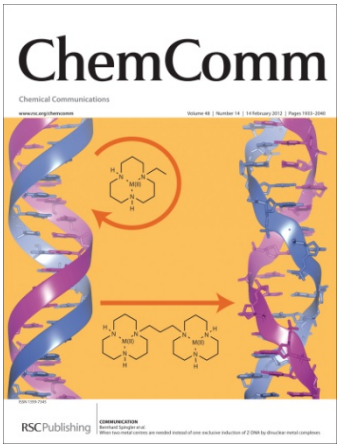
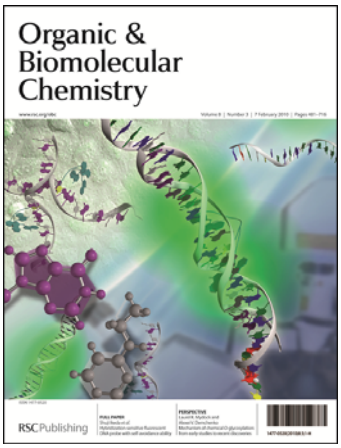




This article is part of the

Nucleic acids: new life, new materials

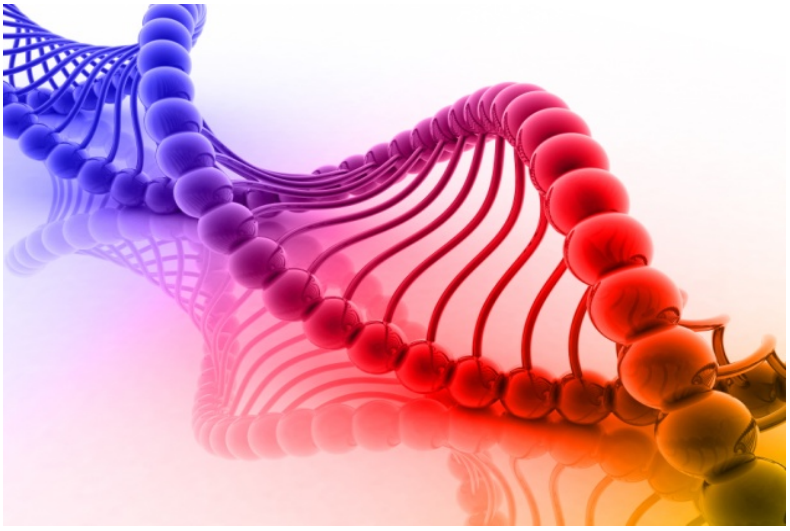
web-themed issue

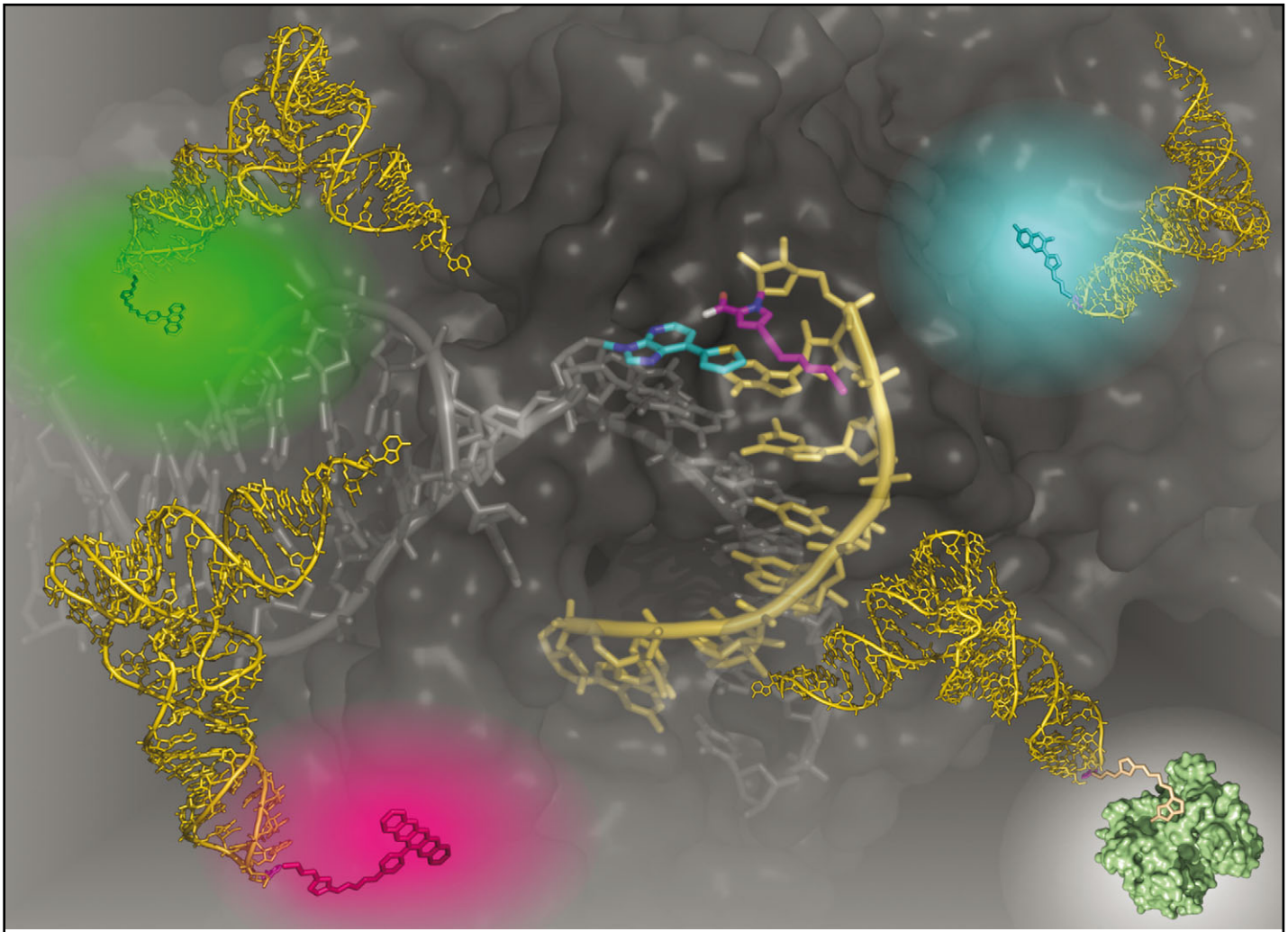
Guest edited by:

Mike Gait Medical Research Council, Cambridge, UK	**Ned Seeman** New York University, USA	**David Liu** Harvard University, USA	**Oliver Seitz** Humboldt-Universität zu Berlin, Germany	**Makoto Komiyama** University of Tsukuba, Japan	**Jason Micklefield** University of Manchester, UK

All articles in this issue will be gathered online at

www.rsc.org/nucleic_acids



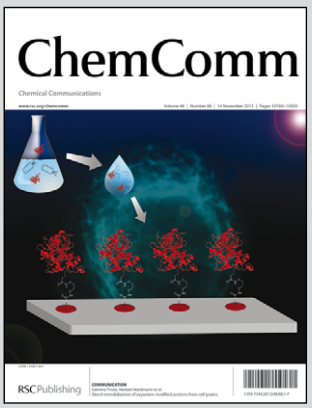


Showcasing research of Dr. Takumi Ishizuka *et al*. from the Nucleic Acid Synthetic Biology Research Team of Dr. Ichiro Hirao, Systems and Structural Biology Center (SSBC), RIKEN, Japan.

Site-specific functionalization of RNA molecules by an unnatural base pair transcription system *via* click chemistry

A site-specific labeling method for large RNA molecules, by combining the unnatural base pair transcription system and the post-transcriptional modification using click chemistry has been developed.

