Development of New Sensing Technologies toward Non-Invasive Nucleic Acid Analysis

Atsushi Narita

2008

Kyoto University

Preface

The study presented in this dissertation has been carried out under the direction of Professor Yasuhiro Aoyama at the Department of Synthetic Chemistry and Biological Chemistry of Kyoto University from April 2002 to March 2008. The study is focused on development of new sensing technologies toward in-cell nucleic acid analysis, based on biorecognics chemistry.

The author would like to express his sincere gratitude to Prof. Yasuhiro Aoyama for his kind guidance, valuable suggestions, and encouragement throughout this work. The author is deeply grateful to Assistant Prof. Shinsuke Sando for his helpful advice, discussions, and encouragement. The author is also indebted to Associate Prof. Takashi Sera for his helpful suggestions.

The author wishes to thank Prof. Yasuo Mori of Kyoto University and Dr. Takashi Yoshida for fluorescence microscopy in Chapter 1. The author acknowledges Prof. Isao Saito of Nihon University for VersaDoc 3000 in Chapter 2. The author is grateful to Prof. Kazuhiko Nakatani and Assistant Prof. Masaki Hagihara of Osaka University for helpful discussion about *in vitro* selection. The author thanks Prof. Itaru Hamachi and Prof. Yasuo Mori of Kyoto University for use of a confocal fluorescence microscopy and developing machines. The author also thanks Assistant Prof. Shigeki Kiyonaka, Mr. Takashi Miki, Mr. Shinji Matsumoto and Mr. Shohei Fujishima of Kyoto University for their advice and technical supports in Chapter 5.

The author thanks Messrs. Toshinori Sasaki (Chapter 1), Kenji Abe (Chapter 2), Kazumasa Ogawa (Chapter 3), Masayoshi Hayami (Chapter 5) for collaboration. The author specially thanks Dr. Atsushi Ogawa and Dr. Keiichiro Kanatani for their supreme advice. The author is also grateful to Messrs. Hiroyuki Tanaka, Kazuki Matsui, Takahumi Iwamoto, Teruyuki Nishi, Toshihiro Shibata, Kousei Uehira, Hiroki Masu, Keigo Mizusawa, Takeshi Tokunaga, Ms. Chika Hurutani, and other members of Prof. Aoyama's research group for their helpful suggestions and hearty encouragement.

The author thanks Japan Society for the Promotion of Science (JSPS) for financial support (Fellowship for Japanese Junior Scientists).

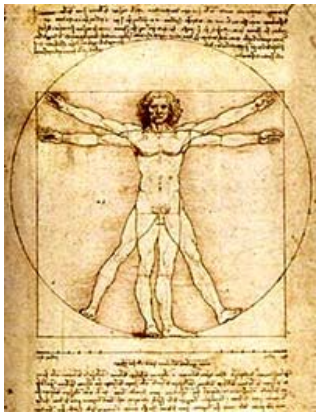
The author expresses his deep appreciations to his parents, Mr. Hiroshi Narita and Mrs. Youko Narita and his grandfather, Mr. Kouichi Sarata for their constant assistance and encouragement.

Atsushi Narita

January, 2008

Contents

General Introduction



"D'ou venous-nous? Que somm es-nous? Ou allons-nous?" (Paul Gauguin, 1897) We have tried to solve this problem through religion and science. Theism holds that humanity was designed for a purpose, while on the other hand, science claims that design and the purposes it serves are an illusion. Though we don't know which way of knowing would give us an unequivocal clue, it would not be an exaggeration to say that substantial ologies such as astrology, philosophy, biology, psychology, anthropology, etc. started from this problem.

Since Watson-Crick discovery 55 years ago,[1] it is now a matter of common knowledge that DNA (deoxyribose nucleic acid) sequences determine an organism's genetic code. There is more and more genetic information available, and the entire human genome has been mapped.[2] Mapping ranged to a genome sequence of even an individual human.[3] Sequence variation in human genes is largely confined to SNPs (single nucleotide polymorphisms), and there are requirements for evaluation of the links between newly discovered gene mutations and human disorders. Furthermore, non-protein-coding genes have still unknown world rather than protein-coding genes.

SNPs Genotyping.

Much effort has focused on methods for detecting various genetic differences in individuals. The most abundant type of DNA sequence variation in the human genome is SNP,[2] which can be characterized as a substitution, insertion, or deletion at a single base position on a DNA strand. There is expected to be on average one SNP for every 500-1000 bases of the human genome, and some variations located in genes are suspected to cause phenotypic diversity, influencing an individual's anthropometric characteristics, and risk of certain diseases including cardiovascular disease, cancer, diabetes, autoimmunity, psychiatric illnesses and many others.[4] So, there has been growing recognition that large collections of mapped SNPs would provide a powerful tool for human genetic studies.[5] SNPs can be used as genetic markers to identify the genes that underlie complex diseases and to realize the full potential of pharmacogenomics in analyzing variable response to drugs.

For major instance, PCR (polymerase chain reaction)-based methods,[6] FRET (fluorescence resonance energy transfer)-based technologies,[7] RCA (rolling circle amplification) assays,[8] Invader™ assays,[9] microarrays[10] and mass spectrometry technologies[11] have been developed for SNP genotyping. These technologies enable global genetic analysis of individual patients.[12] However, there is still requirement for a rapid, simple, low-cost assay for highly parallel genotyping.

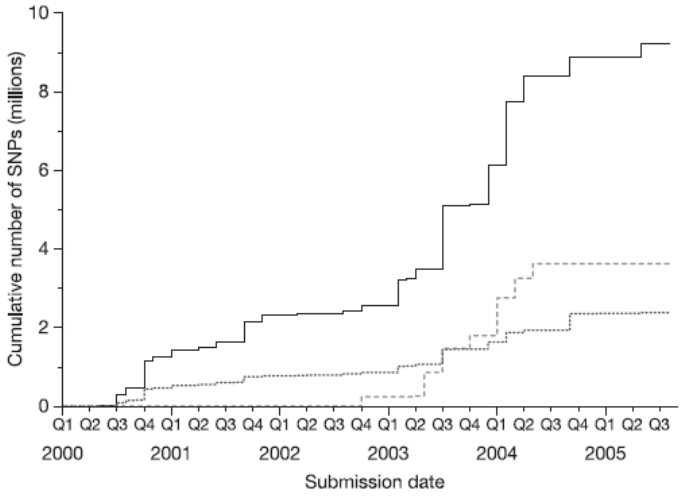


Figure 1. Number of SNPs in dbSNP (SNP database) over time (from Ref. 2). The cumulative number of non-redundant SNPs (each mapped to a single location in the genome) is shown as a solid line, as well as the number of SNPs validated by genotyping (dotted line) and double-hit status (dashed line). Years are divided into quarters (Q1–Q4).

