Synthetic Riboswitches

DOI: 10.1002/anie.201303370

A Synthetic Riboswitch that Operates using a Rationally Designed Ligand–RNA Pair**

Chikara Dohno, Izumi Kohyama, Maki Kimura, Masaki Hagihara, and Kazuhiko Nakatani**

Nature has used small molecules that bind to specific RNA sequences for the regulation of biological functions. Functional RNAs termed riboswitches are *cis*-regulatory elements found in untranslated regions (UTRs) of mRNAs.[1–3] Ligand binding to the aptamer domain of the riboswitch triggers a structural change that results in the modulation of gene expression in a concentration-dependent manner. The regulatory mechanism involving riboswitches is a new approach in which small molecules directly regulate gene expression without the aid of proteins. Artificial riboswitches operating with designed ligand–RNA pairs can, in principle, provide designer gene regulation systems, which are promising for the programming of cellular behavior, detection of bioactive molecules, and controlled expression of therapeutic genes.[4–6]

Several reports have described artificial riboswitches that regulate gene expression in response to specific ligands.[4–10] The RNA aptamer domain that binds to the ligand and the concomitant change in its secondary structure are indispensable for the creation of new riboswitches driven by selected ligands. To date, the ligand–aptamer pair required for this purpose has been obtained using a directed evolution method called systematic evolution of ligands by exponential enrichment (SELEX) or in vitro selection that identifies favorable RNA molecules from a large RNA library[11,12] and reengineering of preexisting aptamer sequences.[10] Along with these molecular biology-based approaches for creating artificial riboswitches, chemistry-based approaches using a rational design of ligand molecules that bind to the predetermined RNA sequences expand the scope and potential of artificial riboswitches. However, the chemistry-based approaches have been hindered by difficulties in the design of the ligand that binds to specific RNA sequences and structures.[13–16] Herein, we present the construction of an artificial riboswitch system based on the rational design of a ligand–RNA pair without using any molecular biology-based selection processes. The success of the chemistry-based approach to the generation of artificial riboswitches described herein should promote a deeper understanding of the fundamental properties of riboswitches for biological and medical purposes.

We have previously developed a series of DNA-mismatch-binding ligands that can induce a transition from single-stranded to double-stranded DNA upon the binding.[17–19] If this system is applicable to RNA, the ligand-induced structural change can be used effectively as an aptamer domain of riboswitches that sense the ligand. The mismatch-binding ligands recognize Watson–Crick surface of DNA bases by forming fully matched hydrogen bonds.[19–22] Although the ligand-accessible surface of Watson–Crick base pairs is almost identical between DNA and RNA, most of the DNA-mismatch-binding ligands developed by us exhibit a reduced or abolished binding affinity to the corresponding RNA mismatches.

We focused on a naphthyridine carbamate tetramer derivative, *Z*-NCTS, as a potential RNA-mismatch-binding ligand. *Z*-NCTS consists of two units of a naphthyridine carbamate dimer (NCD) connected by a rigid *Z*-stilbene linker (Figure 1a).[22] *Z*-NCTS binds to the GG mismatch in the d(CGG)/d(CGG) sequence of DNA with an entropically favored 1:1 stoichiometry and strongly stabilizes the resulting double-stranded DNA (Figure 1b).[22] Four naphthyridine moieties in a single *Z*-NCTS molecule recognize four guanine bases in the d(CGG)/d(CGG) sequence by forming fully

[*] Dr. C. Dohno, I. Kohyama, M. Kimura, Dr. M. Hagihara, Prof. K. Nakatani
The Institute of Scientific and Industrial Research, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka 567-0047 (Japan)
E-mail: cdohno@sanken.osaka-u.ac.jp
nakatani@sanken.osaka-u.ac.jp
Dr. C. Dohno
PRESTO (Japan) Science and Technology Agency
4-1-8 Honcho Kawaguchi, Saitama 332-0012 (Japan)
Dr. M. Hagihara
Graduate School of Science and Technology, Hirosaki University
3 Bunkyo-cho, Hirosaki, Aomori 036-8561 (Japan)

[**] This work was supported by Grant in Aid for Scientific Research on Innovative Areas (21200042) from the Japan Society for the Promotion of Science (JSPS) and Sumitomo Foundation for C.D. and Grant in Aid for Scientific Research (A) (23241073) and the Advanced research for medical products Mining Program of the National Institute of Biomedical Innovation (10–22) for K.N.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201303370.

