Site-specific dual-color labeling of long RNAs for single-molecule spectroscopy

Meng Zhao, Fabio D. Steffen, Richard Börner, Michelle F. Schaffer, Roland K. O. Sigel* and Eva Freisinger*

Department of Chemistry, University of Zurich, Zurich 8057, Switzerland

Received June 28, 2017; Revised September 15, 2017; Editorial Decision October 14, 2017; Accepted October 21, 2017

ABSTRACT
Labeling of long RNA molecules in a site-specific yet generally applicable manner is integral to many spectroscopic applications. Here we present a novel covalent labeling approach that is site-specific and scalable to long intricately folded RNAs. In this approach, a custom-designed DNA strand that hybridizes to the RNA guides a reactive group to target a preselected adenine residue. The functionalized nucleotide along with the concomitantly oxidized 3’-terminus can subsequently be conjugated to two different fluorophores via bio-orthogonal chemistry. We validate this modular labeling platform using a regulatory RNA of 275 nucleotides, the btuB riboswitch of Escherichia coli, demonstrate its general applicability by modifying a base within a duplex, and show its site-selectivity in targeting a pair of adjacent adenines. Native folding and function of the RNA is confirmed on the single-molecule level by using FRET as a sensor to visualize and characterize the conformational equilibrium of the riboswitch upon binding of its cofactor adenosylcobalamin. The presented labeling strategy overcomes size and site constraints that have hampered routine production of labeled RNA that are beyond 200 nt in length.

INTRODUCTION
Single-molecule (sm) spectroscopy along with Förster resonance energy transfer (FRET) has emerged as a versatile tool for probing distance distributions in biomacromolecules on the nanometer scale (1,2). Single-molecule detection can unravel conformational, kinetic, and thermodynamic heterogeneities of biomolecules like DNA, RNA, and proteins to provide information that is lost in ensemble experiments (3–6). An integral part of fluorescence spectroscopy is the choice and incorporation of suitable extrinsic fluorophores. smFRET requires a set of two spectrally overlapping dyes, where the donor transfers its energy to the acceptor via dipole-dipole coupling. Working with long nucleic acids on the single-molecule level still remains challenging due to the lack of a universally applicable method that allows the introduction of a FRET dye pair in an orthogonal and site-specific manner while preserving the correct fold and optimal function of the biomolecule.

Many current approaches utilize modified nucleotides that are incorporated either chemically or enzymatically into the nucleic acid sequence for post-synthetic labeling. Most of these methods have either limitations concerning the size of the nucleic acid or lack site-specificity (7–9). Only recently, a platform for position-selective labeling of RNA has been introduced that combines T7 RNA polymerase transcription with solid-phase synthesis (SPS) (10). While being flexible in the type of labels, sequence upsampling is still problematic as yields are gradually decreasing with the number of repeating cycles. Moreover, ensuring selectivity within patches of consecutive identical nucleotides remains challenging. Currently, labeled RNAs of sizes smaller than 200 nt are usually prepared from custom-designed fragments which are ligated post-synthetically using T4 RNA or DNA ligase (11). Ligase-assisted approaches are, however, inherently restricted in the number and location of ligation sites. Intricate secondary structures complicate the design of the fragments to be ligated. In addition, production of the segments containing functionalized nucleotides still relies on solid phase synthesis and is therefore subject to size limitations. One labeling strategy which is independent of RNA size consists in hybridizing fluorophore that carry complementary sequences (DNA or PNA) to an in vitro transcribed RNA (12–14). However, this strategy often requires the insertion of non-native loops that may potentially alter the folding and function of the RNA. In case of the B12 riboswitch dealt with herein, numerous attempts to introduce such inserts have led to RNA misfolding eventually rendering the switch inactive (data not shown). Ultimately, some methods use naturally occurring or in vitro selected enzymes to modify specific RNA substrates or in-

*Correspondence may also be addressed to Roland K.O. Sigel. Tel: +41 44 65 635 4652; Fax: +41 44 635 6802; Email: roland.sigel@chem.uzh.ch

Present address: Meng Zhao, Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada.

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
assert a labeled mono- or polynucleotide at the 2′-ribose of an existing RNA strand (15–17).