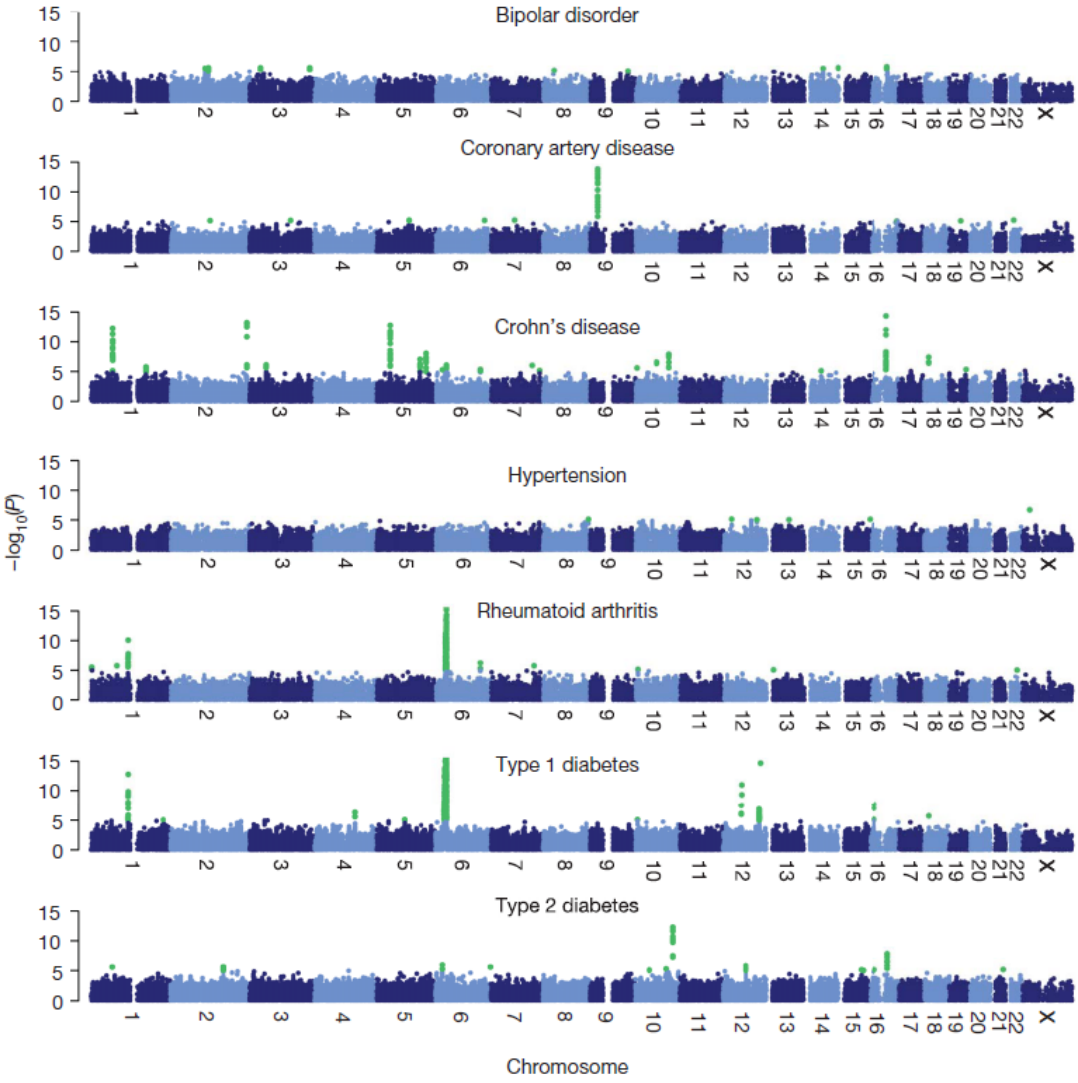


Figure 2. Genome-wide scan for seven diseases (from Ref. 4b). For each of seven diseases $-\log_{10}$ of the trend test *P* value for quality-control-positive SNPs, excluding those in each disease that were excluded for having poor clustering after visual inspection, are plotted against position on each chromosome. Chromosomes are shown in alternating colors for clarity, with *P* values $< 1 \times 10^{-5}$ highlighted in green. All panels are truncated at $-\log_{10}(P \text{ value}) = 15$.

Non-Coding RNA.

The information stored in genome DNA is transmitted, via mRNA (messenger ribonucleic acid), to protein. Traditionally, most RNAs were regarded as carriers in scheme of "Central Dogma". The most prominent exceptions to this are tRNA (transfer RNA) and rRNA (ribosomal RNA), both of which work in the process of translation. However, over the years, other types of non-protein-coding RNAs have been discovered in organisms ranging from bacteria to mammals, which are as important as proteins in the regulation of vital cellar function at many levels.[13] Mammalian cells harbor multitudinous small ncRNAs (non-coding RNAs), which include snoRNAs (small nucleolar RNAs), miRNAs (microRNAs), siRNAs (short interfering RNAs) and small double-stranded RNAs.[14] However, there has been considerably less progress in actually demonstrating the functions of ncRNAs. One reason for the indefinite is that genes have been identified through cDNA (complementary DNA) cloning of polyadenylated mRNAs, EST (expressed sequence tag) sequencing,[15] identification of conserved coding exons by comparative genome analysis, and computational gene prediction. These methods work best for large, highly expressed, evolutionarily conserved protein-coding genes. However, they essentially do not work at all for one class of genes — the ncRNA genes.[16] Therefore, further researches on ncRNAs will give us a new framework for considering and understanding the genomic programming of biological complexity.

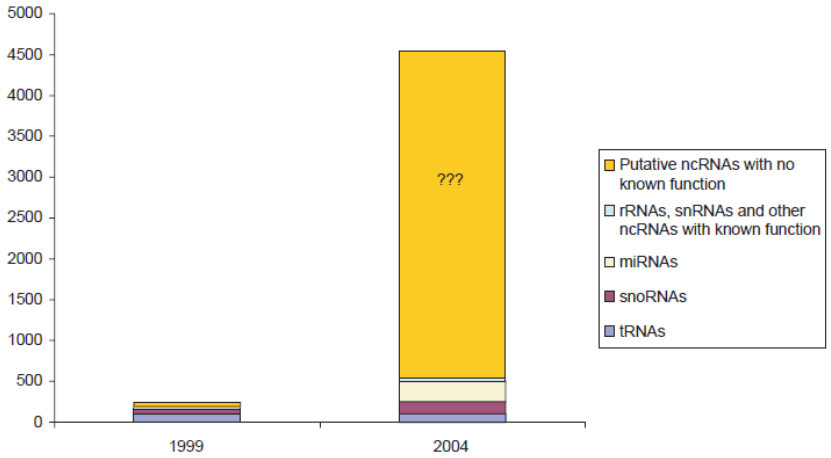


Figure 3. The rapidly increasing number of mammalian ncRNAs and ncRNA candidates from 1999 to 2004 (from Ref. 13c). Estimates of sizes of families of known RNAs based on experimental or computational studies. The numbers for ncRNA candidates are estimated from data from mouse cDNA-transcript libraries. The question marks reflect the current uncertainty regarding the proportion of these transcripts that are actually functional.

