Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Journal Name

COMMUNICATION

Biofunction-assisted aptasensors based on ligand-dependent 3' processing of a suppressor tRNA in wheat germ extract

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

Atsushi Ogawa* and Junichiro Tabuchi

We have developed a novel type of biofunction-assisted aptasensor that harnesses ligand-dependent 3' processing of a premature amber suppressor tRNA and the subsequent amber suppression in a reporter gene in wheat germ extract.

Aptamers are functional nucleic acids that bind to their ligands with high affinity and specificity. Unlike antibodies, which exhibit similar characteristics, a synthetic aptamer for a particular ligand can be relatively easily and quickly obtained *in vitro* from a randomized nucleic acid pool through iterative selection-amplification cycles (i.e., *in vitro* selection).[1] Therefore, aptamers have received considerable attention as analyte-recognizing elements of biosensors (so-called aptasensors)[2] in addition to their potential as inhibitory drugs.[3] In fact, an array of methods has been developed to convert the ligand (=analyte) input signal, through a structural rearrangement of the aptamer induced by aptamer-ligand complex formation, to a detectable output signal.[2,4]

We have thus far constructed several unique RNA aptasensors that use a synthetic riboswitch, which acts as a ligand-dependent gene regulator, in a cell-free translation system.[5] As in the case of natural riboswitches, in our synthetic riboswitch, a ligand binds to its aptamer implanted in an untranslated region of mRNA to promote the downstream reporter gene expression in a cis manner. In addition, we previously reported a different type of biofunction-assisted aptasensor based on a tRNA switch that regulates suppression of a nonsense codon in an open reading frame (ORF) in a "trans" manner.[6] In this aptasensor, a ligand induces the cleavage of a premature suppressor tRNA (sup-tRNA) that has a ligand-responsive self-cleaving aptazyme (a conjugate of an aptamer and a ribozyme) in the 5' leader, to produce the mature sup-tRNA for expression of the full-length reporter protein via nonsense-codon suppression. This tRNA switch-based aptasensor has advantages in terms of its design simplicity and facile engineering compared to the riboswitch-based aptasensors described above, mainly because the tRNA probe in the former is much shorter than the mRNA probes in the latter. On the other hand, it has the disadvantages of utilizing an aptazyme that is more difficult to obtain and more sensitive to the detection conditions than a normal aptamer[5c] and in requiring a costly reconstituted cell-free translation system to enhance the aptazyme (precisely, ribozyme) activity and circumvent the enzymatic end-processing of the tRNA probe. We herein exploit a premature sup-tRNA probe that has a ribozyme-free pure aptamer in the 3' trailer, to develop a novel tRNA switch-based aptasensor, which conversely (i.e., actively) utilizes the end processing of the probe for its maturation in a relatively low-cost wheat germ extract (WGE),[7] a eukaryotic cell-free translation system.

Previously, we identified an *in vitro*-transcribed amber sup-tRNA$^{Ser}_{CUA}$ (named S2-$G_{27}C_{43}$-G_{73} or **t86**) that showed a high suppression efficiency (60-85%) in WGE,[8] through step-by-step improvement of an *Oryza sativa* nuclear tRNA^{Ser}. In addition, we have recently used this highly active sup-tRNA as the framework to investigate the end processing of tRNA in WGE.[9] Specifically, we added a 5' leader and/or a 3' trailer to **t86** to prepare premature tRNAs and evaluated their end processing with their amber suppression efficiencies and gel band intensities. As a result, the 3' processing (the RNase Z-mediated 3' trailer cleavage and the CCA addition) was found to proceed rapidly in contrast to the relatively slower 5' processing (the RNase P-mediated 5' leader cleavage). Moreover, it was elucidated that the premature tRNAs underwent *exo*-degradation from their premature end(s) at a speed slightly faster than the 5' processing, as did the general *in vitro* transcripts, whereas the processed mature tRNA (i.e., **t86**) was extremely stable. Based on these results, we herein designed a novel type of tRNA switch-based aptasensor (**Fig. 1**). The **t86**-derived premature sup-tRNA probe has three distinguishing segments in the 3' trailer: an inhibitory sequence (IS), an aptamer, and a modulator sequence (MS) in this order from 5' to 3'. The first segment is inserted to distort