Building on our recently developed strategy for DNA-templated generation of etheno-adducts within short, single-stranded DNA (18) and RNA (19), we describe here a novel post-transcriptional and covalent labeling procedure for long RNAs with the ability to specifically graft an alkyne group on adenine residues buried in internal double-stranded regions (Figure 1A). In order to increase the accessibility of the target sites, complementary DNA oligonucleotides, known as helper sequences (dHSs) (20), are employed to disrupt the local secondary structure of the target RNA. During activation of the reactive strand (dRS) precursor, a concurrent oxidative opening of the 3′-terminal ribose ring to a dialdehyde moiety occurs and offers the possibility to place a second fluorophore at this position. The distinct orthogonal chemistry of the two functional groups at the internal and the 3′-terminal site allows for a site-specific dual-color labeling of RNA without the specter of cross-reactivity.

Here, we target a non-coding 275 nt long riboswitch that is embedded in the 5′-untranslated region (UTR) of the btuB gene from Escherichia coli (21,22). This B12-responsive riboswitch regulates the expression of an outer-membrane transport protein that shuffles corrinoids like adenosylcobalamin (Ado-Cbl) into the periplasmic space (23,24). The RNA dynamics are monitored on the single-molecule level. The presented covalent and site-specific labeling protocol is novel in terms of RNA size (>200 nt) and its flexibility in choosing the most informative labeling sites while leaving the original sequence untouched.

MATERIALS AND METHODS
dHS assisted hybridization of a dRS analog
A mixture of btuB riboswitch (1 µl, 1 µM) and dHS (1 µl, 0.1–50 µM) buffered in 1 M sodium acetate, pH 5.5, was heated to 70°C for 5 min followed by incubation at 25°C for 30 min. Upon addition of the 32P-labeled dRS analog (1 µl, 1 µM) the sample was put on ice for 10 min before loading it on a 6% (w/v) native polyacrylamide gel. The gel was run at room temperature, dried in vacuo and transferred onto a phosphor screen for scanning.

Site-specific modification of the btuB riboswitch
The btuB riboswitch (80 µl, 36 µM) buffered in 1 M sodium acetate, pH 5.5, was annealed (70°C, 5 min) with the respective dHS (1 eq. to btuB riboswitch). After incubation (25°C, 30 min) the freshly prepared dRS precursor (25 µl, in 1 M sodium acetate, pH 5.5) was added and the solution was placed on ice for 10 min. Activation with sodium periodate (15 µl, 20 mM, in 20 mM sodium acetate, pH 5.5) was carried out under mild agitation (25°C, 500 rpm, 90 min). The reaction was quenched by addition of ethylene glycol (30 µl, 15 mM, in 20 mM sodium acetate, pH 5.5) and continued to incubate overnight. Desalting over a NAP-5 column, lyophilization and purification with 5% (w/v) denaturing PAGE afforded the functionalized btuB riboswitch devoid of dHS and dRS. The RNA was recovered from the gel by soaking the crushed pieces in a buffer containing 3-(N-morpholino)propanesulfonic acid (MOPS, 10 mM, pH 6), EDTA (1 mM) and NaCl (250 mM), followed by EtOH precipitation. The internal alkyne modification was detected based on its characteristic fluorescence emission profile with a maximum around 415 nm upon excitation at 275 nm. The presence of the 3′-dialdehyde was demonstrated indirectly via coupling to the Cy3-hydrazide dye (vide infra).

Coupling of carboxylic dyes to the modified btuB riboswitch
In order to label the 3′-terminal dialdehyde, the modified btuB riboswitch (25 µl, 25 µM, in 0.5 M triethylammonium acetate (TEAA), pH 7) was mixed with formamide (5 µl, purum) and DMSO (5 µl, purum), flushed with argon for 1 min and combined with argon-flushed Cy3-hydrazide (2 µl, 10 mM in DMSO). Following incubation in the dark (25°C, 500 rpm, overnight), the RNA was precipitated with EtOH (75%). Subsequently, the sample was conjugated with sCy5-azide (sulfonated, Lumiprobe) without further purification to generate the double-labeled btuB riboswitch (vide infra).