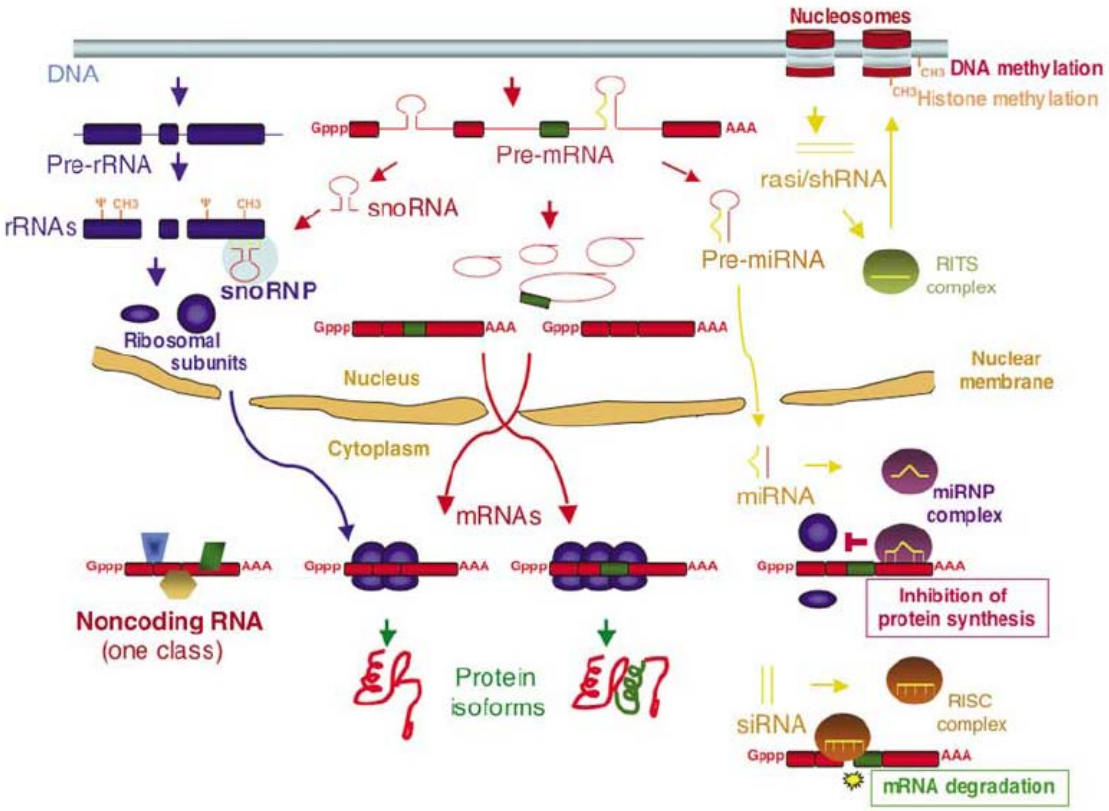


Figure 4. An overview of the eukaryotic transcriptome through examples of its products (from Ref. 13a).

RNA Imaging in Living Cells.

The differences in the genetic sequence, the amount of transcripts, and the timing of transcription among various cells are vital information for our understanding of cell-based events such as cell–cell communication.[17] The development of a method that can be used for nucleic acid sensing without destruction of the cell membrane is highly desirable. PCR-free homogeneous nucleic acid sensing is a conceptually new method which has recently gained much attention for rapid and simple sensing. To obtain accurate information on gene activation, RNA localization, and processing, it is necessary to be able to observe RNAs in living cells.[18] Current methods of fluorescent RNA-imaging rely on the use of labeled molecules that can bind to RNA of interest.

Fluorescent *In-Situ* Hybridization Technology: FISH (fluorescent *in-situ* hybridization) assays, which are useful for localization of specific nucleic acids sequences in native context, are a 25-year-old technology that have developed on an ongoing basis.[19] This approaches using fluorescence-labeled oligonucleotides have been improved to optimize the detection of DNA and RNA on the single cell level.[20] The oligonucleotide probes selectively hybridize to target nucleic acids with Watson-Crick base pairing interactions. However, since FISH technologies essentially require fixing the cells and have an intrinsic susceptibility to degradation by nucleases in cells, it is especially problematic to see dynamic process or behavior of RNAs. Furthermore, since both the bound and free probes equally provide fluorescent, FISH techniques are also plagued by high background fluorescence.

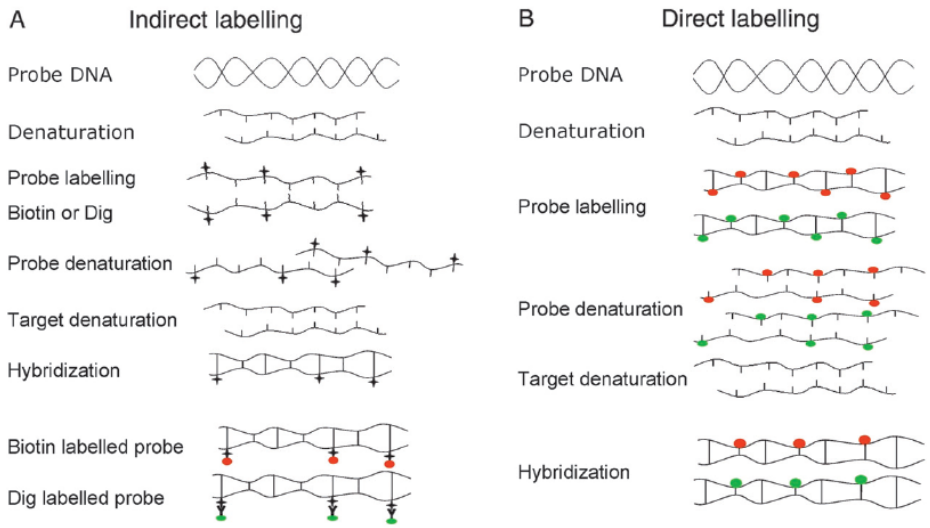


Figure 5. Principles of direct and indirect FISH (from Ref. 19b).

Molecular Beacon Technology: To attempt to detect and monitor specific nucleic acids in real time in homogeneous solutions, molecular beacon (MB) probes have been developed.[21] MBs are hairpin-shaped oligonucleotide probes that contain an essential target-complementary sequence in a loop domain and are linked to a FRET pair at the 5'- and 3'-ends, so that the fluorescence is quenched when not hybridized to the target, thereby reducing background fluorescence. Upon binding to the target, MBs undergo opening of the stem-loop structure (and hence the FRET pair) to restore the fluorescence. MBs have been used for the visualization and localization of RNA in living cells with much success.[22] However, MBs tend to nonspecifically bind to proteins and RNA,[23] and are basically subject to degradation by nucleases.

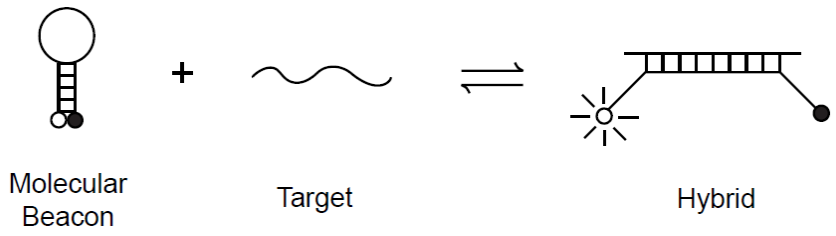


Figure 6. Principles of operation of molecular beacons (from Ref. 21d).

To enhance the specificity and sensitivity, the modifications of conventional MBs are still in progress.[24] Scorpion probe,[25] catalytic MB,[26] PNA (peptide nucleic acid) MBs,[27] quencher-modified MBs,[28] ribozyme-coupled MBs,[29] and binary probe[30] are representative examples of the variants of such modifications and applications. Binding of the target results in opening of the stem-loop structure to restore fluorescence or to switch on the enzymatic or electrochemical catalysis, thus allowing signal amplification. However, There is anxiety for *in vivo* application about biocompatibility and cell-membrane permeability because the MB approach needs to covalently attach fluorophores (and quenchers) or small molecules to the oligonucleotide probes which have polar anionic backbones.

RNA Binding Protein Fusion Technology: To follow specific mRNA in cells, RNA binding protein fusion technology has been reported. GFP-MS2 fusion protein is a representative method to visualize native RNA movement in real time in living cells.[31] This technique utilizes interaction between a 19-nt (nucleotide) RNA hairpin and a phage capsid protein MS2.[32] The MS2-binding sites were inserted into a mRNA of interest, to provide increased signal from multiple bound GFP-MS2 fusion protein. GFP-MS2 protein can be expressed in cells by introduction of a plasmid. Expressed GFP-MS2 protein is restricted to the nucleus if not complexed to the mRNA which has MS2-binding site. Although this GFP-MS2 mRNA-tagging system has been used to study real-time mRNA analysis,[33] it has several disadvantages. Most significantly, the system inherently suffers from a low signal-to-noise ratio because of the presence of unbound fluorescent GFP-MS2.