Proteo-Science Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan. E-mail: ogawa.atsushi.mf@ehime-u.ac.jp; Fax: +81 89 927 8450; Tel: +81 89 927 8450

† Electronic Supplementary Information (ESI) available: Materials and methods, supplementary figures. See DOI: 10.1039/x0xx00000x

the acceptor stem and sequester the cleavage site from RNase Z by forming a stem-loop (IS-SL) with a 3' region of the tRNA body in the absence of the ligand (OFF state, Fig. 1, left). It is clear that neither aminoacylation nor amber suppression takes place without 3' processing. The second segment, of course, is needed for recognizing and binding to the ligand. An acceptable aptamer is a well-minimized one that requires the proximity of both ends for binding.[5d] The last segment is a sequence complementary to a 3' region of the IS. Given that its length is appropriately adjusted, the aptamer-ligand binding should induce duplex formation between the MS and the IS (MS-IS), which restores the tRNA structure and exposes the cleavage site, thereby promoting the rapid 3' processing (ON state, Fig. 1, right). In the OFF state, due to the inhibition of tRNA maturation, the probe is highly likely to undergo *exo*-degradation in WGE,[9] which leads to a reduction in the false-positive rate. In contrast, in the ON state, if the structural restoration through aptamer-ligand complex formation occurs relatively rapidly, the probe is expected to be successfully processed to the stable mature **t86**, which has already been shown to highly suppress the amber codon (UAG) in an ORF to express a full-length reporter protein (firefly luciferase (FLuc) encoded in **amber-mRNA(FLuc)** in this work) in WGE.[8]

We started with optimization of the first 3'-trailer segment of the probe. To ensure that RNase Z efficiently cleaved out the 3'-trailer in the ON state, a 7-nt 3' trailer (U_{74}UUGCUU) derived from the downstream sequence of the original $tRNA^{Ser}$ gene was first added to the 3' end of 3' CAA-free **t86** to prepare **3pt** (Fig. 2A, left), which exhibited a relatively high suppression efficiency (70% of that by **t86**), due to the rapid 3' processing.[9] A sequence for actually inhibiting the 3' processing in the OFF state was then ligated to the 3' end of **3pt** so that an IS-SL with a stable ACAAC loop was formed with the cleavage site (G_{73}-U_{74}) around the center of the stem (Fig. 2A, right). Because it was clear that the stem length in the IS-SL (n) contributes to the inhibition of the RNase Z-mediated cleavage, we examined its effect on the 3' processing with the suppression efficiency in 1-h incubation in WGE by varying it in increments of one base pair (bp) (Fig. 2B). As shown in Fig. 2C, premature tRNAs with an IS-SL composed of more than a 14-bp stem (**3pt-IS(14)** and **3pt-IS(15)**) showed sufficiently low suppression efficiencies (approx. 10% that of **3pt**), indicating effective inhibition of the 3' processing. In contrast, the suppression efficiency increased in a sensitive manner in response to shortening of the stem length down to 11 bp and plateaued at an approx. 1.3-fold higher efficiency than that of **3pt**. This high plateau efficiency suggested that the acceptor stem was correctly formed instead of the short IS-SL to completely expose the cleavage site in **3pt-IS(9)~(11)**.[10]