The coupling of an azide-functionalized carboxylic dye sCy5 (or sCy3, Figure 2) to the internal alkyne group was performed via a copper-catalyzed [3+2] cycloaddition (CuAAC). The 3′-modified btuB riboswitch (25 µl, 25 µM) was buffered with 0.5 M TEAA, pH 7, treated with formamide (10 µl, purum), DMSO (55 µl, purum), sCy5-azide (2 µl, 10 mM in DMSO) and ascorbic acid (10 µl, 5 mM) and flushed with argon for 1 min. A solution containing CuSO4 and tris(benzyltriazolylmethyl)amine (TBTA) in a 1:1 ratio (5 µl, 10 mM in 55% (v/v) DMSO) was added, followed by argon-flushing (1 min) and incubation in the dark (25°C, 500 rpm, overnight). Excess free dyes were removed by precipitating the RNA in EtOH, desalting the redissolved RNA over a NAP-5 column and purifying the crude product by 5% denaturing PAGE. Labeling yields were calculated by comparing the absorption maxima at 260 nm (RNA) and 550 nm (Cy3) or 650 nm (sCy5) weighted by the respective extinction coefficient

\[
\psi = \frac{A_{550 or 650/\epsilon_{btuB}}}{A_{260/\epsilon_{Cy3 or sCy5}}} \tag{1}
\]

with \(\epsilon_{btuB} = 3022 \text{mM}^{-1} \text{cm}^{-1}\) (25), \(\epsilon_{Cy3} = 150 \text{mM}^{-1} \text{cm}^{-1}\) and \(\epsilon_{sCy5} = 250 \text{mM}^{-1} \text{cm}^{-1}\) (26).

Additional methodological information is available in the supporting information.

RESULTS AND DISCUSSION
In the following, we will outline the steps involved in the design of suitable reactive and helper strands, their hybridization to the riboswitch and subsequent coupling to the fluorophores. We further evaluate the site-specificity of the modification and prove the integrity of both the RNA and the dyes upon bioconjugation via fluorescence spectroscopy.

Selection of the labeling sites
The btuB riboswitch construct comprises the full-length aptamer and expression platform (240 nt) and is extended...
Figure 1. Site-specific dual-color labeling of long RNAs. (A) Schematic illustration of the double labeling workflow at an internal adenine and the 3′-end. (i) dHS annealing, (ii) hybridization of the dRS precursor, (iii) activation by sodium periodate, (iv) transfer of the reactive group and modification of an internal adenine, (v) orthogonal fluorophore coupling with Cy3-hydrazide at the 3′-end dialdehyde and sCy5-azide at the internal alkyne-εA via CuAAC. (B) Depiction of the translational control by the btuB riboswitch (22): the equilibrium between two conformers is shifted in response to the cofactor adenosylcobalamin (Ado-Cbl). In the apo form, the Shine–Dalgarno sequence (or ribosomal binding site, RBS) is accessible to the ribosome. Conversely, when Ado-Cbl levels increase, the RNA establishes a kissing loop interaction whereupon the RBS is sequestered through formation of an RBS hairpin motif. This equilibrium can be probed via single-molecule FRET by placing a donor fluorophore (Cy3) at the 3′-terminus and an acceptor (sCy5) at A213 or A35 (indicated with arrows).

at the 5′-end by 35 nucleotides of the natural genomic sequence to allow surface immobilization for total internal reflection fluorescence (TIRF) microscopy (Figure 1B and Supplementary Figure S1). Adenines 35 (A35, note that the 5′-elongation is not part of the numbering scheme) and adenine 213 (A213) are chosen as the labeling sites for either one of the sulfo-carbocyanine fluorophores sCy5 or sCy3. A35 is located in an internal hairpin loop in close proximity to the kissing loop that forms upon binding of the cofactor Ado-Cbl (Figure 1B). The expression platform harbors a second modification site, A213, which is particularly challenging to target as it is situated in a duplex region. In conjunction with a suitable fluorophore at the 3′-end, the FRET intensity reflects the underlying conformational equilibrium that characterizes the sequestration of the ribosomal binding site (RBS) during translational control by the riboswitch.

The synthesis of the reactive group (RG), as part of the reactive strand (dRS), was described previously (Supplementary Figures S2 and S3) (19). No adenines and cytosines are placed next to the RG in the dRS in order to prevent self-modification (Supplementary Figures S6 and S7) (18,19). For reasons of clarity, we refer to the target ribonucleotide opposite the one linked to the RG as being the n position (Figure 1A). The RNA residue which is next to n but not hydrogen-bonded to the dRS is in the n + 1 position. Both dRSs where designed such that the target adenines are situated in the n + 2 position, which is the ideal position due to the linker length of the RG (Supplementary Figure S3) and fulfills the criterion that the dRS lacks a terminal A or C base. We previously demonstrated that labeling efficiencies for the n + 1 and n + 2 positions are similar and then progressively diminish with increasing distance from the RG (18,19). This makes designing the dRS more flexible since more labeling sites are tolerated.

