction; B1−B5: 1−5 nt bulge loop; and I11−I33: 1 × 1−3 × 3 nt internal loop). (h) Examples of steric effects modulating acylation selectivity. Pyridyl reagents with increased steric bulk show more localized acylation within a tetranucleotide bulge loop.

differed with the structures of reagents. We examined the positional reactivity of each reagent in all loop nucleotides of the folded RNAs, evaluated by variance in the positional reactivities (Figure 3c,d and detailed methods in Supporting Information). We found that the reagents C6, C7, and C2 exhibited very high localized preferences among loop nucleotides compared with others including mapping reagents NAIN3 and 1M7 (Figure 3d). In addition, the nucleotides with the highest reactivity are all located in bulge loops (BLP) and internal loops (ILP) as compared with hairpin loops (HP) or three-way junctions (TWJ) for all three reagents (Figure 3c). Examining positional reactivity among these most highly reactive loops (Figure 3e,f), we find that acylation by C6, C7, and C2 was localized to certain nucleotides (chiefly the first and the second nucleotide) in BLP and ILP cases. In addition, reagents preferred to acylate the over-numbered side of asymmetric internal loops. Taken together, the results suggest that the combination of a small bulge loop or internal loop with one of the new highly selective reagents (C6, C7, and C2) could potentially yield improved selectivity for site-specific RNA conjugation (see further investigations below).

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151

Having quantified localized nucleotide selectivity with the diversified reagents, we then turned our attention to the converse possibility of more balanced acylation along loop nucleotides. To this end, we examined the positional reactivity of reagents C1 and C5, which displayed the lowest selectivity and the smallest size, respectively (Figure 3g). Overall, C1 and C5 revealed balanced acylation along the loops that was equal to or better than that of the SHAPE reagent NAIN3 (see e.g., the loops of H6, T3, B5, and I33 in Figure 3g). This finding, along with the high loop-stem reactivity ratio for C1 and C5 analyzed above, suggests that these two reagents could be candidates for improved SHAPE mapping of folded RNAs (Figure S5) (see additional experiments below).

To survey the data as a whole, we conducted PCA analysis to study the relationship of physicochemical characteristics for all reagents with their activity features (Figure S6, Table S3). Interestingly, we found some observations in common. First, the reagents C8, C9, and C10 are grouped well away from the others, and given their unfavorable selectivity properties, they were ruled out for further analysis (Figure S6a). Second, the reagents with the highest loop-stem reactivity ratio (C1, C5, NAIN3, and 1M7) generally contain fewer hydrogen bond acceptors (Figure S6b). Third, the reagents (C3, C4, C5, and C6) showing higher reactivity in adenosine (A) loops than uridine (U) loops have smaller geometrical sizes (Figure S6c), potentially less blocked by the larger purine bases. Finally, we observe that the reagents with highest positional selectivity are either highly hydrophilic (C6 and C7) (Figure S6d) or sterically bulky (C2) (Figure S7). By further comparing the loop positional reactivity of the reagents with different substitution near the carbonyl, we observed more localized acylation with reagents possessing greater steric bulk (Figure 3h). These observations may offer some hints for future reagent structure−activity relationships, although more studies with greater numbers of compounds are needed for verification.

Improving Site-Specific Conjugation with Selective Reagents. To further test the above discovery of high positional selectivity, we evaluated the reagents C6, C7, and C2 with our recently developed method (RAIL, RNA Acylation at Induced Loops)[8] to look for possible improved site-specific conjugation. The RAIL method employs a complementary DNA to induce a reactive bulge loop in an RNA of interest and protect the remaining residues for localized RNA acylation (Figure S8). In the early description of the method, we performed high-yield site-selective modification of RNA with the NAIN3 reagent; however, significant off-target acylation occurred at the adjacent bases next to the intended looped-out nucleotide (see e.g. Figure 4b, last two lanes). To test whether improvements to yield single-nucleotide selectivity are possible, we examined the new combinations of bulge loop/internal loop and highly selective reagents (C6, C7, and C2) observed above. We conducted the experiments following previous procedures (Figure S8). Briefly, RNA-DNA bulge loop or internal loop complexes were assembled by addition of an inducer DNA, and then, the loop complexes reacted with reagents, followed by DNA removal with DNase I and click chemistry for further labeling. We tested 1 and 3 nt A/U bulge loops (B1A, B1U, B3A, and B3U) and 3 × 2 A/U internal loops (I32A and I32U) with C6, C7, and C2 reagents.