We next designed tRNA probes based on **3pt-IS(14)**, which hardly undergoes the 3' processing. For the first design, a theophylline-binding aptamer was used, because it has been well-minimized and known to require duplex formation between both ends for binding to the ligand.[11] To determine the optimal length of MS (m), some probes with the theophylline aptamer and various lengths of MS were prepared (Fig. S1A), and then their suppression efficiencies were measured after 1-h incubation without or with 1 mM theophylline (theo) in WGE. However, although the efficiencies in the OFF state were on the whole lower than that of **3pt-IS(14)**, probably due to the much longer 3' dangling end causing more rapid *exo*-degradation, those in the ON state were also equal values, meaning that no induction of expression by theophylline was observed with any length of MS (the ON/OFF ratios were approx. 1; Fig. S1B). Because these results suggested that the theophylline ligand had difficulty binding to the aptamer, we inserted one toehold (th) base pair into the ends of the aptamer in the above-mentioned failure probes to facilitate aptamer-ligand complex formation in the ON state (Figs. 3A and B). As a result, although the probes with longer MS showed a slightly higher suppression of the amber codon even in the OFF state, the suppression efficiency increased in response to the ligand in all the examined probes with 3~6-mer MS (Fig. 3C), indicating the ligand-dependent 3' processing of the probes as shown in Fig. 3A. The probe that exhibited the highest ON/OFF induction ratio was **theo(th1)-MS(4)**, with 4-mer MS (7.8 ± 0.9 at 1 mM theophylline).[12] This induction ratio was comparable to that of the previously reported tRNA switch-based sensor utilizing a theophylline-dependent aptazyme.[6] Incidentally, when one more th base pair was inserted (Fig. S1C), similar results were obtained, although the dependence on the MS length was slightly shifted toward the shorter due to the longer th (Fig. S1D).

To characterize the theophylline probe bearing the optimal length of MS, **theo(th1)-MS(4)**, we investigated the reactivity, the specificity and the sensitivity thereof. The first property was evaluated by gel electrophoresis of the probe incubated without mRNA in WGE.[9] As shown in the gel bands (Fig. 4A), approx. 30% amounts of the probe were successfully processed to maturity (though 55% were degraded) during 1-h incubation in the ON state (with 1 mM theophylline), while 80% underwent degradation with almost no processing in the OFF state, as expected. These results of gel analyses clearly indicated that the probe was processed via the relatively fast structural change upon binding to the ligand.[13] Therefore, the induction of expression shown in Fig. 3C could undoubtedly be attributed to the ligand-dependently processed probe. The second property was checked with a theophylline analogue, caffeine (caf), which has a methyl group at the N7 position. As a result, the **theo(th1)-MS(4)** probe showed caffeine unresponsiveness (ON/OFF ratio = 0.86 ± 0.08 at 1 mM caffeine) in contrast to its high theophylline responsiveness (Fig. 4B, left). In addition, when mutations were introduced into the theophylline aptamer in the probe (m-theo in Fig. 3A, right) to construct a mutated probe (**m-theo(the1)-MS(4)**), the theophylline responsiveness was lost (ON/OFF ratio = 0.84 ± 0.09 at 1 mM theophylline, Fig. 4B, left). These results prove that the specific binding between theophylline and its aptamer in the probe promoted the 3' processing and the subsequent amber suppression. The third property was examined with the suppression efficiencies at various concentrations of the ligand. The suppression efficiency was sensitively enhanced in a dose-dependent manner (Fig. 4C), reflecting the increase in

the amount of the processed probe through that of the probe-ligand complex. The detection limit was 3 μM (Fig. 4C, inset), which is comparable to that of the best riboswitch-based aptasensors for detecting theophylline.[5c]