Helper sequences assist dRS annealing

Targeted delivery of a functional group to a nucleic acid by DNA-templated synthesis (27) has so far only been applied to single-stranded DNA (18) or long single-stranded regions of RNA molecules (19). The complex secondary and tertiary architecture of the folded btuB riboswitch restricts the accessibility of certain ribonucleotides and hampers hybridization of the dRS to the RNA. To protect the integrity of both the RNA and the RG, we avoid exposure to high temperatures during the modification process and melt the local base-pairing regime by annealing complementary dHSs (15,20) at moderate temperatures up- or downstream of the RNA sequence complementary to the dRS before activating the RG in situ (Figure 1A). The dHS is designed in such a way to sufficiently disrupt the secondary structure elements in proximity of the desired labeling site. The length
Figure 2. Evaluating the site-specificity of internal labeling by reverse transcription. Internally labeled btuB riboswitch constructs εA35(sCy3)-btuB (A) and εA213(sCy3)-btuB (B) are used as templates for the reverse transcription of short cDNA strands. The autoradiograph of the denaturing, polyacrylamide footprinting gel (18% w/v, 7 M urea) shows the original cDNA (lane 1) and the extended cDNA in presence of the non-labeled btuB riboswitch (lane 2) or the internally labeled btuB riboswitch (lane 3). Short DNA oligonucleotides, cDNA1 +3nt (A) and cDNA2 + 7nt (B) are shown as references (lane 4). The weak band at cDNA + 3 nt (A) or cDNA + 6 nt (B) along with the stronger band at cDNA +4nt (A) and cDNA +7nt (B) reflect a delayed chain termination (CT) upon insertion of one additional nucleotide past the modification site, as previously observed (31).

of the dHS majorly depends on the RNA secondary structure and the GC content and accordingly on the predicted melting temperature.

In case of A35, two helper sequences (dHS1a and dHS1b) are designed to mask short segments on either side of the target nucleotide (Supplementary Figure S4A) and thereby open up the stem loop temporarily. To assess the efficiency of the dHS in promoting dRS hybridization, a 32P-labeled DNA oligonucleotide, referred to as dRS analog, was added to the btuB RNA in presence of increasing amounts of dHS. In the absence of any dHS only 60% of the dRS1 analog bind to the RNA at room temperature. In contrast, treatment with one equivalent of dHS1a and dHS1b each increases the binding efficiency of dRS1 to >95% at ambient temperature (Supplementary Figure S4B and C). By comparison, temperature induced unfolding shows only ~70% hybridization efficiency in absence of any dHS. More evidently in case of A213, only about 20% of the dRS2 analog anneal to the RNA in absence of dHS2 at room temperature, whereas addition of one equivalent of dHS2 leads to ~75% bound dRS2 (Supplementary Figure S5b−c). Again, thermal melting without dHS2 produces no more than 50% annealed DNA. In brief, the dHSs enable efficient hybridization of the dRS at room temperature – a prerequisite for an accurate positioning of the RG close to the target sites A35 and A213 where the functional alkyne group should be transferred onto.

RNA functionalization in a one-pot reaction

After annealing of the dHS and the dRS to the RNA, activation of the RG in situ under mild oxidative conditions to a reactive aldehyde is achieved by sodium periodate (Supplementary Figure S6−S7). At the same time, the two hydroxyl groups of the 3′-terminal ribose unit are oxidized to a di-aldehyde (28,29). At the internal site, the exocyclic amine of adenine forms an etheno adduct (εA) with an additional alkyne moiety. Etheno bridge formation is detected by exciting the RNA at either 275 nm or 308 nm to produce an emission peak around 415 nm that is characteristic of εA (Supplementary Figure S8) (30). As a result of this one-pot reaction the riboswitch is modified with two chemically orthogonal functional groups, one at an internal adenine and another at the 3′-terminus. Fluorophore coupling can now be achieved by treating the functionalized RNA with azide-
and hydrazide-functionalized fluorophores without worrying about dye cross-reactivity.