After optimization of conditions (Figure S8), we found that an induced A-loop in RNA is more locally selective than a U-loop for acylation, and new methoxyethoxyacetyl reagent C7 performed the best among the reagents (Figure 4a). PAGE gel analysis (Figure 4b) with reverse transcriptase (RT) primer extension for the 1 nt A-bulge samples (B1A-C7) revealed that the band only appeared at C, which is the position corresponding to RT stopping immediately before the acylated nucleotide. No off-target acylation was observed in the adjacent nucleotides for all conditions with increased concentrations and reaction times. This reveals that the acylation occurs virtually exclusively at the intended bulged A residue. Analysis of single-adduct acylation with 3 nt A-bulge loop RNA (B3A-C7 vs B3A-NAIN3) also exhibited higher selectivity for C7 over NAIN3 (Figure S9a,b). The localization phenomenon is also striking in the induced 3 × 2 internal loop cases (I32A-C7 vs. I32A-NAIN3); the acylation was completely selective to the center nucleotide of the AAA triloop for reagent C7, while the reaction was scattered over all three nucleotides for NAIN3. To enable further functionalization, we synthesized a terminal azide-substituted analogue of C7 (C7N3, see Supporting Information for details), which showed the same positional selectivity as C7 (Figure 4a,c). MALDI-TOF mass spectra independently showed a single adduct for C7N3 in the I32A-type induced loop, while up to four adducts for NAIN3 were seen (Figure S10). We improved the acylation yield of C7N3 by reducing the temperature and