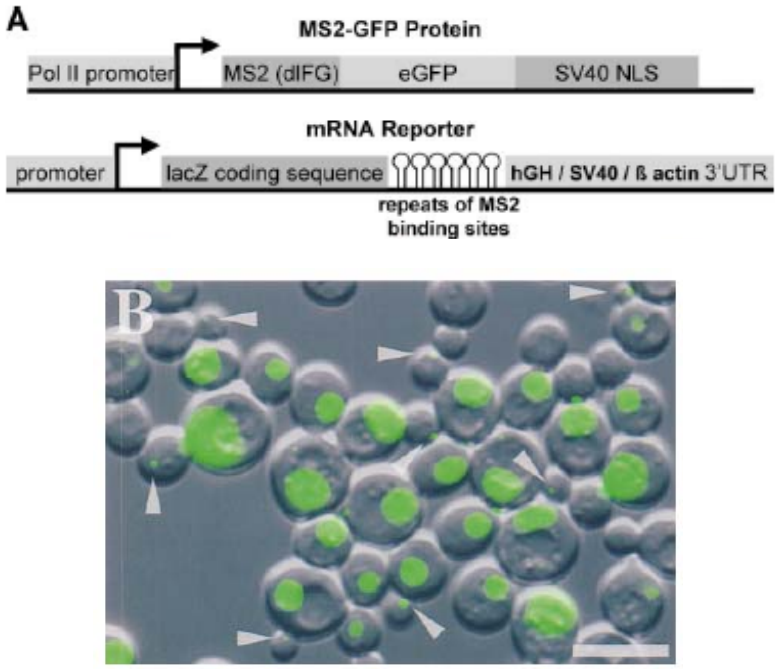


Figure 7. (A) A schematic of GFP-MS2 mRNA-tagging system. The cassettes expressing the MS2-GFP fusion protein and the reporter mRNA are shown (from Ref. 33a). (B) Sample of live cells expressing the GFP-MS2 protein and the reporter mRNA. Arrows indicate some of the particles, usually in the bud. Bar, 5 μm (from Ref. 31).

Problems and Prospects of Current Nucleic Acids Sensing.

Although a number of nucleic acid sensing and imaging strategies, a fraction of which are described above, have resulted in substantial success, these current methods still have some problems to be solved in some cases. Therefore, there is room to be improved for wider use or the purpose of the particular study such as *in vivo* imaging. While other approaches toward genotyping and RNA imaging have been reported,[34-40] the situations of their utilization is restricted. Currently, the methods for *in vivo* sensing and imaging of nucleic acids are noted, hence biocompatibility and cell-membrane permeability are noticeable properties. So, in many instances, the specific features of DNA,[41] RNA[16,42] and protein[32,43] in their natural forms and the functions of the cell-original system[44] have a potential to be turned to good accounts for the sensing in living cells.

Survey of This Thesis.

This thesis is composed of six chapters mainly on the studies toward overcoming these problems. The first three chapters describe the development of catalytic sensing methods: "Locked TASC probe" for catalytic sensing of nucleic acids (Chapter 1), "artificial riboregulation system" for biocompatible sensing of nucleic acids utilizing prokaryote translation system (Chapter 2), "RNase H activity-coupled riboregulation system" for multiply catalytic sensing of nucleic acids (Chapter 3). The following two chapters describe a new strategy to generate light-up fluorophore/nucleic acid sequence pair: "modified Hoechst dye/DNA-aptamer light-up pair" for convenient genotyping by use of unmodified DNA (Chapter 4), and "blue fluorescent RNA tag" for imaging in real time in living cells (Chapter 5). The last chapter describe the development of catalytic nucleic acid-detection methods: "artificial MB-type shRNA" for biocompatible detection strategy utilizing eukaryote gene silencing system (Chapter 6).

Each chapter is summarized as follows.

Chapter 1 is focused on the development of second-generation TASC (target-assisted self-cleavage) probe, named "Locked TASC probe", for suppression of off-target activity. Addition of lock-domain which suppresses *cis*-acting DNAzyme activity can allow more accurate nucleic acid detection by the Locked TASC probe. Furthermore, the Locked TASC probe with a FRET pair can be used to image fixed *E. coli* cells.

Chapter 2 concerns a MB (molecular beacon)-type riboregulation system for multiply catalytic nucleic acid sensing. This sensing, using an unmodified RNA or even stabler dsDNA as a probe with a chemiluminescence output, allowed doubly catalytic or amplifiable with sensitivity at ≤50 fmol in respect to target with 4.5 fmol (1 ng/μL) of probe and a single-nucleotide resolution.

Chapter 3 deals with the development of RNase H activity-coupled riboregulation system. The utilization of RNase H activity in combination with MB-type riboregulator described in Chapter 2 allows more multiple genotyping under isothermal conditions. The improved system allows to color visualize the target nucleic acid sequence. What is needed for this color-based sensing is simply a genetically encodable unmodified RNA or DNA which works in a prokaryotic translation system.

Chapter 4 deals with the development of a light-up fluorophore/DNA-aptamer pair by a down-modification of Hoechst dye and re-selection of DNA-aptamer that binds to the modified Hoechst derivative. The selected DNA-aptamers worked as triggers to enhance the fluorescence of an otherwise nonfluorescent Hoechst derivative. An optimized DNA-aptamer sequence was split into binary probes, thus enabling us to detect a target nucleic acid sequence with a single-nucleotide resolution.

Chapter 5 concerns the development of a bioorthogonal light-up fluorophore/RNA-aptamer pair by a down-modification of Hoechst dye and re-selection of RNA-aptamer that binds to the modified Hoechst derivative. All of RNA aptamers, *in vitro*-selected against the immobilized Hoechst dye, induced a fluorescent enhancement of the Hoechst dye. It was demonstrated that the optimized RNA-tag/Hoechst derivative pair can be used as a dynamic fluorescent label of mRNA transcription by fusing the tag sequence at 3'-side of mRNA.

References

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953,** *171*, 737-738.

(2) The international HapMap consortium *Nature* **2005,** *437*, 1299-1320.

(3) Levy, S.; Sutton, G.; Ng, P. C.; Feuk, L.; Halpern, A. L.; Walenz, B. P.; Axelrod, N.; Huang, J.; Kirkness, E. F.; Denisov, G.; Lin, Y.; MacDonald, J. R.; Pang, A. W. C.; Shago, M.; Stockwell, T. B.; Tsiamouri, A.; Bafna, V.; Bansal, V.; Kravitz, S. A.; Busam, D. A.; Beeson, K. Y.; McIntosh, T. C.; Remington, K. A.; Abril, J. F.; Gill, J.; Borman, J.; Rogers, Y. H.; Frazier, M. E.; Scherer, S. W.; Strausberg, R. L.; Venter, J. C. *PLoS Biol.* **2007,** *5*, e254.

(4) (a) Goldstein, D. B.; Cavalleri, G. L. *Nature* **2005,** *437*, 1241-1242. (b) The Wellcome Trust Case Control Consortium **2007,** *447*, 661-678.

(5) Cargill, M.; Altshuler, D.; Ireland, J.; Sklar, P.; Ardlie, K.; Patil, N.; Lane, C. R.; Lim, E. P.; Kalyanaraman, N.; Nemesh, J.; Ziaugra, L.; Friedland, L.; Rolfe, A.; Warrington, J.; Lipshutz, R.; Daley, J. Q.; Lander, E. S. *Nat. Genet.* **1999,** *22*, 231-238.