Site-specificity of the internal labeling scheme

As stated previously, the site-specificity of the internal nucleotide modification is crucial for subsequent single-molecule experiments. A polymerase stop assay was applied (18,19), in which a cDNA primer annells downstream of the putative labeling site and is extended by a reverse transcriptase (RT). Two dye labeled constructs, \( e_{A35(sCy3)}-btuB \) and \( e_{A213(sCy3)}-btuB \), are used to assess site-specificity of the internal labeling. The footprinting pattern (Figure 2a) around fluorophore position A35 shows distinct bands corresponding to the cDNA primer extended by 1, 2 and 4 nt. The RT is not stalled at the presumed labeling site (cDNA1 + 3 nt) but inserts one additional nucleotide before stopping and releasing the nascent cDNA, a process also known as delayed chain termination (CT) (31). For \( e_{A213} \) (Figure 2b), several bands representing cDNA2 + 1–3 nt are observed for both the non-labeled riboswitch and the \( e_{A213(sCy3)}-btuB \) construct, which reflects the stalling of the RT during initiation of primer extension (32). The additional bands in the vicinity of the labeling site (cDNA2 + 6 nt) are very weak. Major pausing of the RT occurs again one nucleotide after the modification site. In combination with the fluorescence emission profile of \( e_{A} \) (Supplementary Figure S8), this is a clear indication that etheno bridge formation occurs predominantly at A35 and A213, whereas \( e_{A36} \), and \( e_{C212} \) are present in trace amounts only, if at all. Moreover, labels at those sites would only slightly affect smFRET measurements as mono-labeling at \( n \pm 1 \) or 2 nt changes the transfer efficiency by about 0.08 in the worst case (calculated for a helical rise of 0.28 nm with maximal sensitivity around the Förster radius \( R_0 = 5.4 \) nm). Importantly, and in contrast to current ligation approaches, our labeling strategy prohibits non-templated nucleotide incorporation and thus the underlying RNA sequence remains unchanged.

Double labeling yields are sufficient for single-molecule experiments

The reaction of 3'-terminal periodate-activated sugars with hydrazide functionalized fluorophores proceeds with relatively high yields (29). To avoid reduction and hence inactivation of the 3'-dialdehyde during the double-labeling procedure the functionalized butuB riboswitch was first incubated with a hydrazide derivatized Cy3-dye. Only subsequently, CuAAC Click chemistry under reductive conditions was used to couple the azide derivative of sCy5 to the alkyne modified internal adenines A35 or A213, respectively (Supplementary Figure S9). Co-localized green and red emission on denaturing gels (Figure 3A and Supplementary Figure S12A) along with the slightly red-shifted absorption and fluorescence emission spectra, as compared to the free dyes (Supplementary Figure S10), indicates that both fluorophores are covalently attached to the riboswitch constructs 3'-Cy3)-eA35(sCy5)-btuB (1) and 3'- (Cy3)-eA213(sCy5)-btuB (2). Labeling efficiencies (Eq. 1) are 26% (Cy3) and 14% (sCy5) for construct 1 and 37% (Cy3) and 17% (sCy5) for construct 2, respectively. Double labeling yields are about 5% which is generally sufficient for single-molecule fluorescence spectroscopy. In comparison, classical SPS of RNA sequences proceeds with ~98% yield per nucleotide, affording an overall yield of ~0.4% for an RNA of 275 nt in length. This yield will be further reduced by post-synthetic labeling. In our case, the final yields result from four independent reactions: (i) internal adenine modification, (ii) 3'-end modification, (iii) internal dye coupling via CuAAC reaction, and (iv) 3'-end dye coupling. Step (i) in particular is sequence dependent and yields are expected to vary depending on the accessibility of the modification site. We purposely chose sites that are buried and thus not easily tractable, thereby making the use of HS sequences essential and pushing the method to its limits. Labeling of a 633 nt long RNA in single stranded regions has been shown previously (19). Here, we demonstrate that RNAs, irrespective of their length, can be double-labeled and not only in internal loops but also in helical duplex regions. Proper folding of the single- and dual-color-labeled RNA constructs was evaluated by fluorescent PAGE (Supplementary Figure S11). All constructs show the same migration behavior as the non-labeled one and hence adopt similar 3D folds. Accordingly, the dyes do not interfere with the folding of the RNA.