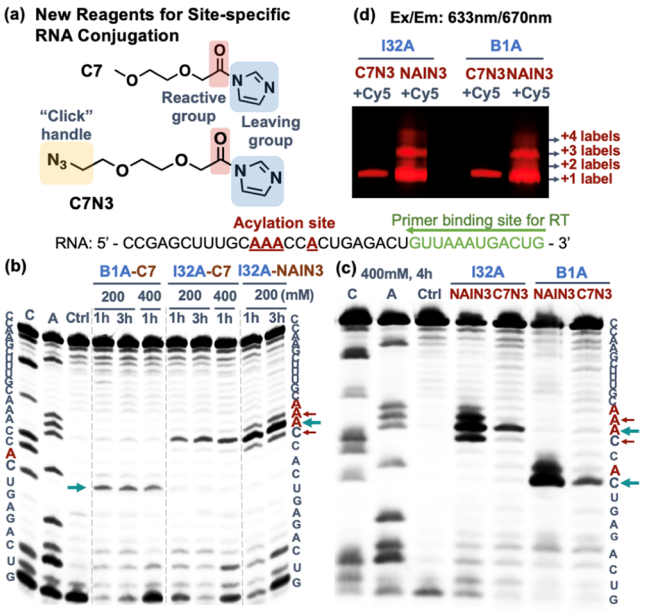


Figure 4. Site-specific RNA labeling in vitro by acylation in induced adenosine loops (RAIL+), employing new reagents with improved specificity. (a) Schematic of new reagents (C7 and C7N3) with elevated selectivity for improved conjugation at a specific loop nucleotide via "RAIL+" method. RNA bulge loops or internal loops were induced by a DNA oligonucleotide through hybridization. (b,c) Gel electrophoretic analysis of reverse transcriptase (RT) stops for RNA samples reacting with C7, C7N3, and NAIN3 under different conditions, revealing presence and absence of reactivity at positions adjacent to the intended site. RNA was constructed to either 1 nt A bulge (B1A) or 3 × 2 nt A internal loop (I32A) (Figure S8). The sequences and sites of the bands in the gel are shown. Note that RT stops (and corresponding bands) occur at the nucleotide immediately 3′ to the acylated residue. (d) Fluorescent image of gel analysis of Cy5-labeled RNAs via RAIL + method with NAIN3 and C7N3. RAIL + with C7N3 show significantly higher site-specificity compared with the prior acylating reagent, NAIN3.