(6) Gibson, N. J. *Clin Chim Acta* **2006,** *363*, 32-47.

(7) Ihara, T.; Mukae, M. *Anal Sci.* **2007,** *23*, 625-629.

(8) Nilsson, M.; Dahl, F.; Larsson, C.; Gullberg, M.; Stenberg, J. *Trends Biotechnol.* **2006,** *24*, 83-88.

(9) Olivier, M. *Mutat. Res.* **2005,** *573*, 103-110.

(10) Fan, J. B.; Chen, X.; Halushka, M. K.; Berno, A.; Huang, X.; Ryder, T.; Lipshutz, R. J.; Lockhart, D. J.; Chakravarti, A. *Genome Res.* **2000,** *10*, 853-860.

(11) Ragoussis, J.; Elvidge, G. P.; Kaur, K.; Colella, S. *PLoS Genet.* **2006,** *2*, e100.

(12) Fan, J. B.; Chee, M. S.; Gunderson, K. L. *Nat. Rev. Genet.* **2006,** *7*, 632-644.

(13) (a) Soares, L. M. M.; Valcarcel, J. *EMBO J.* **2006,** *25*, 923-931. (b) Mattick, J. S.; Makunin, I. V. *Hum. Mol. Genet.* **2006,** *15*, R17-R29. (c) Huttenhofer, A.; Schattner, P.; Polacek, N. *Trends Genet.* **2005,** *21*, 289-297.

(14) Mattick, J. S.; Makunin, I. V. *Hum. Mol. Genet.* **2005,** *14*, R121-R132.

(15) (a) Ewing, B.; Green, P. *Nature Genet.* **2000,** *25*, 232-234. (b) Liang, F.; Holt, I.; Pertea, G.; Karamycheva, S.; Salzberg, S. L.; Quackenbush, J. *Nat. Genet.* **2000,** *25*, 239-240. (c) Wiemann, S.; Weil, B.; Wellenreuther, R.; Gassenhuber, J.; Glass, S.; Ansorge, W.; Bocher, M.; Blocker, H.; Bauersachs, S.; Blum, H.; Lauber, J.; Dusterhoft, A.; Beyer, A.; Kohrer, K.; Strack, N.; Mewes, H. W.; Ottenwalder, B.; Obermaier, B.; Tampe, J.; Heubner, D.; Wambutt, R.; Korn, B.; Klein, M.; Poustka, A. *Genome Res.* **2001,** *11*, 422-435.

(16) Eddy, S. R. *Nat. Rev. Genet.* **2001,** *2*, 919-929.
(17) Mihalcescu, I.; Hsing, W.; Leibler, S. *Nature* **2004,** *430*, 81-85.
(18) Tal, Y. S.; Singer, R. H.; Darzacq, X. *Nat. Rev. Mol. Cell Biol.* **2004,** *5*, 856-862.
(19) (a) Ekong, R.; Wolfe, J. *Curr. Opin. Biotechnol.* **1998,** *9*, 19-24. (b) Lambros, M. B. K.; Natrajan, R.; Filho, J. S. R. *Hum. Pathol.* **2007,** *38*, 1105-1122. (c) Halling, K. C.; Kipp, B. R. *Hum. Pathol.* **2007,** *38*, 1137-1144.
(20) Levsky, J. M.; Singer, R. H. *J. Cell Sci.* **2003,** *116*, 2833-2838.
(21) (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996**, *14*, 303-308. (b) Tyagi, S.; Bratu, D. P.; Kramer, F. R. *Nat. Biotechnol.* **1998**, *16*, 49-53. (c) Bonnet, G.; Tyagi, S.; Libchaber, A.; Kramer, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6171-6176. (d) Marras, S. A. E.; Kramer, F. R.; Tyagi, S. *Genotyping single nucleotide polymorphisms with molecular beacons*, The Humana Press Inc., Towada, NJ **2003,** *212*, 111–128.
(22) (a) Sokol, D. L.; Zhang, X.; Lu, P.; Gewirtz, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11538-11543. (b) Vargas, D. Y.; Raj, A.; Marras, S. A. E.; Kramer, F. R.; Tyagi, S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17008-17013. (c) Santangelo, P.; Nitin, N.; Bao, Gang. *Ann. Biomed. Eng.* **2006,** *34*, 39-50.
(23) (a) Fang, X.; Li, J. J.; Tan, W. *Anal. Chem.* **2000**, *72*, 3280-3285. (b) Molenaar, C.; Marras, S. A.; Slats, J. C. M.; Truffert, J. C.; Lemaitre, M.; Raap, A. K.; Dirks, R. W.; Tanke, H. J. *Nucleic Acids Res.* **2001**, *29*, e89.
(24) (a) Tan, W.; Wang, K.; Drake, T. J. *Curr. Opin. Chem. Biol.* **2004**, *8*, 547-553. (b) Goel, G.; Kumar, A.; Puniya, A. K.; Chen, W.; Singh, K. *J. Appl. Microbiol.* **2005**, *99*, 435-442. (c) Santangelo, P.; Nitin, N.; Bao, Gang. *Ann. Biomed. Eng.* **2006,** *34*, 39-50.
(25) Whitcombe, D.; Theaker, J.; Guy, S. P.; Brown, T.; Little, S. *Nat. Biotechnol.* **1999**, *17*, 804-807.
(26) Stojanovic, M. N.; Prada, P.; Landry, D. W. *ChemBioChem* **2001**, *2*, 411-415.
(27) (a) Kuhn, H.; Demidov, V. V.; Coull, J. M.; Fiandaca, M. J.; Gildea, B. D.; Kamenetskii, M. D. F. *J. Am. Chem. Soc.* **2002**, *124*, 1097-1103. (b) Xi, C.; Balberg, M.; Boppart, S. A.; Raskin, L. *Appl. Environ. Microbiol.* **2003**, *69*, 5673-5678.
(28) (a) Du, H,; Disney, M. D.; Miller, B. L.; Krauss, T. D. *J. Am. Chem. Soc.* **2003**, *125*, 4012-4013. (b) Brunner, J.; Kraemer, R. *J. Am. Chem. Soc.* **2004**, *126*, 13626-13627.
(29) Hartig, J. S.; Grune, I.; Shoushtari, S. H. N.; Famulok, M. *J. Am. Chem. Soc.* **2004**, *126*, 722-723.