As featured in:



See Ichiro Hirao *et al*., *Chem. Commun*., 2012, **48**, 10835.

RSC Publishing

www.rsc.org/chemcomm

Registered Charity Number 207890

ChemComm

Dynamic Article Links

Cite this: *Chem. Commun.*, 2012, **48**, 10835–10837

www.rsc.org/chemcomm

COMMUNICATION

Site-specific functionalization of RNA molecules by an unnatural base pair transcription system *via* click chemistry†‡§

Takumi Ishizuka,[a] Michiko Kimoto,[ab] Akira Sato[a] and Ichiro Hirao*[ab]

Received 30th August 2012, Accepted 18th September 2012
DOI: 10.1039/c2cc36293g

Toward new biotechnology by genetic alphabet expansion, we developed an efficient site-specific labeling method for large RNA molecules. The combination of unnatural base pair transcription and post-transcriptional modification by click chemistry enables simple RNA labeling with a wide variety of functional groups at desired positions, in a one-pot reaction.

The interest of many researchers has been focused on the development of methodologies for site-specific modifications at defined positions in functional RNA molecules, which would be useful to clarify their biological functions. The site-specific modification of large functional RNA molecules, such as coding, non-coding, and artificial RNA species including riboswitches, ribozymes and aptamers, has become an attractive topic because of its potential to enable the simple visualization of their local structural properties in solution.[1] Although a combination of direct chemical synthesis and enzymatic ligation is the present mainstream approach, it is still a time-consuming and complex task to prepare RNA fragments with long chains. Other methods for terminal or internal labeling at a specific position of RNA are still restricted, and efficiently adjusting the labeling positions is difficult.[2]

A newly developed method, an unnatural base pair system, enables the site-specific incorporation of extra components into RNA molecules by transcription mediated by an artificial third base pair. If the unnatural base pair works along with the natural A–T(U) and G–C base pairs in transcription, then an unnatural base substrate could be incorporated as an extra base at specific positions in RNA molecules, opposite the counterpart of the unnatural base in a DNA template, by transcription.

Several unnatural base pairs for *in vitro* transcription have been reported so far.[1e,l,3] Among them, we developed a hydrophobic unnatural base pair between 7-(2-thienyl)-imidazo[4,5-*b*]pyridine (**Ds**) and pyrrole-2-carbaldehyde (**Pa**), as a third base pair working in replication and transcription (Fig. 1a).[3f,4] Both unnatural base substrates (**Ds**TP and **Pa**TP) are complementarily incorporated into transcripts by T7 RNA polymerase, with high selectivity (>95%). In addition, modified **Pa**TPs linked *via* a propynyl linker to functional molecules, such as amino groups, fluorophores, and biotin, are also site-specifically incorporated into RNA molecules. Long DNA templates containing **Ds** are easily prepared by combining chemical synthesis and PCR amplification using another pairing partner of **Ds**, 2-nitro-4-propynyl-pyrrole (**Px**), which exhibits high efficiency and selectivity in PCR.[5]

Here, we present a modified version of transcription using the **Ds–Pa** pair, involving a post-transcriptional modification by click chemistry[6] (Fig. 1a), a simple and versatile labeling method for large RNA molecules. Although the previous **Ds–Pa** pair transcription system is a potential tool for the