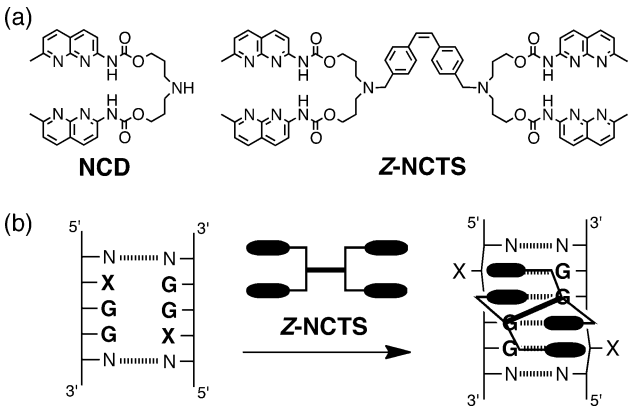

Figure 1. a) Chemical structures of NCD and *Z*-NCTS. b) Illustration of the binding of *Z*-NCTS to d(XGG)/d(XGG) (X = C, G, T, or A).

matched hydrogen bonds. First, we evaluated the binding of *Z*-NCTS to the RNA GG mismatch in the r(CGG)/r(CGG) sequence by measuring the melting temperature (T_m) of a 11-mer double-stranded RNA (5′-r(CUAA *CGG* AAUG)-3′/5′-r(CAUU *CGG* UUAG)-3′ (Figure 2a). The presence of *Z*-NCTS only slightly increased the T_m value of the RNA duplex ($\Delta T_m = +2.1\,°C$; Supporting Information, Table S1), indicating its weak binding to the r(CGG)/r(CGG) sequence. This result was in marked contrast with the large ΔT_m value of $+21.4\,°C$ observed for the corresponding DNA duplex.[22] In our related studies, the binding of the ligand to the DNA mismatches was accompanied by helix unwinding and base-pair rearrangements at the binding site (Figure 1b).[21] Such structural changes would be less favorable in RNA, because the A-form RNA duplex is more stable than is a duplex DNA with the same sequence and tolerates the presence of mismatched base pairs.[23,24]

Because the single GG mismatch in the RNA duplex did not render a good binding pocket for *Z*-NCTS, we then introduced a perturbation of the local structure near the mismatched site to gain a toehold for the binding of the ligand and to generate a sufficient free energy change upon binding. Thus, the two GC base pairs flanking the GG mismatch in the r(CGG)/r(CGG) sequence were replaced with non-Watson–Crick GU base pairs, providing the less stable helix containing a r(UGG)/r(UGG) sequence. The perturbation strategy was effective, as shown by the increase in T_m from 14.0 to 41.9 °C observed for the RNA duplex containing r(UGG)/r(UGG) after the addition of *Z*-NCTS ($\Delta T_m = +27.9\,°C$) (Figure 2b). To gain insight into the structural guideline for the design of RNA-binding ligands, we assessed the binding of NCD, which is one of the building blocks of *Z*-NCTS, to the r(UGG)/r(UGG) and found a much decreased ΔT_m value of $+12.6\,°C$ (Supporting Information, Figure S1). A geometrical isomer of *Z*-NCTS (*E*-NCTS) showed much lower affinity to the RNA.