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151

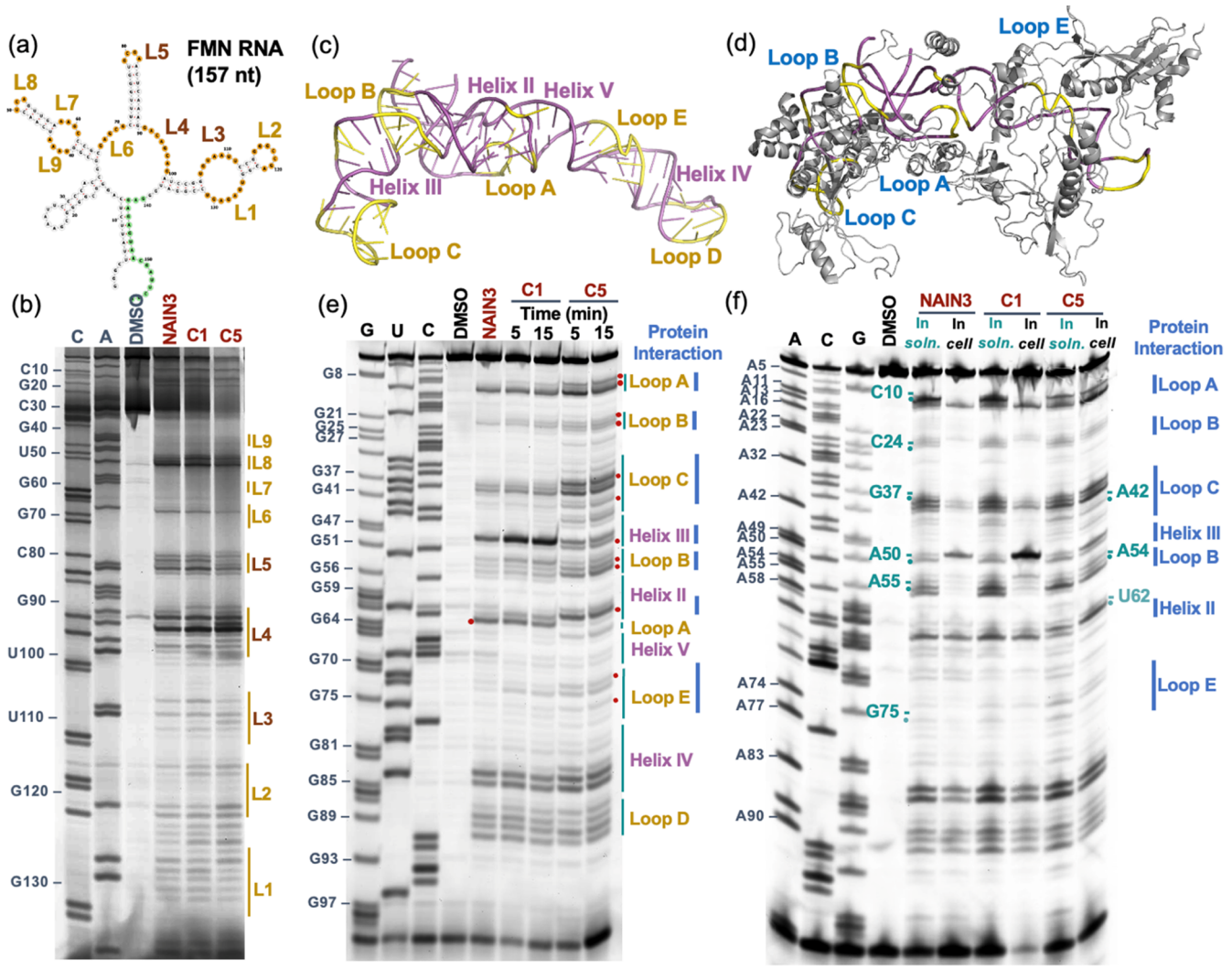


Figure 5. In-solution and in-cell RNA structure mapping with new acylating reagents. (a) Secondary structure of flavin mononucleotide (FMN) riboswitch predicted by the RNAFOLD folding algorithm. Primer binding site (light green) for reverse transcriptase primer extension and detected loop region (orange: L1-L9) are shown. (b) In-solution SHAPE probing of FMN RNA by our acylating reagents (200 mM C1, 100 mM C5) compared with a common mapping reagent (200 mM NAIN3) at 37 °C for 10 min. Sequencing lanes labeled with C and A are used to calibrate the acylation position at nucleotide resolution. (c) Structure of human 5S rRNA (PDB 4UG0). The loop regions (yellow) and helix regions (magenta) of 5S rRNA are shown. (d) Structure of human 5S rRNA with interacting ribosomal proteins (PDB 4UG0). Gray: interacting proteins; yellow: loops in 5S rRNA; and magenta: helices in 5S rRNA. (e) In-cell SHAPE probing of human 5S rRNA with 100 mM C1 and C5 for 5−15 min and 100 mM NAIN3 for 5 min in HEK293 cells. The loop regions (yellow), helix regions (magenta), and protein interaction sites (blue) are denoted in the gel corresponding with the 5S rRNA structure. Significant differences of probing between C5 and C1, NAIN3 are noted in the gel (red dots). Sequencing lanes are labeled with G, U, and C for each nucleotide identification. (f) In-solution and in-cell SHAPE probing comparison of human 5S rRNA. Nucleotides showing significant differences between solution and cellular experiments are labeled in green. Sequencing lanes are labeled with A, C, and G for nucleotide identification.

dispensing the reagent in portions (Figure S9) and further confirmed the site selectivity of C7N3 by comparing the RT-stops of acylated RNAs with NAIN3 under similar acylation yield (Figure S9c). To further test utility for in vitro conjugation of RNAs, we carried out a post-acylation strain-promoted click reaction to label the RNA with Cy5 fluorophore at a single position. Fluorescence images of the Cy5-labeled RNA revealed a single band in the gel (Figure 4d). These results demonstrated that C7N3 can be a useful acylating reagent for local RNA labeling with single-nucleotide specificity.

Application of New Acylating Reagents for RNA Structure Mapping in Solution and in Cells. The two types of commonly used RNA mapping reagents are based on isatoic anhydride and acylimidazole scaffolds (e.g., 1M7 and NAIN3, respectively).[17,18] Our above data suggested that compounds C1 and C5 have potentially reduced selectivity biases than the above reagents, while preserving high dynamic range. Since C5 has a smaller size, we further confirmed the efficiency of generating RT stops by comparison with NAIN3 in titration experiments (Figure S11). To verify the ability of C1 and C5 in profiling RNA structure, we first probed a 157 nt flavin mononucleotide (FMN) riboswitch RNA[21] in solution with C1, C5, and NAIN3 (Figure 5a,b). All three reagents identified the expected loop regions correctly. We found that the reactivity profile of C1 was almost identical with that of NAIN3; however, C5 showed more bands in loops, disclosing

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151

more loop information as compared with NAIN3 [see, e.g., L3, L4, and L5 (Figure 5b)].