Photophysics of cyanine dyes within the riboswitch environment

Carbocyanine fluorophores are known to be sensitive to their local environment which is reflected in changes of the fluorescence lifetime, quantum yield and anisotropy (33–35). The torsional flexibility of the polymethine chain of carbocyanines causes an isomerization to a non-fluorescent mono-cis state. The photoisomerization reaction is an excited state process that is conditioned by sterical constraints imposed on the fluorophore. Interactions with the biomolecular environment define the relative populations of trans and cis states, which are experimentally accessible via a change in molecular brightness. In other words, the RNA shapes the photophysics of the fluorophore, a phenomenon referred to as RNA-induced fluorescence enhancement (RIFE) (34). In order to demonstrate the integrity of the dyes, we measured fluorescence lifetimes of the dyes prior to and after coupling to the riboswitch (Figure 3B and Supplementary Figure S12B). While the fluorescence of the free azide form of sCy5 decays monoexponentially with a lifetime of \( \tau = 0.76 \pm 0.03 \) ns, a second lifetime component (\( \tau_2 = 1.9 \pm 0.1 \) ns) emerges upon covalent attachment to \( e_{A213} \) (Figure 3B). Both components have a relative weight of roughly 50% (Supplementary Table S2). The lifetime decay of the 3’-terminal Cy3 can also be represented by a biexponential function (\( \tau_1 = 0.67 \pm 0.02 \) ns and \( \tau_2 = 2.6 \pm 0.1 \) ns). Compared to the free dye in solution (\( \tau = 0.11 \pm 0.01 \) ns) the lifetime is drastically increased, indicating a considerable reduction in torsional flexibility and photoisomerization as expected.

The extent of motional restriction by the RNA environment is illustrated by recording the fluorescence anisotropy as a function of time (Figure 3C and Supplementary Figure S12C). The decays are modeled in terms of a fast, hindered rotation (wobbling-in-cone model) governed by the
fluorophore linker ($\tau_{\text{local}}$) and a decoupled, global tumbling ($\tau_{\text{global}}$) component along with the entire construct (Eq. S2) (34). The long global correlation time ($\tau_{\text{global}} > 40 \text{ ns}$) and high residual anisotropy ($r_{\infty} > 0.22$) of the fluorophores suggest considerable stacking of the fluorophores on the ribonucleic acid backbone as observed previously for various RNA systems (Supplementary Table S2) (34). The non-sulfonated Cy3 is intrinsically more hydrophobic and displays a higher interaction propensity with the RNA. Moreover, it is directly coupled to the 3′-end which allows the dye to additionally stack onto the terminal guanine (36). The reduced local correlation time of the 3′-Cy3 ($\tau_{\text{local}} \sim 0.2 \text{ ns}$) in comparison to Cy3 free in solution ($\tau_{\text{local}} = 0.92 \text{ ns}$) as well as the high residual anisotropy ($r_{\infty} = 0.35$) reflect the dye wobbling within a small cone angle and support the discussion above. The internal modification with sCy5, however, shows about 50% free dye rotation and no reduction of the local correlation time. Even though the tumbling of the biomolecule-tethered fluorophores is spatially constrained, the dyes retain full functionality. It is noteworthy that local confinements on the fluorophore movement affect the orientation parameter $\kappa^2$ in the Förster radius $R_0$ and thus need to be considered for exact distance measurements.

Riboswitch dynamics monitored by smFRET

Single-molecule spectroscopy has the key advantage to dissect the ensemble and monitor individual molecules over time. Total internal reflection fluorescence (TIRF) microscopy usually requires the molecule of interest to be tethered to the surface of a quartz slide (37). For that purpose, a biotin-tagged DNA oligonucleotide is hybridized to the 5′-elongation of the riboswitch (Supplementary Figure S13a). Applying an alternating-laser excitation (ALEX) (38) scheme with illumination at 532 and 640 nm, time traces of the dual-color-labeled RNA are recorded in the absence and presence of the metabolite Ado-Cbl. The latter is reported to quench the fluorescence emission of the Cy3 donor to some extent (4,23). Nonetheless, we are able to detect the molecules upon Ado–Cbl binding and observe the energy transfer. The 3′-(Cy3)-eA213(sCy5)-btuB (2) construct probes the formation of the RBS hairpin (Figure 4A). Abrupt single-step photobleaching of Cy3 and sCy5 within the observation time along with a stable stoichiometry of Cy3 (SCy3) fluctuating around 0.5 suggests that both dyes are covalently attached to the same RNA molecule (Supplementary Figure S13B). The riboswitch interconverts between a low and a high FRET state in the presence of 3 mM Mg2+ (Figure 4B and Supplementary Figure S14) (39). Upon addition of the cofactor the relative populations of the states shift in favor of the high transfer efficiencies. FRET time traces are analyzed with a home-written Matlab software (40) to build cumulative histograms (Figure 4C). In addition to the high and low FRET states, we find a substantial contribution of intermediate FRET states originating from static molecules (Supplementary Figure S14, static molecule 2). Additionally, we observe dynamic molecules that show short excursions to either the high or the low FRET state (Figure 4B and Supplementary Figure S14, dynamic molecule 2 and 3). Interestingly, consecutive transitions from low to intermediate to high FRET and vice versa are barely visible. Intermediate states between 0.2 and 0.8 may arise from time-averaging (41), but can also be attributed to trapped conformations of the riboswitch expression platform. Further, shifting FRET states can arise from quantum yield changes of carbocyanine fluorophores due to differences in the photoisomerization probability in the local RNA environment, as reported
Figure 4. Single-molecule FRET characterization of 3′-Cy3-εA213(sCy5)-btuB (2). (A) Formation of the RBS hairpin upon addition of Ado-Cbl (B12). (B) FRET trajectories in the absence and presence of 0.1 mM Ado-Cbl (50 mM KCl, 3 mM MgCl2). FRET trajectories have been partly rescaled to visualize fast state transitions from/to intermediate states (indicated by black arrows). (C) Normalized cumulated FRET histograms built up from about 90 molecules showing a predominant occupancy of high FRET values in the presence of the metabolite (red).