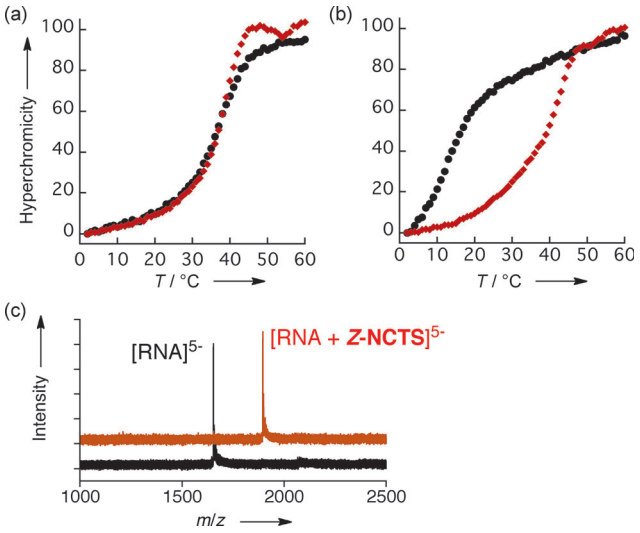

Figure 2. Thermal melting profiles of an 11-mer RNA duplex (4.5 μM) in the absence (black) and presence (red) of *Z*-NCTS (11.4 μM). a) r(CGG)/r(CGG) sequence. b) r(UGG)/r(UGG) sequence. c) CSI-TOF/MS of a 26-mer RNA containing an r(UGG)/r(UGG) sequence (20 μM) in the absence (black) and presence (red) of *Z*-NCTS (40 μM).

Among the naphthyridine-based ligands tested, *Z*-NCTS exhibited the highest ΔT_m value for the GG mismatch in the r(UGG)/r(UGG) sequence. The exceptionally high ΔT_m value obtained for *Z*-NCTS can be attributed to the covalent connection of NCD to the proper linker. *Z*-NCTS also increased the T_m values of RNA duplexes containing other three consecutive mismatch sequences, such as AGG and GGG (Supporting Information, Figure S2, Table S2). The requirement of the tandem mismatch site probably reflects the more stable and rigid nature of RNA duplexes compared with DNA duplexes.

The binding between *Z*-NCTS and r(XGG)/r(XGG) was analyzed using cold-spray ionization time-of-flight mass spectrometry (CSI-TOF/MS).[21,22,25] A 26-mer hairpin RNA containing a r(UGG)/r(UGG) site gave a single major peak corresponding to the 5^- ion of the RNA (Figure 2c, black). In the presence of *Z*-NCTS, an ion peak corresponding to a 1:1 complex (calcd 1894.5, found 1895.2) appeared, together with the loss of the free RNA ion peak (Figure 2c, red). Similar data were obtained for r(AGG)/r(AGG) sequence, whereas 2:1 (*Z*-NCTS:RNA) binding prevailed for r(GGG)/r(GGG) sequence (Supporting Information, Figure S3). These data showed that *Z*-NCTS bound to r(XGG)/r(XGG) (X = U and A) preferentially with 1:1 stoichiometry. Based on the 1:1 binding model, we calculated the dissociation constant (K_d) of the *Z*-NCTS-r(XGG)/r(XGG) complex by circular dichroism (CD) titration experiments (Supporting Information, Figure S4). *Z*-NCTS bound more strongly to r(UGG)/r(UGG) with a submicromolar K_d value (3×10^{-7} M) than r(AGG)/r(AGG) (2×10^{-6} M). An isothermal titration calorimetry (ITC) experiment for the r(UGG)/r(UGG) sequence provided the apparent K_d value of the same order (2.7×10^{-7} M; Supporting Information, Figure S5).[26]

Having developed an RNA-binding ligand that was selective for the r(XGG)/r(XGG) sequence, we then attempted to design a *Z*-NCTS-responsive regulatory RNA element. The key principle used for the design of this RNA sequence was based on the finding that binding to *Z*-NCTS drastically increased the thermodynamic stability of the target RNA duplex containing an r(XGG)/r(XGG) sequence (Figure 2b). A pair of unhybridized RNA regions containing the XGG sequence were stuck together into a thermodynamically stable double-stranded form with the aid of *Z*-NCTS. To link the binding to *Z*-NCTS with translational regulation, we incorporated the *Z*-NCTS-dependent hairpin-forming sequence into the synthetic riboswitch, as an aptamer domain, and integrated a regulatory mechanism consisting of blocking and sequestering ribosome binding sites (RBSs) (Figure 3a).[4–6,27,28] The hairpin-forming site incorporated into the 5′-UTR consists of the newly introduced UGG sequence and a guanine-rich RBS containing the AGG sequence. T_m measurement indicated that the resulting r(UGG)/r(AGG) sequence can be a binding site of *Z*-NCTS (Supporting Information, Table S2). In the presence of *Z*-NCTS, the stabilized hairpin structure would suppress translation initiation by blocking the access of ribosomes to the RBS (Figure 3a).