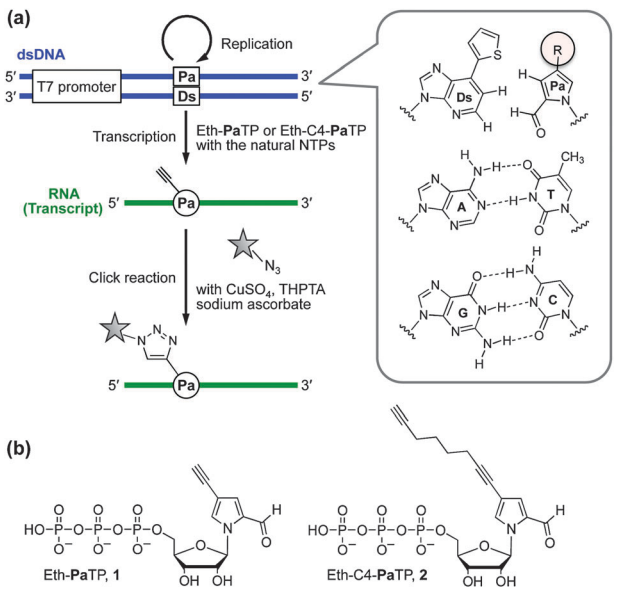


Fig. 1 An unnatural **Ds–Pa** base pair system for generating site-specifically labeled RNA molecules. (a) Overview of the combination of the specific transcription involving the **Ds–Pa** pair using ethynyl-modified **Pa** substrates and the post-transcriptional modification of the transcripts by the click reaction with functional azide reagents. R = ethynyl-modified position. (b) Chemical structures of Eth-**Pa**TP (**1**) and Eth-C4-**Pa**TP (**2**).

[a] *RIKEN Systems and Structural Biology Center (SSBC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan*
[b] *TagCyx Biotechnologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. E-mail: ihirao@riken.jp; Fax: +81-45-503-9645; Tel: +81-45-503-9644*

† The manuscript is dedicated to the memory of Professor H. Gobind Khorana.

‡ This article is part of the 'Nucleic acids: new life, new materials' web themed issue.

§ Electronic supplementary information (ESI) available: Detailed experimental methods and characterization data. See DOI: 10.1039/c2cc36293g

site-specific labeling of large RNA molecules, some modified **Pa** substrates with bulky functional groups reduce the transcription efficiency and selectivity.[4] Thus, we first incorporate **Pa** modified with ethynyl groups into RNA by T7 transcription mediated by the **Ds–Pa** pairing, and then post-transcriptionally modify the transcripts by copper(I)-catalyzed azide–alkyne cycloaddition, using azide derivatives with functional groups. The clickable **Pa** substrates are efficiently and site-specifically incorporated into transcripts opposite **Ds** by T7 RNA polymerase, and a wide variety of functional groups can be introduced by the following post-transcriptional modification.

We chemically synthesized two types of clickable **Pa** substrates, Eth-**Pa**TP and Eth-C4-**Pa**TP (**1** and **2**, Fig. 1b), which were prepared from the ribonucleoside derivatives of the 4-ethynyl-linked **Pa** base (Fig. S1, ESI§). The ribonucleoside, 1-(β-D-ribofuranosyl)-4-iodopyrrole-2-carbaldehyde, was reacted with trimethylsilyl acetylene or 1,7-octadiyne by Sonogashira coupling, using tetrakis(triphenylphosphine) palladium, copper(I) iodide and triethylamine. After deprotection of the trimethylsilyl group, both nucleoside derivatives were converted to the 5′-triphosphates by a conventional method.[7] The products were purified by DEAE Sephadex column chromatography and C18 reversed-phase HPLC. The structures of substrates **1** and **2** were confirmed by ^{1}H and ^{31}P NMR and high resolution mass spectrometry (see ESI§). Their molar absorption coefficients (ε) were determined by quantitative analysis of the phosphorus in the compound,[8] and the UV spectra are shown in Fig. S2 (see ESI§).