(30) Kolpashchikov, D. M. *J. Am. Chem. Soc.* **2006**, *128*, 10625-10628.
(31) Bertrand, E.; Chartrand, P.; Schaefer, M.; Shenoy, S. M.; Singer, R. H.; Long, R. M. *Mol. Cell* **1998**, *2*, 437-445.
(32) Fouts, D. E.; True, H. L.; Celander, D. W. *Nucleic Acids Res.* **1997**, *25*, 4464-4473.
(33) (a) Fusco, D.; Accornero, N.; Lavoie, B.; Shenoy, S. M.; Blanchard, J. M.; Singer, R. H.; Bertrand, E. *Curr. Biol.* **2003**, *13*, 161-167. (b) Shav-Tal, Y.; Darzacq, X.; Shenoy, S. M.; Fusco, D.; Janicki, S. M.; Spector, D. L.; Singer, R. H. *Science* **2004**, *304*, 1797-1800. (c) Janicki, S. M.; Tsukamoto, T.; Salghetti, S. E.; Tansey, W. P.; Sachidanandam, R.; Prasanth, K. V.; Ried, T.; Shav-Tal, Y.; Bertrand, E.; Singer, R. H.; Spector, D. L. *Cell* **2004**, *116*, 683-698. (d) Golding, I.; Paulsson, J.; Zawilski, S. M.; Cox, E. C. *Cell* **2005**, *123*, 1025-1036.
(34) Nakatani, K. *ChemBioChem* **2004**, *5*, 1623-1633.
(35) (a) Xu, Y.; Karalkar, N. B.; Kool, E. T. *Nat. Biotechnol.* **2001**, *19*, 148-152. (b) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002**, *124*, 2096-2097. (c) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002**, *124*, 9686-9687. (d) Sando, S.; Abe, H.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 1081-1087. (e) Silverman, A. P.; Kool, E. T. *Trends Biotechnol.* **2005,** *23*, 225-230. (f) Silverman, A. P.; Kool, E. T. *Nucleic Acids Res.* **2005**, *33*, 4978-4986. (g) Abe, H.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 263-268.
(36) Saghatelian, A.; Guckian, K. M.; Thayer, D. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2003**, *125*, 344-345.
(37) (a) Xiao, Y.; Pavlov, V.; Niazov, T.; Dishon, A.; Kotler, M.; Willner, I. *J. Am. Chem. Soc.* **2004**, *126*, 7430-7431. (b) Pavlov, V.; Shlyahovsky, B.; Willner, I. *J. Am. Chem. Soc.* **2005**, *127*, 6522-6523.
(38) Penchovsky, R.; Breaker, R. R. *Nat. Biotechnol.* **2005**, *23*, 1424-1433.
(39) Hasegawa, S.; Gowrishankar, G.; Rao, J. *ChemBioChem* **2006**, *7*, 925-928.
(40) Beyer, S.; Simmel, F. C. *Nucleic Acids Res.* **2006**, *34*, 1581-1587.
(41) For examples of the specific DNA-feature: about a stem-loop DNA, see (a) Broude, N. E. *Trends Biotechnol.* **2002,** *20*, 249-256. About a DNAzyme, see (b) Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4262-4266. About triplex-forming oligonucleotides (TFOs), see (c) Duval-Valentin, G.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 504-508.
(42) For examples of the specific RNA-feature: about RNA folding and assembly, see (a) Schroeder, R.; Barta, A.; Semrad, K. *Nat. Rev. Mol. Cell Biol.* **2004,** *5*, 908-919. About a ribozyme, see (b) Tang, J.; Breaker, R. R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5784-5789. About an aptamer, see (c) Babendure, J. R.; Adams, S. R.;

Tsien, R. Y. *J. Am. Chem. Soc.* **2003**, *125*, 14716-14717. About antisense oligonucleotides, see (d) Crooke, S. T. *Annu. Rev. Med.* **2004,** *55*, 61-95. About the hairpin loop-loop interaction, see (e) Horiya, S.; Li, X.; Kawai, G.; Saito, R.; Katoh, A.; Kobayashi, K.; Harada, K. *Chem. Biol.* **2003**, *10*, 645-654.

(43) For examples of the specific protein-feature: for review of transcription factors, see (a) Pabo, C. O. *Annu. Rev. Biochem.* **1992,** *61*, 1053-1095. For review of assessing protein, see (b) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2004**, *304*, 1797-1800. About zinc finger proteins, see About the Tat protein, see (c) Yamamoto, R.; Katahira, M.; Nishikawa, S.; Baba, T.; Taira, K.; Kumar, P. K. R. *Genes to Cells* **2000**, *5*, 371-388.

(44) For examples of the functions of the cell-original system: for review of the mechanism of translation initiation in eukaryotes, see (a) Algire, M. A.; Lorsch, J. R. *Curr. Opin. Chem. Biol.* **2006**, *10*, 480-486. For review of internal ribosome entry sites (IRES), see (b) Komar, A. A.; Hatzoglou, M. *J. Biol. Chem.* **2005**, *280*, 23425-23428. For review of RNA swithes, see (c) Gottesman, S. *Genes Dev.* **2002,** *16*, 2829-2842. About ribosome binding site (RBS), see (d) Ringquist, S.; Shinedling, S.; Barrick, D.; Green, L.; Binkley, J.; Stormo, G. D.; Gold, L. *Mol. Microbiol.* **1992,** *6*, 1219-1229. About *trans*-splicing, see (e) Hasegawa, S.; Choi, J. W.; Rao, J. *J. Am. Chem. Soc.* **2004**, *126*, 7158-7159. About a short interfering RNA (siRNA) function, see (f) Chiu, Y. L.; Rana, T. M. *RNA* **2003**, *9*, 1034-1048. About an antigene RNA (agRNA), see (g) Janowski, B. A.; Huffman, K. E.; Schwartz, J. C.; Ram, R.; Hardy, D.; Shames, D. S.; Minna, J. D.; Corey, D. R. *Nat. Chem. Biol.* **2005,** *1*, 216-222.

Chapter 1

Locked TASC Probes for Homogeneous Sensing of Nucleic Acids and Imaging of Fixed *E. coli* Cells

Abstract

We designed a second-generation TASC probe. It is based on switch-on of the incorporated *cis*-acting DNAzyme activity upon target-induced conformational change of the otherwise inactive off-target probe locked in an intrastrand base-paired hairpin geometry. With *E. coli* 16S ribosomal RNA-relevant oligonucleotides as targets, the locked TASC probe exhibits an allosteric factor of k_{on}/k_{off} = 65 and the sequence selectivity in terms of single nucleotide difference is high when the sequence and length of targets are particularly chosen. Preliminary experiments with fixed E. coli cells show that the locked TASC probe with a FRET pair can be used to image fixed *E. coli* cells.

Introduction

As an increasing volume of sequence information continues to reveal important genetic markers, there have been increasing demands on rapid and accurate nucleic acid sensing especially at a single nucleotide resolution.[1] Although a number of nucleic acid sensing strategies have been reported during the last decade,[2] most of them are applicable only to PCR-isolated/purified specimen as targets. These isolation/purification processes normally require cell membrane destruction, resulting in loss of individual information of various cells to allow only a mean genetic information of cell assemblies. Difference in genetic sequences, amount of transcripts, and timing of transcription among various cells are indispensable informations for our understanding of cell-based events such as cell-cell communications.[3] It is now highly desirable to develop a method that can be used for in-situ, especially in-cell nucleic acid sensing without destruction of the cell membrane.

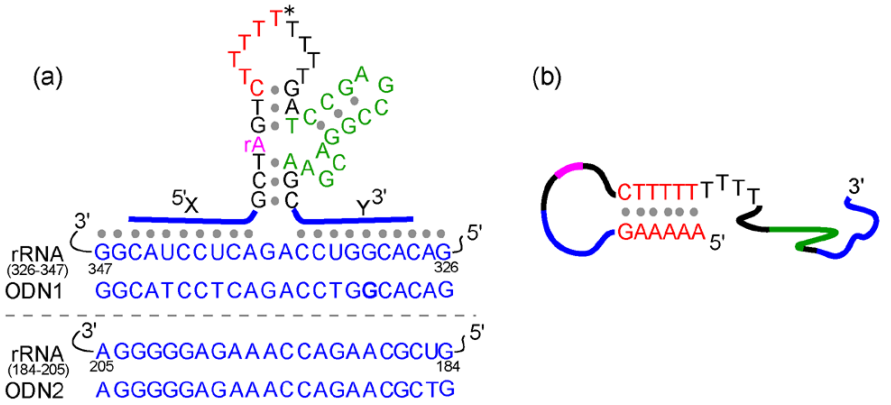
The "non-enzymatic and reagent-free nucleic acids detection" is a conceptually new method of much recent attention for rapid, simple, and *in-situ* (in-cell) sensing.[4-7] The FISH (fluorescence in situ hybridization) approach[5] using fluorescence-labelled hybridization probe has been broadly used on a single cell level. A problem here is post-hybridization washing; off-target (unbound) fluorescence probe must be carefully and thoroughly washed out,[8] thus preventing the method from being used in homogeneous *in-situ* sensing. Recently, Kool and co-workers have developed a new DNA/RNA-sensing method, known as QUAL (quenched autoligating) based on proximity-dependent chemical ligation, which proceeds only when two probes bind on target DNA/RNA side-by-side with concomitant "light up" upon nucleophilic displacement.[6a] They succeeded in detecting DNA on solid surface[6a] and in discriminating ribosomal RNA (rRNA) sequences in fixed or non-fixed *Escherichia coli* (*E. coli*) cells with a single nucleotide resolution without post hybridization washings.[6b,6c] Widely used molecular beacons (MBs) are also successful examples of in-cell nucleic acid imaging.[7] MBs are hairpin-shaped oligonucleotide probes with an essential loop domain complementary to target DNA/RNA and a FRET pair at the 5' and 3' ends. Upon binding to target, MBs undergo opening of the hairpin structure and hence FRET pair to restore fluorescence. In a recent study, RNase-resistant MBs were used to visualize distribution and transport of *oskar* mRNA in living *Drosophia melanogaster* oocytes.[7b] With such good aspects in mind, however, these methods have been confronted with difficulties in terms of *sensitivity*, because one can expect only one signal from one target at best in a stoichiometric manner. A next challenge