We prepared a series of mRNAs (lucWT, lucG2, and lucG3) encoding firefly luciferase for expression in a bacterial

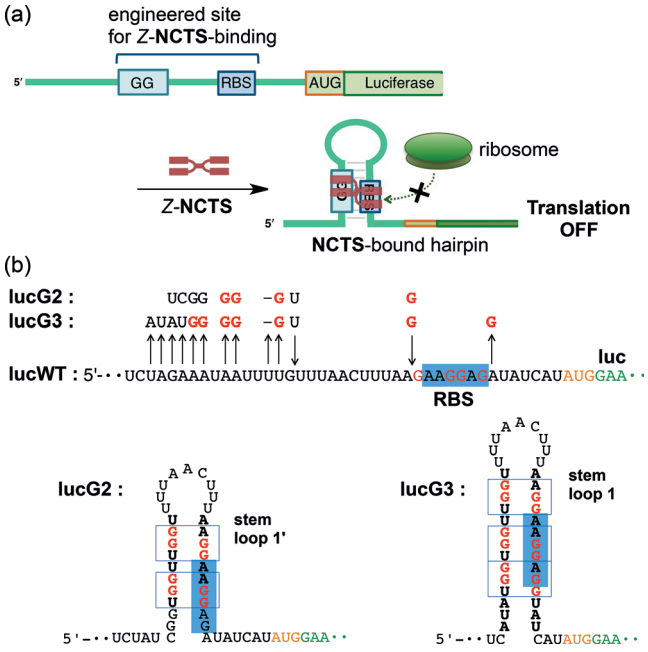

Figure 3. a) Illustration of *Z*-NCTS-mediated translational regulation. The template mRNA contained a multiple guanine site (boxed GG) and an RBS upstream of the luciferase coding region, which can form a hairpin structure only in the presence of *Z*-NCTS, resulting in the inhibition of translation initiation. b) Partial sequences in the 5′-UTR of the mRNAs and the expected hairpin secondary structure after binding to the *Z*-NCTS binding sites (box). Wild-type lucWT was engineered into mutant lucG2 and lucG3, as indicated by the upward arrows (substitution), downward arrows (insertion) and the minus (−) symbols (deletion). The RBS is indicated by a blue box. The guanines in the *Z*-NCTS-binding sites are shown in red.

system. lucG2 and lucG3 were 5′-UTR-engineered mRNAs containing multiple UGG sequences upstream of the RBS and were designed to form a hairpin structure involving the RBS only in the presence of *Z*-NCTS, albeit with different efficiencies. lucG3 had three potential *Z*-NCTS-binding sites (r(UGG)/r(AGG)), whereas lucG2 had two of these sites (Figure 3b). The multiple binding sites are expected to be effective for ligand-induced structural change from single-stranded form to hairpin structure.[18,19] Indeed, thermal melting profiles of 5′-UTR sequences of lucG2 and lucG3 showed the presence of stable secondary structure only in the presence of *Z*-NCTS (Supporting Information, Figure S6).