The incorporation efficiency and selectivity of substrates **1** and **2** by T7 transcription were examined by our conventional method, using chemically-synthesized DNA templates (35-mer) containing **Ds**.[3f,4] Full-length transcripts (17-mer) were obtained with 40–50% relative yields when using 1 mM substrate **1** or **2** and 1 mM natural NTPs, and the efficiency was boosted to 74% by increasing the concentration of **2** (2 mM) (Fig. S3, ESI§). The site-specific incorporation of each unnatural substrate in the transcripts was confirmed by a 2D-TLC nucleotide-composition analysis of the transcripts (Fig. S4, ESI§), and their incorporation selectivities were as high as 99%, irrespective of the differences between substrates **1** and **2** (Table S1, ESI§). In addition, in the transcription using a natural-base template (N = A in Fig. S4, ESI§) in the presence of substrate **1** or **2**, no unnatural base nucleotide corresponding to **1** or **2** was observed on the 2D-TLC of the transcript digestion. Therefore, these modified **Pa** substrates were site-specifically incorporated into the transcripts, only opposite **Ds** in the templates, by T7 RNA polymerase.

The post-transcriptional modification by the click reaction was first examined by using internally-^{32}P-labeled 17-mer transcripts (5 μM), containing Eth-**Pa** or Eth-C4-**Pa** at position 13 by treatments with fluorescent-dye azide reagents, such as Alexa Fluor 488 azide (488-N_3) and Alexa Fluor 594 azide (594-N_3) or Biotin-N_3 (each 0–250 μM). The click reaction was performed in the presence of a water-soluble copper(I) stabilizing ligand, tris[1-(3-hydroxypropyl-1*H*-1,2,3-triazol-4-yl)-methyl]amine (THPTA, 500 μM) (see ESI§),[6i] together with $CuSO_4$ (100 μM) and sodium ascorbate (1 mM) in 100 mM sodium phosphate buffer (pH 7.0) at 37 °C for 3 h, and the products were analyzed on a gel (Fig. S5 and S6, ESI§). The modification efficiency with Eth-C4-**Pa** was higher than that with Eth-**Pa**, and Eth-C4-**Pa** in the RNA fragment was completely modified with a 10-fold molar excess of 488-N_3, 594-N_3 or Biotin-N_3 reagent (Fig. S5, ESI§). Under the same conditions, transcripts consisting of only the natural bases were not modified (Fig. S5, ESI§ lanes 6 and 12), indicating that the ethynyl moiety attached to the **Pa** base was selectively modified with these azide reagents. We also confirmed the site-specific incorporation of the modified **Pa** substrates opposite **Ds** by the post-transcriptional modification, using transcripts synthesized in the presence of either substrate **1** or **2** with the natural-base template (N = A in Fig. S6, ESI§), in which no modified products were observed.