is amplification of sensing signals under physiological (isothermal) conditions.

Catalytic molecular beacons (catMBs)[9] and signal-amplifying ribozymes,[10] wherein DNA/RNAzyme activity is initially suppressed by the MB sequence embedded in the same strand, are one of promising candidates for amplified sensing of nucleic acids.[11] Binding to target results in opening of the MB structure, activates the *trans*-acting DNA/RNAzyme, triggers catalytic cleavage of target FRET-probes,[12,13] and ultimately leads to amplification of unquenched fluorescence signals over a stoichiometric level in homogeneous solution.[9,10] However, such a *trans*-acting system requiring synchronized association of DNAzyme and FRET probe on the target DNA/RNA would be less suitable for sensing in crowded in-cell environments. We reported the TASC (Target-Assisted Self-Cleavage) approach as a new concept for *catalytic* mix-and-read sensing.[14] The *cis*-acting TASC probes incorporating a DNAzyme function are intended to undergo Mg^{2+}-assisted, on-target strand cleavage with concomitant dissociation of the probe fragments from the target, thus allowing multiple-turnover self-cleavage under conditions of [probe] > [target], where the essential role of target is that of an allosteric effector. Catalytic self-cleavage should result in amplified fluorescence emission under isothermal and enzyme/reagent-free conditions. This is in principle indeed the case.[14] For in-cell use, there are a couple of concerns. First, non-sensing or off-target self-cleavage of the probe *in excess* should be highly suppressed. The first-generation TASC probe is not satisfactory in this respect; it is flexible and still possesses a substantial off-target activity. Second, on the other hand, the conformation and hence the activity of the off-target probe might be affected by intracellular components in an unpredictable way. We are concerned here about conformational fixation of the probe by itself.[9,10] The present work aims at shedding light on these two concerns. We report here that the second-generation TASC probe having a hairpin-locked geometry exhibits a better allosteric (target-on/off) performance and can be used for rRNA imaging in *E.coli*.

Results and discussion

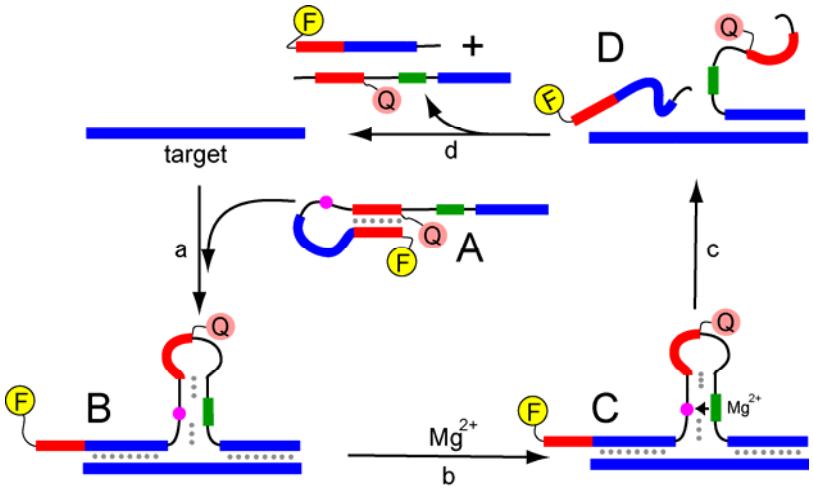
Design of locked TASC probes. Probes **1-5** used here (Figure 1) are fluorescence-labelled at the 5'-end. TASC probe **1**, designed along the concept previously reported[14] with a basic design principle of Wang et al.,[13b-d] is composed of a self-cleaving sequence containing a DNAzyme domain (green) and a ribonucleotide rA as a specific cleavage site (pink) and a target-binding sequence (blue), split into two parts ($^{5'}$X and $^{3'}$Y) (10 + 10 = 20 nucleotides), complementary to the accessible 326-347 region of 16S rRNA[15] of the *E. coli* K12 strain MG1655 or the corresponding 22-mer oligodeoxyribonucleotide (ODN1) as a target in blue (Figure 1a).[16] Self-cleaving reactions were conducted in a Tris-HCl buffer (50 mM, pH 7.2) containing Mg^{2+} (25 mM) with [probe] = 1 μM and [target] = 5 (5-fold molar excess) or 0 μM at 37 °C and the reaction kinetics were analyzed on PAGE. Incubation of TASC probe **1** with target 22-mer ODN1 afforded two fragmented ODNs (one is fluorescent 14-mer and the other is nonfluorescent 41-mer). The PAGE-monitored self-cleavage of this probe in the absence and presence of target ODN1 with Mg^{2+} followed the first-order kinetics with $k_{obs}^{-} = 1.8 \times 10^{-3}$ min^{-1} and $k_{obs}^{+} = 5.0 \times 10^{-3}$ min^{-1}, respectively (Table 1). This poor allosteric activation factor of $k_{obs}^{+}/k_{obs}^{-} \cong 2.8$ is not surprising in view of the flexible nature of the probe[16] and may be compared with the relatively large allosteric activation of the previously reported c-*fos* targeting TASC probe.[14] This conflict may be explained in terms of intramolecular or intrastrand folding of the off-target probes. Prediction of secondary structure based on the RNAfold program shows that the present TASC probe **1** forms no well-definable secondary, i.e., internary hybridized, structure,[16] while the previous c-*fos* targeting probe forms a stable intrastrand hairpin-lock structure with total ~15 bps (bp = base pair). Thus, the low off-target activity of the previous probe is due to unexpected folding and hence somehow accidental. The intrastrand folding or base-pairing depends on the base sequence of variable parts of the probe, particularly the "stem" part, which are constructed on an arbitrary basis[16] and the target-binding domain which is complementary to the particular target in concern. Removal of such an ambiguous and target-dependent nature of the off-target activity of the first-generation TASC probe is what is aimed at in the design of the second-generation TASC probes for more general use.