Because the aptamer segment is basically independent from the other two 3'-trailer segments (IS and MS) and from the th base pair in the present probe design, it is highly expected that efficient probes can be readily constructed for other ligands just by replacing the theophylline aptamer in **theo(th1)-MS(4)** with the appropriate one.[5d] To confirm this assumption, we prepared two tRNA probes for tetracycline (tc) and FMN (**tc(th1)-MS(4)** and **FMN(th1)-MS(4)**, respectively) by using their aptamers,[14,15] both of which meet the above-mentioned requirements as an implanted aptamer (Fig. 5A). As a result of the suppression experiments in WGE, each probe clearly exhibited the corresponding ligand-dependent induction of expression (Figs. 5B and C). In particular, **tc(th1)-MS(4)** showed a surprisingly high ON/OFF ratio (81 ± 5 at 100 μM tetracycline) mainly due to the extremely low suppression efficiency in the OFF state (0.3% of that by **3pt**). The reason for this reduction of false-positives was probably that this probe with the much longer aptamer sequence in the 3' trailer is more susceptible to degradation.[16] In fact, it was found from gel analyses that **tc(th1)-MS(4)** underwent more severe degradation than **theo(th1)-MS(4)** (Fig. S2A). In contrast, **FMN(th1)-MS(4)** with the shorter aptamer more highly suppressed the amber codon in the OFF state.[17] These results suggest that a longer aptamer is more ideal for implantation into a tRNA probe to achieve a higher induction ratio.[18] Nevertheless, the high ON/OFF ratio of **tc(th1)-MS(4)** even at low ligand concentrations is also due to its high sensitivity, which can be attributed to more rapid (rather than stronger) aptamer-ligand complex formation: it detected 100 nM (approx. 50 ppb) of tetracycline with an induction ratio of 1.5 (Fig. S2B).

Conclusions

We harnessed several functions of WGE (particularly tRNA processing and degradation, and cell-free translation) to develop a novel type of tRNA switch-based aptasensor that is composed of a tRNA probe with an aptamer in the 3' trailer, an amber-mutated reporter gene (FLuc here), and WGE. The 3' premature tRNA probe was designed to be mainly degraded in the absence of the ligand, but to undergo rapid 3' processing in its presence to become a mature and highly-active amber suppressor tRNA (**t86**), through the structural rearrangement induced by aptamer-ligand complex formation in the 3' trailer. Because the resulting stable suppressor enables complete expression of the reporter protein though the unprocessed probe does not, the ligand can be indirectly detected with the activity of the translated reporter.

A key to the success of the ligand-dependent modulation of the 3' processing is appropriate configuration of the surrounding bases of the aptamer segment in the 3' trailer. In fact, we constructed a tRNA probe for detecting theophylline (**theo(th1)-MS(4)**) by optimizing these bases around its aptamer. However, their optimization is no longer needed for other aptamers because these bases are mechanistically independent from the aptamer segment and thus can be shared with other probes. In other words, tRNA probes for other ligands can be readily created simply by exchanging the aptamer segment, as in the case of **tc(th1)-MS(4)** and **FMN(th1)-MS(4)**. Although the only aptamers that are acceptable for implantation are aptamers that require the proximity of both ends for binding, such aptamers can be obtained through *in vitro* selection using an appropriate initial pool and the subsequent minimization. In light of the modulating mechanism, a longer and/or more rapidly binding aptamer would be ideal for detecting its ligand with a higher sensitivity.

The present probes are designed to actively exploit the endogenous 3' processing of tRNA, which generally occurs in the nucleoplasm in eukaryotes.[19] Therefore, they may be applicable for detecting an analyte in the eukaryotic nucleoplasm by using an appropriate sup-tRNA framework with some optimization. Moreover, these tRNA-based switches can upregulate the expression of any amber-mutated gene in response to the ligands in a trans manner. They are thus also promising as trans-acting, ligand-dependent gene regulators[20] for use in the recently expanding areas of cell-free or cell-based synthetic biology.

This work was supported by JSPS KAKENHI Grant Number 25708027.