Encouraged by the in-solution mapping experiments, we next evaluated whether C1 and C5 would be amenable to probe RNA structures in cells, which requires high cell permeability to achieve robust signals in a short time.[18] We chose human 5S ribosomal RNA (rRNA) as a target since the structure is well studied. Human 5S rRNA consists of five loop regions (two hairpin loops C and D, two internal loops B and E, and one junction loop A as a hinge) and five helix regions (Helix I–V), with two conserved adenosine bulges (A49 and A50) in helix III and extrahelical cytosine (C63) flipped out of the helix II in two different configurations[22] (Figure 5c). HEK293 cells were treated with 100 mM C1 or C5 for 5 min or 15 min, similar to the published use of 100 mM NAIN3 for 5–10 min, which has been used to map intracellular 5S rRNA structures previously.[18] We found that we could achieve 5S rRNA modification in 5 min for all the reagents, with a suitable signal-to-noise ratio at accessible regions, suggesting that, in addition to NAIN3, C1 and C5 are viable as cell-permeable probes of RNA structures (Figure 5e). The SHAPE profile of pyridyl compound C1 in cells is overall similar to that of NAIN3, and comparison of its in-cell and in-solution profiles reveals protected positions in cells where associated ribosomal proteins block reagent access, as was previously seen for NAIN3[18] (Figure 5f).

Interestingly, the acylation profile of reagent C5 (acetylimidazole, the smallest compound in the study) with 5S rRNA in cells is quite distinct from those of C1 and NAIN3, disclosing more nucleotides in protein-contacting regions, such as loops A, B, and C (Figure 5e, highlighted with red dots). Comparing our SHAPE data to the crystal structure of the human 80S ribosome, which contains the 5S rRNA and its interacting proteins (PDB 4UG0) (Figure 5c,d), we find that the additional sites revealed by C5 relative to C1 and NAIN3 are all interacting with ribosomal proteins in cells. We hypothesized that due to its small size, C5 may have enhanced accessibility to probe positions that are near in space to protein surfaces. To explore this further, we compared the in-cell data with in-solution SHAPE profiles of 5S rRNA to exclude protein interaction effects (Figure 5f). Strikingly, the in-cell SHAPE reactivity profile of C5 is almost identical with its in-solution data and is similar to the other in-solution profiles generated by NAIN3 and C1, adding support to the hypothesis that C5 can profile RNA structures with greatly reduced influence by associated proteins (Figures 5f and S12b). We further examined published ribosome crystal structures to examine specific nucleotides (Fig. 5f, labeled in green) that can only be detected in solution with the cell-permeable mapping reagent NAIN3 (Figure S13). We found that the 2′-OH groups at these positions are all facing protein surfaces nearby, providing limited spaces and/or dynamic accessibility for acylation by C5 while apparently excluding reagents with larger size. Additional experiments were also carried out probing human 18S rRNA in HEK293 cells, showing similar differences in accessibility by the smaller reagent (see additional analysis and experiments in Figures S14 and S15).

In sum, the reagents C1 and C5 are both applicable for RNA structure mapping both in solution and in cells. In particular, C5 may have added benefit over the commonly used cell-permeable reagent NAIN3 by showing less bias along loop positions. More strikingly, the acetylimidazole reagent (C5), at least in this limited initial test, exhibits an extraordinary ability to map in-cell RNA structures that are associated with proteins that block access by other larger reagents. We conjecture that it may be broadly useful in mapping intracellular RNA structure; more studies are required to confirm this. Acetylimidazole was early on studied by Sutherland as a reagent for protecting a terminal 2′-OH group in solution with oligoribonucleotides.[23] More recently, we showed that it can react with RNA 2′-OH groups more generally in superstoichiometric yields[25] and also showed that it could be used for SHAPE mapping of FMN riboswitch RNA. In addition, Weeks and co-workers recently showed that acetylimidazole can be employed for in-solution SHAPE analysis by nanopore sequencing.[24] Overall, acetylimidazole appears to have a number of promising properties for mapping RNA structure based on its cell permeability and small size.