recently (34,42). Indeed, Stoichiometry-FRET histograms show substantial broadening which supports an influence of dye photophysics (Supplementary Figure S13C). A more detailed analysis and characterization of the btuB folding pathway will be addressed elsewhere. Importantly, different population levels of lower and higher FRET states in response to Ado-Cbl provide evidence of a tunable equilibrium between the apo and the bound conformation as observed previously for a hydroxocobalamin riboswitch (4).

CONCLUSION
Current efforts in co- and post-transcriptional labeling of nucleic acids are directed towards finding the best trade-off between site-specificity, programmability and size of the target molecule while preserving the integrity of the native RNA structure and folding. While PLOR is a versatile platform adaptable to many different applications, the practicability for very long RNAs remains elusive. Ligase-assisted approaches on the other hand have been used to prepare riboswitch domains with up to 200 nt, but suffer from restrictions in the number and positioning of appropriate ligation sites. Non-covalent hybridization approaches are generally independent of the RNA dimensions, however they usually require changes made to the native sequence in order to accommodate the fluorescently labeled complement.

We developed a minimally invasive, nucleotide- and position-selective labeling strategy which is scalable to RNAs beyond 200 nt in length. Our method utilizes the selective annealing of custom-designed DNA probes to modify an in vitro transcribed RNA in situ in principle at any specific adenine or cytosine. Regions that feature buried nucleotides such as internal stem loops and duplex structures are tackled by introducing additional dHSSs, which work synergistically with the dRSs to disrupt the local secondary structure. The DNA-templated transfer of a functional group to an adenine and subsequent coupling with a fluorescent probe is shown to preserve the overall RNA fold and function. Cross-contamination in dual-color labeling is avoided by using two different bio-orthogonal coupling reactions. The presented labeling platform has been validated on the ensemble and single-molecule level by means of TSCPC and smFRET using a long riboswitch as a model system. The modular nature of the method allows the design of alternative functional handles including azides, amines, or thiols to pave the way for three-color FRET. On the RNA side, the scope of possible modification targets can be readily extended to cytosines which also feature an exocyclic amine group. Overall, this study provides a novel approach to fluorescently label long nucleic acids. It is highly customizable with respect to the location of the labeling sites and hence particularly suitable for studying large and sensitive RNA systems such as riboswitches or ribozymes.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
The authors thank Igor A. Oleinich for valuable discussions. M.Z. thanks Mélodie C.A.S. Hadzic for her help with the smFRET data analysis in MASH.

FUNDING
Financial support from the Swiss National Science Foundation [to E.F. and R.K.O.S.]; European Research Council [ERC to R.K.O.S.]; SystemsX.ch [to R.K.O.S.]; UZH Forschungskredit [FK-14-096, FK-15-095 to R.B.]; UZH Stiftung für wissenschaftliche Forschung [to R.K.O.S. and R.B.]; University of Zurich and the SBFI [COST Action CM1105 to E.F. and R.K.O.S]. Funding for open access charge: University of Zurich.

Conflict of interest statement. None declared.

REFERENCES
1. Roy, R., Hohng, S. and Ha, T. (2008) A practical guide to single-molecule FRET. Nat. Methods, 5, 507–516.
2. Schuler, B. (2013) Single-molecule FRET of protein structure and dynamics - a primer. *J. Nanobiotechnol.*, 11, 52.