Notes and references

1 (a) C. Tuerk and L. Gold, *Science*, 1990, **249**, 505; (b) A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818.
2 (a) S. Song, L. Wang, J. Li, J. Zhao and C. Fan, *Trends Anal. Chem.*, 2008, **27**, 108; (b) M. Famulok and G. Mayer, *Acc. Chem. Res.*, 2011, **44**, 1349; (c) C. Feng, S. Dai and L. Wang, *Biosens. Bioelectron.*, 2014, **59**, 64.
3 A. D. Keefe, S. Pai and A. Ellington, *Nat. Rev. Drug Discov.*, 2010, **9**, 537.
4 (a) M. N. Stojanovic and D. M. Kolpashchikov, *J. Am. Chem. Soc.*, 2004, **126**, 9266; (b) J. Liu and Y. Lu, *Angew. Chem. Int. Ed. Engl.*, 2005, **45**, 90; (c) S. Tombelli, M. Minunni, E. Luzi and M. Mascini, *Bioelectrochemistry*, 2005, **67**, 135; (d) Y. Xiao, T. Uzawa, R. J. White, D. DeMartini and K. W. Plaxco, *Electroanalysis*, 2009, 21, 1267; (e) C. Furutani, K. Shinomiya, Y. Aoyama, K. Yamada and S. Sando, *Mol. Biosystems*, 2010, **6**, 1569; (f) A. Ogawa, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 155; (g) P. S. Lau, C. K. Lai and Y. Li, *ChemBioChem*, 2013, **14**, 987; (h) E. Sharon, N. Enkin, H. B. Albada and I. Willner, *Chem. Commun.*, 2015, **51**, 1100.
5 (a) A. Ogawa and M. Maeda, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 3156; (b) A. Ogawa, *ChemBioChem*, 2009, **10**, 2465; (c) A. Ogawa, *RNA*, 2011, **17**, 478; (d) A. Ogawa, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 1639; (e) A. Ogawa, *ChemBioChem*, 2013, **14**, 1539.
6 A. Ogawa and M. Maeda, *ChemBioChem*, 2008, **9**, 2204.
7 K. Madin, T. Sawasaki, T. Ogasawara and Y. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 559.
8 A. Ogawa, Y. Doi and N. Matsushita, *Org. Biomol. Chem.*, 2011, **9**, 8495.
9 A. Ogawa and Y. Doi, *Org. Biomol. Chem.*, 2015, **13**, 1008.
10 This high suppression efficiency comparable to that of **t86** was likely attributable to the moderately long 3' trailer being

favorable for RNase Z-mediated cleavage. It has been reported that a much longer 3' trailer adversely affects the RNase Z-mediated cleavage; see: S. Schiffer, M. Helm, A. Théobald-Dietrich, R. Giegé and A. Marchfelder, *Biochemistry*, 2001, **40**, 8264.

11 (a) R. D. Jenison, S. C. Gill, A. Pardi and B. Polisky, *Science*, 1994, **263**, 1425; (b) G. R. Zimmermann, R. D. Jenison, C. L. Wick, J.-P. Simorre and A. Pardi, *Nat. Struct. Biol.*, 1997, **4**, 644.

12 The suppression efficiency in the OFF state (2.6% of that by **3pt**) was as low as those of the th-free probes.

13 Given the fact that the maturation efficiency of a 5' premature sup-tRNA (**5pt**) was 21% with 1-h incubation in WGE, the structural change was expected to be faster than the 5' processing ($k = 2.7 \times 10^{-2}$ min^{-1}), which was slightly slower than the degradation of **3pt** ($k = 5\sim7 \times 10^{-2}$ min^{-1}).[9]

14 M. Müller, J. E. Weigand, O. Weichenrieder and B. Suess, *Nucleic Acids Res.*, 2006, **34**, 2607.

15 P. Burgstaller and M. Famulok, *Angew. Chem. Int. Ed. Engl.*, 1994, **33**, 1084.

16 Another explanation is that a tRNA probe with a much longer 3' trailer is less susceptible to the RNase-Z mediated cleavage.[10] However, **tc(th1)-MS(4)** was well processed in the ON state, even at low concentrations of the ligand.

17 Nonetheless, the suppression efficiency in the OFF state was still lower than that of **3pt-IS(14)**, due to the longer 3' dangling end.

18 However, a much longer aptamer may slightly inhibit the RNase Z-mediated cleavage in the ON state.[10]

19 E. M. Phizicky and A. K. Hopper, *Genes Dev.,* 2010, **24**, 1832.

20 B. Berscheider, M. Wieland, M. Rubini and J. S. Hartig, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 7564.

21 Both ligands adversely affected the suppression efficiency at higher concentrations, unlike in the case of theophylline.