In vitro translation experiments were performed using the template mRNAs and a reconstituted *E. coli* cell-free translation system (PURE system).[29] The translational efficiency of the mRNA was evaluated by measuring the chemiluminescence generated by the luciferase-catalyzed reaction. The results of the luciferase assay are summarized in Figure 4. Luminescence intensity decreased with increasing concentration of *Z*-NCTS; the magnitude of this variation was dependent on the RNA sequences. Luminescence intensity of the lucG3 RNA template decreased by half in the presence of 5 μM *Z*-NCTS (Figure 4c). Under the same conditions, luciferase expression from lucWT was almost unaffected, whereas expression from lucG2 was decreased by 15%.[30] With a higher concentration of *Z*-NCTS, inhibition of translation became more pronounced for lucG3, but an inhibitory effect was also detectable for lucWT possibly due to unintended binding of *Z*-NCTS to G-rich sites of RNAs.[30] GFP reporter RNAs with the same 5′-UTR sequences showed a similar response to *Z*-NCTS (Supporting Information, Figure S7), indicating the 5′-UTR sequences play a crucial role in the *Z*-NCTS-dependent translational regulation.

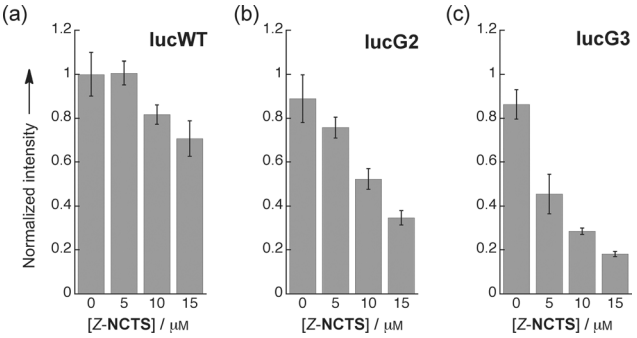

Figure 4. The luciferase assay of a) lucWT, b) lucG2, and c) lucG3 in the presence of *Z*-NCTS. In vitro translation experiments were performed for each template RNA in varying *Z*-NCTS concentrations (0, 5, 10, and 15 μM). The obtained chemiluminescence intensities were normalized to that of lucWT without *Z*-NCTS. Error bars represent the standard deviation of at least three independent experiments.

To clarify whether translational regulation was attributable to the *Z*-NCTS-induced formation of the hairpin structure in the 5′-UTR, we performed a reverse transcriptase (RTase) stop assay.[31] In the presence and absence of *Z*-NCTS, primer extension reactions on the template RNAs using RTase were analyzed by polyacrylamide gel electrophoresis (Figure 5). The presence of stable secondary structures in the template RNA interfered with the RTase reaction and was identified by paused bands on the gel. As all template RNAs have an intrinsic 5′-terminal stem-loop structure, which stabilizes mRNAs in *E. coli*,[32] the extension reaction mediated by RTase was partially blocked at that site, regardless of the presence of *Z*-NCTS (stem-loop 2; Supporting Information, Figure S8). An increase in *Z*-NCTS concentration led to the appearance of additional paused bands for lucG2 and lucG3, which indicated the existence of stable secondary structure induced by *Z*-NCTS. The locations of the RTase-paused sites were consistent with the 3′-end of the expected *Z*-NCTS-bound hairpin structure (stem-loop 1 and 1′, Figure 3b). The inhibitory effect of *Z*-NCTS on RTase was more pronounced for lucG3 than it was for lucG2, which is consistent with the greater stability of the *Z*-NCTS-bound hairpin in lucG3 vs lucG2. These data indicate that *Z*-NCTS binding produced a stable hairpin structure selectively at the intended site in the 5′-UTR, supporting the idea of *Z*-NCTS-mediated translational regulation.

In summary, we have successfully created a synthetic riboswitch using the chemistry-based design of a specific ligand molecule and its target RNA sequence. A ligand–aptamer pair, which is usually obtained using an in vitro selection technique, is an essential element in the creation of a designer riboswitch. This research showed clearly that such a ligand–aptamer pair could be obtained by the rational design of a synthetic RNA-binding ligand without the use of