3. Haller, A., Altman, R.B., Soulère, M.F., Blanchard, S.C. and Micura, R. (2013) Folding and ligand recognition of the TPP riboswitch aptamer at single-molecule resolution. *Proc. Natl. Acad. Sci. U.S.A.*, 110, 4188–4193.

4. Holmstrom, E.D., Polaski, J.T., Batey, R.T. and Nesbitt, D.J. (2014) Single-molecule conformational dynamics of a biologically functional hydroxocobalamin riboswitch. *J. Am. Chem. Soc.*, 136, 16832–16843.

5. Karunatilaka, K.S. and Rueda, D. (2009) Single-molecule fluorescence studies of RNA: a decade’s progress. *Chem. Phys. Lett.*, 476, 1–10.

6. Börner, R., Kowerko, D., Miserachs, H.G., Schaffer, M.F. and Sigel, R.K.O. (2015) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coord. Chem. Rev.*, 327–328, 123–142.

7. Rao, H., Tanpure, A.A., Sawant, A.A. and Srivatsan, S.G. (2012) Enzymatic incorporation of an azide-modified UTP analog into oligoribonucleotides for post-transcriptional chemical functionalization. *Nat. Protoc.*, 7, 1097–1112.

8. Lavergne, T., Lamichhane, R., Malysh, D.A., Li, Z., Li, L., Sperling, E., Williamson, J.R., Millar, D.P. and Romesberg, F.E. (2016) FRET characterization of complex conformational changes in a large 165 ribosomal RNA fragment site-specifically labeled using unnatural base pairs. *ACS Chem. Biol.*, 11, 1347–1353.

9. Seidu-Larry, S., Krieg, B., Hirsch, M., Helm, M. and Domingo, O. (2012) A modified guanosine phosphoramidite for click functionalization of RNA on the sugar edge. *Chem. Commun.*, 48, 11014–11016.

10. Liu, Y., Holmstrom, E., Zhang, J., Yu, P., Wang, J., Dyba, M.A., Chen, D., Ying, J., Lockett, S., Nesbitt, D.J. et al. (2015) Synthesis and applications of RNAs with position-selective labelling and mosaic composition. *Nature*, 522, 368–372.

11. Lang, K. and Micura, R. (2008) The preparation of site-specifically modified riboswitch domains as an example for enzymatic ligation of chemically synthesized RNA fragments. *Nat. Protoc.*, 3, 1457–1466.

12. Smith, G.J., Sosnick, T.R., Scherer, N.F. and Pan, T. (2005) Efficient fluorescence labeling of a large RNA through oligonucleotide hybridization. *RNA*, 11, 234–239.

13. Steiner, M., Karunatilaka, K.S., Sigel, R.K.O. and Rueda, D. (2008) Single-molecule studies of group II intron ribozymes. *Proc. Natl. Acad. Sci. U.S.A.*, 105, 13853–13858.

14. Schmitz, A.G., Zelger-Paulus, S., Gasser, G. and Sigel, R.K.O. (2015) Strategy for internal labeling of large RNAs with minimal perturbation by using fluorescent PNA. *ChemBioChem*, 16, 1302–1306.

15. Bütter, N., Javadi-Zarnaghi, F. and Höbartner, C. (2014) Site-specific labeling of RNA at internal ribose hydroxyl groups. Terbium-Assisted Deoxyribozymes at Work. *J. Am. Chem. Soc.*, 136, 8131–8137.

16. Plotnikova, A., Olenich, L.A., Masevičius, V., Vilaitis, G. and Klimasauskas, S. (2014) Selective covalent labeling of miRNA and sRNA duplexes using HEIN1 methyltransferase. *J. Am. Chem. Soc.*, 136, 13550–13553.

17. Baum, D.A. and Silverman, S.K. (2007) Deoxyribozyme-catalyzed labeling of RNA. *Angev. Chem. Int. Ed.*, 46, 3502–3504.

18. Egloff, D., Olenich, L.A. and Freisinger, E. (2015) Sequence-specific generation of 1, N6-ethenoadenine and 3, N4-ethenocytosine in fluorescencely labeled oligonucleotide probes. *Appl. Environ. Microbiol.*, 86, 3603–3607.

19. Nahvi, A., Barrick, J.E. and Breaker, R.R. (2004) Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.*, 32, 143–150.