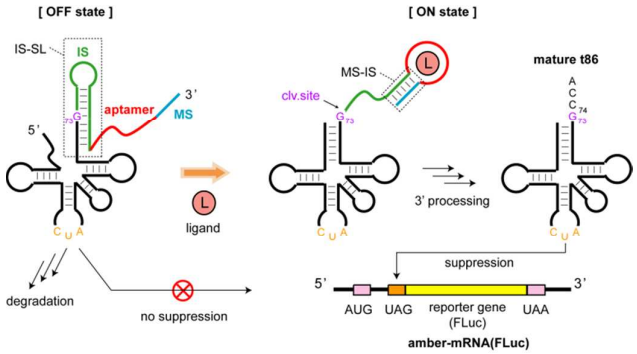
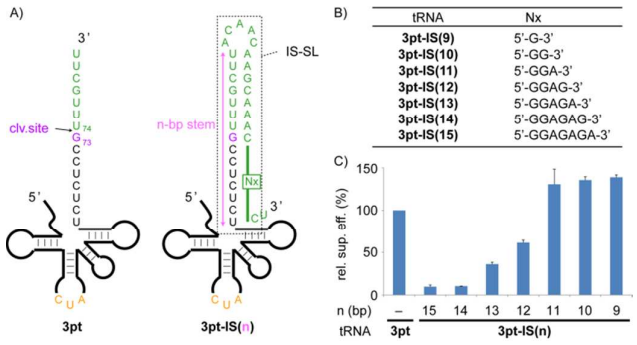
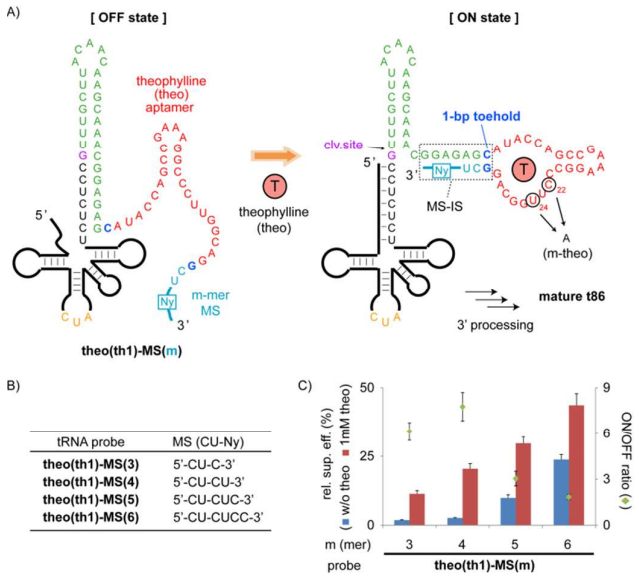


Fig. 1 Schematic illustration of the tRNA switch-based aptasensors in this study. The structure-distorted 3' premature tRNA probe degrades without undergoing processing (i.e., without suppressing the amber codon) in the absence of the ligand in WGE (left, OFF state). In contrast, in the presence of the ligand, the probe is processed through the structural restoration induced by the ligand binding to become a stable mature sup-tRNA (**t86**), which highly suppresses the amber codon in **amber-mRNA(FLuc)** to express the easily detectable FLuc (right, ON state). RNase Z cleaves the phosphodiester bond (clv. site) between G_{73} and N_{74} during the 3' processing.



tRNA	Nx
3pt-IS(9)	5'-G-3'
3pt-IS(10)	5'-GG-3'
3pt-IS(11)	5'-GGA-3'
3pt-IS(12)	5'-GGAG-3'
3pt-IS(13)	5'-GGAGA-3'
3pt-IS(14)	5'-GGAGAG-3'
3pt-IS(15)	5'-GGAGAGA-3'