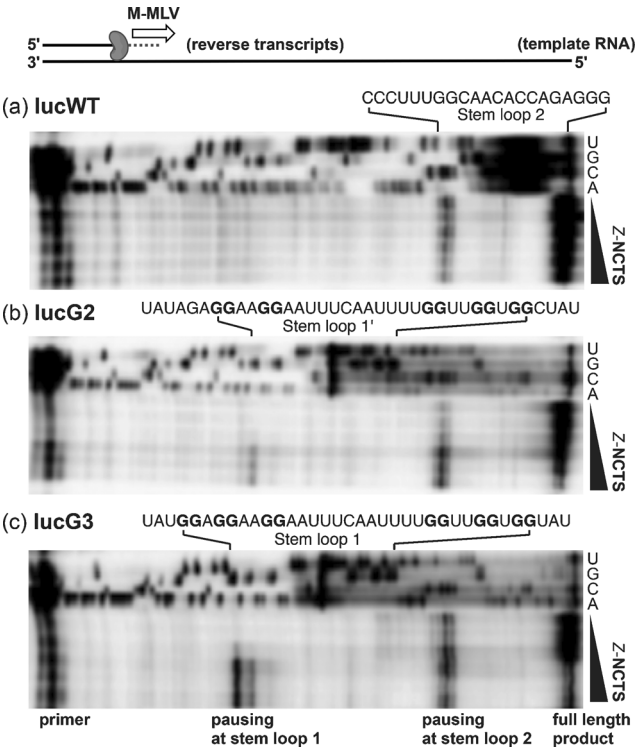

Figure 5. Reverse transcriptase (RTase) stop assay. The interruptions of RTase (M-MLV)-mediated primer extension reactions on each template RNA were analyzed using polyacrylamide gel electrophoresis with varying concentrations of *Z*-NCTS (0, 1, 2.5, 5, 7.5, and 10 μM). The expected pausing sites (stem-loops 1 and 1′ (*Z*-NCTS-dependent hairpin) and stem-loop 2 (intrinsic 5′-terminal hairpin)) are shown above and below the gel images. a) lucWT; b) lucG2; and c) lucG3. The lane markers U, G, C, and A indicate the bases on the template RNA.

selection methods. The chemistry- and molecular biology-based approaches are complementary, and both approaches will expand the flexibility of the design and application of riboswitches. At this stage, in contrast to the well-developed selection technique, there is still room for further growth and improvement in the generation of small ligand molecules that bind selectively to target RNA sequences and structures. The present study marked the first step toward the generation of designer riboswitches by a chemistry-based approach.

Received: April 21, 2013
Revised: July 4, 2013

Keywords: naphthyridine · riboswitches · RNA-binding ligand · synthetic ligands · translational regulation