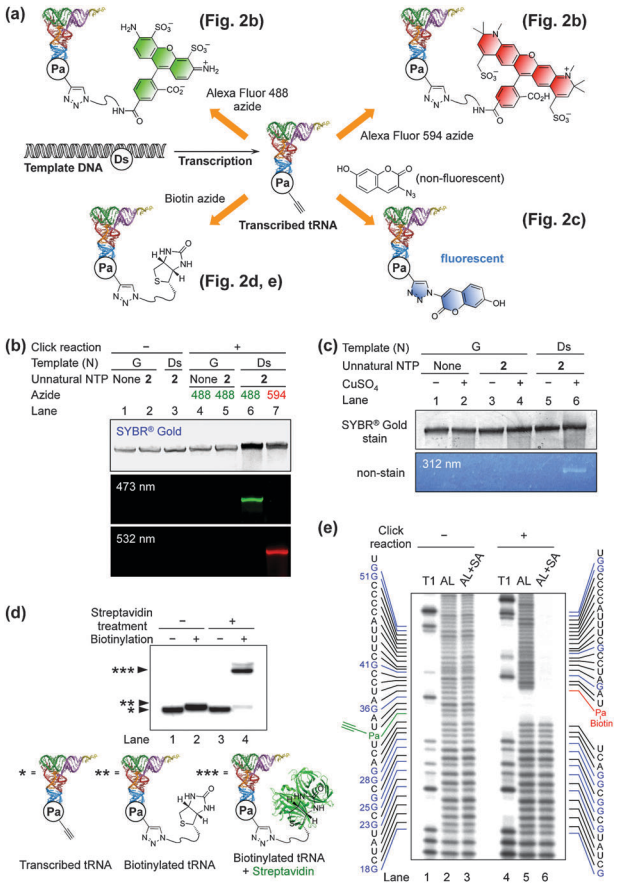


Fig. 2 Site-specific labeling of a tRNA transcript containing Eth-C4-**Pa**. (a) Experimental scheme for post-transcriptional labeling by the click reaction. Eth-C4-**Pa** was incorporated at position 33 in the tRNA molecule (75-mer) by T7 transcription using DNA templates (N = **Ds** or G), and the transcript was modified with several azide reagents. (b) The tRNA transcripts (75-mer) containing Eth-C4-**Pa** were labeled with 488-N_3 or 594-N_3. (c) Fluorogenic click reaction of the tRNA transcript. (d) Gel-mobility shift assay for the analysis of biotinylated transcripts with streptavidin. (e) Sequence analysis of the tRNA transcripts containing biotinylated or non-biotinylated **Pa** at position 33 in tRNA. The transcripts labeled with [γ-^{32}P]ATP were partially digested with RNase T1 (T1) or with alkali (AL). A portion of the partially alkali-digested transcripts was treated with streptavidin magnetic beads to capture the RNA fragments containing biotin-modified **Pa** (AL + SA).

The versatility of this transcription and post-transcriptional modification system was examined by the site-specific labeling of tRNA molecules (75-mer) (Fig. 2a). We prepared tRNA transcripts[9] containing Eth-C4-**Pa** at position 33, by T7 RNA

polymerase in the presence of 2 mM substrate **2** and the natural-base NTPs (Fig. S7, ESI§). The gel-purified transcript (3 μM) was then treated with either the 488-N_3 or 594-N_3 reagent (90 μM each), and the products were analyzed on a gel (Fig. 2b). Only the tRNA transcripts obtained from the **Ds**-containing DNA template with substrate **2** were labeled with the fluorescent azide reagents. The modification with non-fluorescent 3-azido-7-hydroxycoumarin[10] yielded its fluorescent product, which emitted at 480 nm upon 358 nm irradiation (Fig. 2c).

The site-specific modification of the tRNA molecules was confirmed by using biotinylated products (Fig. 2d and e). The modification of the tRNA transcript with the biotin azide reagent was examined by a gel-shift assay, and about 90% of the transcripts were shifted by binding streptavidin (Fig. 2d). The 5′-^{32}P-labeled transcripts with or without the biotin modification were digested under alkaline conditions, and half of the digested products were treated with streptavidin-coated magnetic beads to remove the biotinylated fragments. The remaining fragments were then analyzed by gel electrophoresis (Fig. 2e). The positions of each band on the ladder were determined by comparison with the ladder yielded by G-specific nuclease T1 digestion. On the lane without the streptavidin treatment of the digested products obtained by biotin modification, the fragment ladders larger than 33-mer were considerably shifted (Fig. 2e, lane 5) as compared to the control ladder (Fig. 2e, lane 2), and on the lane with the streptavidin treatment, the ladders larger than 33-mer disappeared (Fig. 2e, lane 6). These results indicated the site-specific biotin labeling at position 33 of the tRNA transcript.

In conclusion, we have developed a site-specific labeling method for large RNA molecules, by combining the unnatural base pair transcription system and the conventional click reaction of Cu(I)-catalyzed azide-alkyne cycloaddition. The modification efficiency is significantly improved, relative to that using amino-linked unnatural bases[3d] in oligonucleotides with *N*-hydroxysuccinimides. A variety of functional groups including bulky groups can be site-specifically and efficiently introduced into RNA molecules. This labeling method is performed by a one-pot T7 transcription reaction involving the **Ds–Pa** pair, the ethanol precipitation of transcripts, and click reaction with functional azide reagents (Fig. S8, ESI§). Thus, this labeling method could be used for a wide range of applications with site-specifically labeled RNA molecules.