CONCLUSIONS

Our profiling data reveal that reagent structure can indeed have substantial effects on selectivity in acylating RNA 2′-OH groups. Different RNA secondary structures can also play a critical role in this selectivity. We find that all reagents in our study react preferentially with unpaired loops over base-paired stems in folded RNAs; however, selectivity varies greatly with the reagent structure, and we find high dynamic ranges for our reagents C1 and C5, which are as high or higher than those of existing reagents NAIN3 and 1M7. This property may prove useful in future structure mapping applications, aided by the finding that they show lower positional bias among loop nucleotides as compared with existing mapping reagents. Moreover, C1 and C5 are cell permeable and react with RNAs in intact mammalian cells within 5 min, providing useful new tools for in-cell analysis of RNA structures and interactions. Recently, the Incarnato group has introduced a 2-aminopyridyl imidazolide reagent (2A3), as a probe for in-cell RNA structure analysis.[15] This reagent provided improved loop nucleotide information in bacterial ribosomal RNA compared with the compound lacking the amino group (here denoted C1). Thus, both that earlier study and our current results show that small structural changes near the reactive carbonyl group can have large effects on acylation selectivity profiles.

Our new data reveal that new alkoxy reagents C6 and C7 and sterically bulky reagent C2 exhibit high positional loop nucleotide selectivity (chiefly at first and second nucleotides) particularly in bulge loops and internal loops. This is in contrast to C1 and C5, which show more balanced reactivity along loops. We conclude that the physicochemical characteristics of reagent structures (e.g., steric bulk and reactivity) and the local structural constraints of RNA together govern the access of the reagent carbonyl group to specific 2′-OH groups, thus defining the selectivity of acylation. Understanding this reagent/RNA interaction will provide new insights into future efforts to leverage reagent designs for both conjugation and structure mapping.

Acylation positional selectivity is critical for site-specific conjugation of RNAs. We have applied the highly selective combination of reagent C7 with bulge loops or asymmetric internal loops to conjugate RNAs in vitro at specific nucleotide positions, with significant reduction of unwanted reactions at neighboring nucleotides. Employing our recent DNA-directed loop induction strategy (RAIL),[8] we achieved precise RNA labeling with azido-C7 (C7N3) by inducing a single nucleotide bulge loop (B1A) or an internal loop (I32A) in RNA strands. This enabled post-synthesis placement of a fluorescent label in

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143–151

the RNA at a single nucleotide position. These data suggest an improved RNA functionalization method (RAIL+) with high site specificity. Future studies will explore this methodology with larger biologically relevant RNA targets.

Our initial observation that C5 (acetylimidazole) may have the ability to map RNA structure underlying associated proteins is notable and suggests that the reagent is potentially applicable for broader intracellular analysis. We report here that it is cell permeable and reacts rapidly with intracellular RNAs. Although more work is needed to confirm generality, our early data suggest that the small size of the reagent may enable it to access 2′-OH groups in sterically crowded spaces as compared to larger cell-permeable reagents. Overall, our findings expand the choice of mapping reagents and can assist users for improved RNA structure and interaction analysis.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.2c09040.

Oligonucleotides used in this work, reagents, materials, supplementary data, strategy for probing folded RNA libraries with acylation and sequencing, experimental details, and NMR spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Eric T. Kool − *Department of Chemistry and Sarafan ChEM-H, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0002-7310-2935; Email: kool@stanford.edu

Authors

Lu Xiao − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0003-1934-4711

Linglan Fang − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0003-2637-090X

Sayantan Chatterjee − *Department of Chemistry, Stanford University, Stanford, California 94305, United States;* orcid.org/0000-0001-9265-595X

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.2c09040

Author Contributions

All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the U.S. National Institutes of Health (GM130704 and GM145357) for support.