[1] W. Winkler, A. Nahvi, R. R. Breaker, *Nature* **2002**, *419*, 952–956.
[2] a) M. Mandal, R. R. Breaker, *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 451–463; b) A. Roth, R. R. Breaker, *Annu. Rev. Biochem.* **2009**, *78*, 305–334.
[3] a) A. Serganov, E. Nudler, *Cell* **2013**, *152*, 17–24; b) A. Serganov, D. J. Patel, *Nat. Rev. Genet.* **2007**, *8*, 776–790.
[4] K. H. Link, R. R. Breaker, *Gene Ther.* **2009**, *16*, 1189–1201.
[5] M. Wieland, J. S. Hartig, *ChemBioChem* **2008**, *9*, 1873–1878.
[6] S. Topp, J. P. Gallivan, *ACS Chem. Biol.* **2010**, *5*, 139–148.
[7] G. Werstuck, M. R. Green, *Science* **1998**, *282*, 296–298.
[8] I. Harvey, P. Garneau, J. Pelletier, *RNA* **2002**, *8*, 452–463.
[9] a) B. Suess, S. Hanson, C. Berens, B. Fink, R. Schroeder, W. Hillen, *Nucleic Acids Res.* **2003**, *31*, 1853–1858; b) A. Wittmann, B. Suess, *FEBS Lett.* **2012**, *586*, 2076–2083.
[10] a) N. Dixon, J. N. Duncan, T. Geerlings, M. S. Dunstan, J. E. McCarthy, D. Leys, J. Micklefield, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2830–2835; b) N. Dixon, C. J. Robinson, T. Geerlings, J. N. Duncan, S. P. Drummond, J. Micklefield, *Angew. Chem.* **2012**, *124*, 3680–3684; *Angew. Chem. Int. Ed.* **2012**, *51*, 3620–3624.
[11] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818–822.
[12] C. Tuerk, L. Gold, *Science* **1990**, *249*, 505–510.
[13] J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171–1224.
[14] L. R. Guan, M. D. Disney, *ACS Chem. Biol.* **2012**, *7*, 73–86.
[15] Y. Tor, *ChemBioChem* **2003**, *4*, 998–1007.
[16] D. M. Chenoweth, J. L. Meier, P. B. Dervan, *Angew. Chem.* **2013**, *125*, 433–436; *Angew. Chem. Int. Ed.* **2013**, *52*, 415–418.
[17] a) T. Peng, C. Dohno, K. Nakatani, *Angew. Chem.* **2006**, *118*, 5751–5754; *Angew. Chem. Int. Ed.* **2006**, *45*, 5623–5626; b) C. Dohno, H. Atsumi, K. Nakatani, *Chem. Commun.* **2011**, *47*, 3499–3501.
[18] C. Dohno, S. N. Uno, K. Nakatani, *J. Am. Chem. Soc.* **2007**, *129*, 11898–11899.
[19] C. Dohno, K. Nakatani, *Chem. Soc. Rev.* **2011**, *40*, 5718–5729.
[20] K. Nakatani, *Bull. Chem. Soc. Jpn.* **2009**, *82*, 1055–1069.
[21] K. Nakatani, S. Hagihara, Y. Goto, A. Kobori, M. Hagihara, G. Hayashi, M. Kyo, M. Nomura, M. Mishima, C. Kojima, *Nat. Chem. Biol.* **2005**, *1*, 39–43.
[22] C. Dohno, I. Kohyama, C. Hong, K. Nakatani, *Nucleic Acids Res.* **2012**, *40*, 2771–2781.
[23] a) P. J. Hagerman, *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 139–156; b) A. Noy, A. Perez, F. Lankas, F. J. Luque, M. Orozco, *J. Mol. Biol.* **2004**, *343*, 627–638.
[24] R. Kierzek, M. E. Burkard, D. H. Turner, *Biochemistry* **1999**, *38*, 14214–14223.
[25] a) K. Yamaguchi, *J. Mass Spectrom.* **2003**, *38*, 473–490; b) K. Nakatani, S. Hagihara, S. Sando, S. Sakamoto, K. Yamaguchi, C. Maesawa, I. Saito, *J. Am. Chem. Soc.* **2003**, *125*, 662–666.
[26] The apparent stoichiometry (*n*) was calculated to be 1.4, suggesting the existence of an additional minor binding mode under the ITC conditions.
[27] A. Serganov, Y. R. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A. T. Phan, C. Hobartner, R. Micura, R. R. Breaker, D. J. Patel, *Chem. Biol.* **2004**, *11*, 1729–1741.
[28] B. Suess, B. Fink, C. Berens, R. Stentz, W. Hillen, *Nucleic Acids Res.* **2004**, *32*, 1610–1614.
[29] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751–755.
[30] To minimize tube-to-tube variation, the translation experiments in Figure 4 were carried out with the same lot of translation reagent. To facilitate more accurate comparison between the 5′-UTR sequences, a dual reporter system was employed (Supporting Information, Figure S7). Under the conditions, *Z*-NCTS-dependent translational regulation was again evident for lucG3 and that for lucWT was much less pronounced.
[31] a) M. Hagihara, L. Yamauchi, A. Seo, K. Yoneda, M. Senda, K. Nakatani, *J. Am. Chem. Soc.* **2010**, *132*, 11171–11178; b) M. Hagihara, H. He, K. Nakatani, *ChemBioChem* **2011**, *12*, 1686–1689.
[32] S. A. Emory, P. Bouvet, J. G. Belasco, *Genes Dev.* **1992**, *6*, 135–148.