Fig. 2 Optimization of the inhibitory sequence (IS) for repressing the 3' processing. (A) Schematic diagram of the positive-control 3' premature tRNA (left, **3pt**), which has been known to be rapidly processed in WGE,[9] and its derivatives with various lengths of IS (right, **3pt-IS(n)**), wherein n represents the stem length of IS-SL. A dinucleotide CU that does not interfere with the IS-SL formation was added to the 3' terminus for precise PCR-amplification of the templates for **3pt-IS(n)**. (B) The sequences of Nx in **3pt-IS(n)**. (C) The relative suppression efficiencies of **3pt-IS(n)**.



tRNA probe	MS (CU-Ny)
theo(th1)-MS(3)	5'-CU-C-3'
theo(th1)-MS(4)	5'-CU-CU-3'
theo(th1)-MS(5)	5'-CU-CUC-3'
theo(th1)-MS(6)	5'-CU-CUCC-3'

Fig. 3 Theophylline-detecting tRNA probes with a 1-bp toehold (**theo(th1)-MS(m)**), wherein m represents the base length of the modulator sequence (MS). (A) Schematic diagram of the theophylline-dependent structural switches of **theo(th1)-MS(m)**. Nucleotides in the aptamer domain are numbered according to a previous report.[11b] (B) The sequences of the MS including Ny in **theo(th1)-MS(m)**. (C) The relative suppression efficiencies of **theo(th1)-MS(m)** in the absence or presence of 1 mM theophylline (theo) and their ratios (ON/OFF).

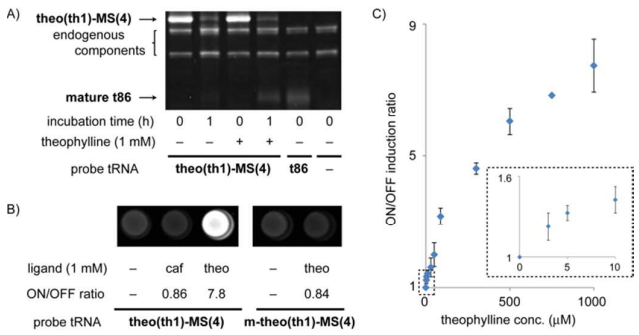


Fig. 4 Characterization of the optimal theophylline probe, **theo(th1)-MS(4)**. (A) Gel analyses of **theo(th1)-MS(4)** incubated in the absence or presence of 1 mM theophylline for 0 or 1 h in WGE. (B) Chemiluminescence images of translated FLuc in the absence or presence of 1 mM theophylline (theo) or caffeine (caf) when using **theo(th1)-MS(4)** (left) or **m-theo(th1)-MS(4)** with mutations in the aptamer domain (right) as a probe. (C) The ON/OFF induction ratios of **theo(th1)-MS(4)** at various concentrations of theophylline. (*Dotted box*) Expansion at low concentrations.

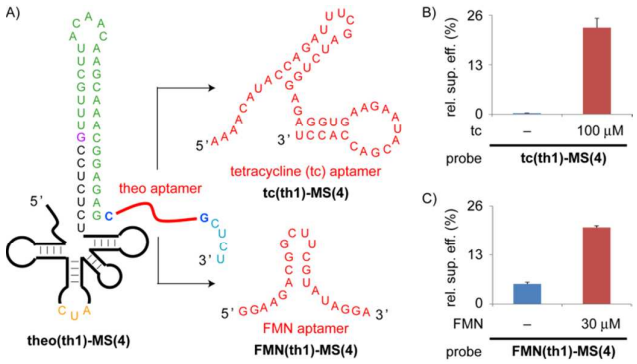


Fig. 5 Versatility of the tRNA-probe design. (A) Schematic diagram for replacement of the aptamer domain of **theo(th1)-MS(4)** with a tetracycline-binding aptamer (upper right) or FMN-binding one (lower right) to construct the corresponding tRNA probes (**tc(th1)-MS(4)** and **FMN(th1)-MS(4)**, respectively). (B) and (C) The relative suppression efficiencies of **tc(th1)-MS(4)** and **FMN(th1)-MS(4)**, respectively, in the absence or presence of 100 μM tetracycline (tc) and 30 μM FMN, respectively.[21]