■ REFERENCES

(1) (a) Ganser, L. R.; Kelly, M. L.; Herschlag, D.; Al-Hashimi, H. M. The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 474−489. (b) Sharp, P. A. The Centrality of RNA. *Cell* **2009**, *136*, 577−580.

(2) Kole, R.; Krainer, A. R.; Altman, S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discovery* **2012**, *11*, 125−140.

(3) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. RNA Structure Analysis at Single Nucleotide Resolution by Selective 2′-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* **2005**, *127*, 4223−4231.

(4) (a) Spitale, R. C.; Flynn, R. A.; Zhang, Q. C.; Crisalli, P.; Lee, B.; Jung, J.-W.; Kuchelmeister, H. Y.; Batista, P. J.; Torre, E. A.; Kool, E. T.; et al. Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **2015**, *519*, 486−490. (b) McGinnis, L.; Liu, Q.; Lavender, A.; Devaraj, A.; McClory, P.; Fredrick, K.; Weeks, M. In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 2425−2430. (c) Smola, M. J.; Calabrese, J. M.; Weeks, K. M. Detection of RNA−Protein Interactions in Living Cells with SHAPE. *Biochemistry* **2015**, *54*, 6867−6875.

(5) (a) Velema, W. A.; Kool, E. T. The chemistry and applications of RNA 2′-OH acylation. *Nat. Rev. Chem.* **2020**, *4*, 22−37. (b) Velema, W. A.; Kietrys, A. M.; Kool, E. T. RNA Control by Photoreversible Acylation. *J. Am. Chem. Soc.* **2018**, *140*, 3491−3495. (c) Ursuegui, S.; Chivot, N.; Moutin, S.; Burr, A.; Fossey, C.; Cailly, T.; Laayoun, A.; Fabis, F.; Laurent, A. Biotin-conjugated N-methylisatoic anhydride: a chemical tool for nucleic acid separation by selective 2′-hydroxyl acylation of RNA. *Chem. Comm.* **2014**, *50*, 5748−5751. (d) Fessler, A. B.; Fowler, A. J.; Ogle, C. A. Directly Quantifiable Biotinylation Using a Water-Soluble Isatoic Anhydride Platform. *Bioconjugate Chem.* **2021**, *32*, 904−908.

(6) Kadina, A.; Kietrys, A. M.; Kool, E. T. RNA Cloaking by Reversible Acylation. *Angew. Chem., Int. Ed. Engl.* **2018**, *57*, 3059−3063.

(7) Habibian, M.; Velema, W. A.; Kietrys, A. M.; Onishi, Y.; Kool, E. T. Polyacetate and Polycarbonate RNA: Acylating Reagents and Properties. *Org. Lett.* **2019**, *21*, 5413−5416.

(8) Xiao, L.; Habibian, M.; Kool, E. T. Site-Selective RNA Functionalization via DNA-Induced Structure. *J. Am. Chem. Soc.* **2020**, *142*, 16357−16363.

(9) Xiao, L.; Jun, Y. W.; Kool, E. T. DNA Tiling Enables Precise Acylation-Based Labeling and Control of mRNA. *Angew. Chem., Int. Ed. Engl.* **2021**, *60*, 26798−26805.

(10) McGinnis, J. L.; Dunkle, J. A.; Cate, J. H. D.; Weeks, K. M. The Mechanisms of RNA SHAPE Chemistry. *J. Am. Chem. Soc.* **2012**, *134*, 6617−6624.

(11) Cao, J.; Xue, Y. Characteristic chemical probing patterns of loop motifs improve prediction accuracy of RNA secondary structures. *Nucleic Acids Res.* **2021**, *49*, 4294−4307.

(12) Xiao, L.; Fang, L.; Kool, E. T. Acylation probing of "generic" RNA libraries reveals critical influence of loop constraints on reactivity. *Cell Chem. Biol.* **2022**, *29*, 1341.