This work was supported by the Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists (to T.I.). This work was also supported by Grants-in-Aid for Scientific Research (KAKENHI 19201046 to I.H. and 24710256 to M.K.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Notes and references

1 (*a*) S. C. Blanchard, R. L. Gonzalez, H. D. Kim, S. Chu and J. D. Puglisi, *Nat. Struct. Mol. Biol.*, 2004, **11**, 1008; (*b*) S. C. Blanchard, H. D. Kim, R. L. Gonzalez, Jr., J. D. Puglisi and S. Chu, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 12893; (*c*) M. Dorywalska, S. C. Blanchard, R. L. Gonzalez, H. D. Kim, S. Chu and J. D. Puglisi, *Nucleic Acids Res.*, 2005, **33**, 182; (*d*) T. Betteridge, H. Liu, H. Gamper, S. Kirillov, B. S. Cooperman and Y. M. Hou, *RNA*, 2007, **13**, 1594; (*e*) M. Kimoto, T. Mitsui, Y. Harada, A. Sato, S. Yokoyama and I. Hirao, *Nucleic Acids Res.*, 2007, **35**, 5360; (*f*) T. H. Lee, S. C. Blanchard, H. D. Kim, J. D. Puglisi and S. Chu, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 13661; (*g*) J. B. Munro, R. B. Altman, N. O'Connor and S. C. Blanchard, *Mol. Cell*, 2007, **25**, 505; (*h*) J. Fei, P. Kosuri, D. D. MacDougall and R. L. Gonzalez, Jr., *Mol. Cell*, 2008, **30**, 348; (*i*) C. M. Zhang, C. Liu, T. Christian, H. Gamper, J. Rozenski, D. Pan, J. B. Randolph, E. Wickstrom, B. S. Cooperman and Y. M. Hou, *RNA*, 2008, **14**, 2245; (*j*) C. Liu, T. Betteridge and Y. M. Hou, *Methods Enzymol.*, 2009, **469**, 69; (*k*) D. Pan, H. Qin and B. S. Cooperman, *RNA*, 2009, **15**, 346; (*l*) Y. Hikida, M. Kimoto, S. Yokoyama and I. Hirao, *Nat. Protocols*, 2010, **5**, 1312; (*m*) M. Kimoto and I. Hirao, *Methods Mol. Biol.*, 2010, **634**, 355; (*n*) S. Barhoom, J. Kaur, B. S. Cooperman, N. I. Smorodinsky, Z. Smilansky, M. Ehrlich and O. Elroy-Stein, *Nucleic Acids Res.*, 2011, **39**, e129; (*o*) S. Blouin, R. Chinnappan and D. A. Lafontaine, *Nucleic Acids Res.*, 2011, **39**, 3373; (*p*) J. Kaur, M. Raj and B. S. Cooperman, *RNA*, 2011, **17**, 1393; (*q*) S. Kellner, S. Seidu-Larry, J. Burhenne, Y. Motorin and M. Helm, *Nucleic Acids Res.*, 2011, **39**, 7348.

2 (*a*) M. Seki, P. Carninci, Y. Nishiyama, Y. Hayashizaki and K. Shinozaki, *Plant J.*, 1998, **15**, 707; (*b*) F. Huang, *RNA*, 2003, **9**, 1562.

3 (*a*) J. A. Piccirilli, T. Krauch, S. E. Moroney and S. A. Benner, *Nature*, 1990, **343**, 33; (*b*) Y. Tor and P. B. Dervan, *J. Am. Chem. Soc.*, 1993, **115**, 4461; (*c*) I. Hirao, T. Ohtsuki, T. Fujiwara, T. Mitsui, T. Yokogawa, T. Okuni, H. Nakayama, K. Takio, T. Yabuki, T. Kigawa, K. Kodama, K. Nishikawa and S. Yokoyama, *Nat. Biotechnol.*, 2002, **20**, 177; (*d*) R. Kawai, M. Kimoto, S. Ikeda, T. Mitsui, M. Endo, S. Yokoyama and I. Hirao, *J. Am. Chem. Soc.*, 2005, **127**, 17286; (*e*) K. Moriyama, M. Kimoto, T. Mitsui, S. Yokoyama and I. Hirao, *Nucleic Acids Res.*, 2005, **33**, e129; (*f*) I. Hirao, M. Kimoto, T. Mitsui, T. Fujiwara, R. Kawai, A. Sato, Y. Harada and S. Yokoyama, *Nat. Methods*, 2006, **3**, 729; (*g*) M. Kimoto, T. Mitsui, S. Yokoyama and I. Hirao, *J. Am. Chem. Soc.*, 2010, **132**, 4988; (*h*) Y. J. Seo, D. A. Malyshev, T. Lavergne, P. Ordoukhanian and F. E. Romesberg, *J. Am. Chem. Soc.*, 2011, **133**, 19878.

4 N. Morohashi, M. Kimoto, A. Sato, R. Kawai and I. Hirao, *Molecules*, 2012, **17**, 2855.

5 (*a*) M. Kimoto, R. Yamashige, S. Yokoyama and I. Hirao, *J. Nucleic Acids*, 2012, **2012**, 230943; (*b*) R. Yamashige, M. Kimoto, Y. Takezawa, A. Sato, T. Mitsui, S. Yokoyama and I. Hirao, *Nucleic Acids Res.*, 2012, **40**, 2793.

6 (*a*) V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596; (*b*) C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057; (*c*) G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen and T. Carell, *J. Am. Chem. Soc.*, 2006, **128**, 1398; (*d*) J. Gierlich, K. Gutsmiedl, P. M. Gramlich, A. Schmidt, G. A. Burley and T. Carell, *Chem.–Eur. J.*, 2007, **13**, 9486; (*e*) P. M. Gramlich, C. T. Wirges, J. Gierlich and T. Carell, *Org. Lett.*, 2008, **10**, 249; (*f*) C. Y. Jao and A. Salic, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15779; (*g*) A. Salic and T. J. Mitchison, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 2415; (*h*) S. H. Weisbrod and A. Marx, *Chem. Commun.*, 2008, 5675; (*i*) V. Hong, S. I. Presolski, C. Ma and M. G. Finn, *Angew. Chem., Int. Ed.*, 2009, **48**, 9879; (*j*) H. Macickova-Cahova and M. Hocek, *Nucleic Acids Res.*, 2009, **37**, 7612; (*k*) A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388; (*l*) E. Lallana, R. Riguera and E. Fernandez-Megia, *Angew. Chem., Int. Ed.*, 2011, **50**, 8794; (*m*) E. Paredes and S. R. Das, *ChemBioChem*, 2011, **12**, 125; (*n*) H. Rao, A. A. Sawant, A. A. Tanpure and S. G. Srivatsan, *Chem. Commun.*, 2012, **48**, 498; (*o*) H. Rao, A. A. Tanpure, A. A. Sawant and S. G. Srivatsan, *Nat. Protocols*, 2012, **7**, 1097.

7 M. Yoshikawa, T. Kato and T. Takenishi, *Tetrahedron Lett.*, 1967, **50**, 5065.

8 T. Ohtsuki, M. Kimoto, M. Ishikawa, T. Mitsui, I. Hirao and S. Yokoyama, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 4922.

9 R. Fukunaga, Y. Harada, I. Hirao and S. Yokoyama, *Biochem. Biophys. Res. Commun.*, 2008, **372**, 480.

10 K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill and Q. Wang, *Org. Lett.*, 2004, **6**, 4603.