(13) Busan, S.; Weidmann, C. A.; Sengupta, A.; Weeks, K. M. Guidelines for SHAPE Reagent Choice and Detection Strategy for RNA Structure Probing Studies. *Biochemistry* **2019**, *58*, 2655−2664.

(14) Lee, B.; Flynn, R. A.; Kadina, A.; Guo, J. K.; Kool, E. T.; Chang, H. Y. Comparison of SHAPE reagents for mapping RNA structures inside living cells. *RNA* **2017**, *23*, 169−174.

(15) Marinus, T.; Fessler, A. B.; Ogle, C. A.; Incarnato, D. A novel SHAPE reagent enables the analysis of RNA structure in living cells with unprecedented accuracy. *Nucleic Acids Res.* **2021**, *49*, No. e34.

(16) Park, H. S.; Kietrys, A. M.; Kool, E. T. Simple alkanoyl acylating agents for reversible RNA functionalization and control. *Chem. Comm.* **2019**, *55*, 5135−5138.

(17) Mortimer, S. A.; Weeks, K. M. A Fast-Acting Reagent for Accurate Analysis of RNA Secondary and Tertiary Structure by SHAPE Chemistry. *J. Am. Chem. Soc.* **2007**, *129*, 4144−4145.

(18) Spitale, R. C.; Crisalli, P.; Flynn, R. A.; Torre, E. A.; Kool, E. T.; Chang, H. Y. RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* **2013**, *9*, 18−20.

(19) (a) Cao, C.; Wei, P.; Li, R.; Zhong, Y.; Li, X.; Xue, F.; Shi, Y.; Yi, T. Ribosomal RNA-Selective Light-Up Fluorescent Probe for

Rapidly Imaging the Nucleolus in Live Cells. *ACS Sens* **2019**, *4*, 1409−1416. (b) Takenaka, S.; Nishira, S.; Tahara, K.; Kondo, H.; Takagi, M. Synthesis and characterization of novel tris-intercalators having potentially two different DNA binding modes. *Supramol. Chem.* **1993**, *2*, 41−46.
(20) Li, Q.; Kim, Y.; Namm, J.; Kulkarni, A.; Rosania, G. R.; Ahn, Y.-H.; Chang, Y.-T. RNA-Selective, Live Cell Imaging Probes for Studying Nuclear Structure and Function. *Chem. Biol.* **2006**, *13*, 615−623.
(21) Mironov, A. S.; Gusarov, I.; Rafikov, R.; Lopez, L. E.; Shatalin, K.; Kreneva, R. A.; Perumov, D. A.; Nudler, E. Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria. *Cell* **2002**, *111*, 747−756.
(22) (a) Szymanski, M.; Barciszewska, M. Z.; Erdmann, V. A.; Barciszewski, J. 5S Ribosomal RNA Database. *Nucleic Acids Res.* **2002**, *30*, 176−178. (b) Ciganda, M.; Williams, N. Eukaryotic 5S rRNA biogenesis. *WIREs RNA* **2011**, *2*, 523−533.
(23) Bowler, F. R.; Chan, C. K. W.; Duffy, C. D.; Gerland, B.; Islam, S.; Powner, M. W.; Sutherland, J. D.; Xu, J. Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat. Chem.* **2013**, *5*, 383−389.
(24) Stephenson, W.; Razaghi, R.; Busan, S.; Weeks, K. M.; Timp, W.; Smibert, P. Direct detection of RNA modifications and structure using single-molecule nanopore sequencing. *Cell Genomics* **2022**, *2*, 100097.
(25) Fang, L.; Xiao, L.; Jun, Y. W.; Onishi, Y.; Kool, E. T. Reversible 2'-OH Acylation Enhances RNA Stability. *Research Square* **2022**, DOI: 10.21203/rs.3.rs-1483354/v1.

https://doi.org/10.1021/jacs.2c09040
J. Am. Chem. Soc. 2023, 145, 143−151