	5'X	Y3'	target rRNA site
TASC **1**:	CCGTAGGAGT	GGACCGTGTC	326-347
locked TASC **2**:	AAAAAGCCGTAGGAGT	GGACCGTGTC	326-347
locked TASC **3**:	AAAAAGGTAGGAGT	GGACCGTG	328-345
TASC **4**:	TCCCCCTCTT	GGACCGTGAC	184-205
locked TASC **5**:	AAAAAGTCCCCCTCTT	GGACCGTGAC	184-205

Figure 1. Sequences of TASC probes **1**-**5** and their targets. Illustration of (a) the target-probe complexation with gray dots representing base-pairing and (b) the internal hairpin-locked structure of locked probes **2**, **3**, and **5** in the absence of target.

We thus designed locked TASC probes **2** and **3** with a (10 + 10) or (8 + 8) target-binding sequence, respectively, wherein a 5'-AAAAAG-3' sequence (red) was added to the 5'-end. They are designed to form an internal hairpin-locked structure (Figure 1b or **A** in Scheme 1) via intramolecular 6-bp hybridization with the 5'-CTTTTT-3' sequence, shown in red in Figure 1a, to switch-off the DNAzyme activity as in catMBs[9] and signal amplifying ribozymes,[10] keeping the DNAzyme domain[17] (green) and the rA cleavage site (pink) away from each other to prevent the probe from undergoing self-cleavage. In the presence of the target, they become unlocked upon 20-bp (**2**) or 16-bp (**3**) target-probe hybridization (step a) to switch-on the activity by bringing the DNAzyme and the rA moieties in an enforced proximity (**B**) and hence promote cleavage (step c) via activation (**C**) with Mg^{2+} (step b). The fragmented products in **D** would either leave the target (step d) or be expelled therefrom by full-length probe in excess (vide infra),[9] regenerating free target or target-probe complex **B**, respectively, to drive a catalytic cycle when [probe] > [target] (Scheme 1).



Scheme 1. Sensing scheme for locked TASC probe **2**, **3**, and **5**, as specifically illustrated for FRET probe **3** having a FRET pair of fluorescein fluorophore (F) and dabsyl quencher (Q).

Melting behaviour of simplified model probes and their complexes. The validity of Scheme 1 was checked up from a viewpoint of thermal stability in reference to the melting behaviours of model probes **a**-**d** in the presence and absence of target ODN **e** or **f** (Figure 2a). The simplified model probes contain the essential partial structure of probe **2** with rA, if included, changed into non-cleavable deoxy-A (pink); such sequences as the DNAzyme domain are removed to avoid contribution of base-pairing arising therefrom. Target **e** is the same as ODN1, having a (10+10) probe-binding site. Target **f**, on the other hand, contains a shorter (8+8) binding site. Melting curves were recorded by plotting the absorbance at 260 nm vs temperature for a probe or probe-target mixture in the same medium as for sensing experiments (vide infra), i.e., in a Tris-HCl (50 mM, pH 7.2) containing Mg^{2+} (25 mM). Probe **a** (2 μM strand concentration), having full lock domains (red), showed a melting curve **i** (Figure 2c) with a midpoint of T_m = 45 °C for the hairpin-lock structure (**A** in Figure 2b). This may also be taken as a measure of the thermal stability of such a structure (Figure 1b and **A** in Scheme 1) of locked probes **2** and **3**. An equimolar (2 μM) mixture of probe **b** having a full set of target-binding sequences (blue) and target **e** (10+10) or **f** (8+8) exhibited a melting curve **ii** or **iii** with T_m = 62 °C or 53 °C for the (10+10) or (8+8) target-probe complex $\mathbf{B}^{(10+10)}$ or $\mathbf{B}^{(8+8)}$ (Figure 2b), respectively. These melting points must reflect the stabilities of the (10+10)- and (8+8)-binding **2**-ODN1 and **3**-ODN1 complexes (**B** in Scheme 1), respectively. Thus, from a thermal-stability consideration, probes **2** and **3** mostly exist in a 6-bp hairpin-locked structure with T_m = 45 °C (Figure 1b and **A** in Scheme 1) at the experimental temperature (37 °C) but become unlocked upon hybridization with the target to give complex **B** (Scheme 1) with T_m = 62 or 53 °C.

The melting curve **iv** or **v**, respectively, for an equimolar (2 μM) ternary mixture of probes **c** and **d** having a fragmented target-binding sequence (blue) and target **e** or **f** is monophasic with a midpoint of T_m = 50 °C or 43 °C for the ternary complex $\mathbf{C}^{(10+10)}$ or $\mathbf{C}^{(8+8)}$ (Figure 2c). This may also apply to the ternary complex formed upon on-target self-cleavage, referring to **D** in Scheme 1. While the observed melting temperatures are higher than 37 °C at which sensing experiments were carried out, it is easily expected, judging from the less sharp melting behaviour, that partial melting or dissociation of the probe fragments (step d in Scheme 1) takes place at that temperature.

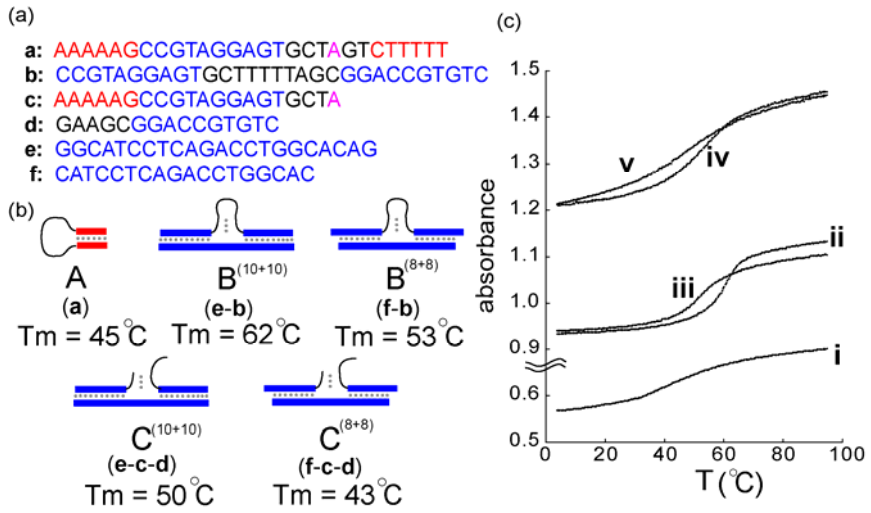


Figure 2. Melting behaviour of simplified probe and target-probe systems. (a) Sequences of simplified model probes **a**-**d** and targets **e** and **f**. The target binding domain, lock domain, and the A site (otherwise cleavable rA moiety in genuine TASC probes) are shown in blue, red, and pink, respectively. (b) Schematic hybridization structures with melting temperatures of probe **a**, binary target-probe complexes **e**-**b** and **f**-**b**, and ternary complexes of target and probe fragments **e**-**c**-**d** and **f**-**c**-**d** in reference to the corresponding structures **A**, **B**, and **C** in Scheme 1. (c) Melting curves for probe **a** (**i**), an equimolar mixture of probe **b** and target **e** (**ii**) or **f** (**iii**), and an equimolar ternary mixture of probes **c** and **d** and target **e** (**iv**) or **f** (**v**). A pair of (10+10) and (8+8) curves, i.e., **ii** and **iii** and **iv** and **v**, are placed so as to have the same absorbance at low temperature just for the sake of comparison of melting temperatures.

Self-cleavage reaction kinetics of locked TASC probes. The self-cleavage reactions of 5'-fluorescein labelled TASC probes were analyzed by fluorescence imaging of the 5'-fragment as fluorescent cleavage product as well as the uncleaved (unreacted) starting probe. The reactions follow the first-order kinetics and Mg^{2+} was essential. An example of the kinetic analysis is shown in Figure 3 for probes 4 and 5 (vide infra). The first-order rate constants for locked probe 2 in the presence (5-fold molar excess) and absence of target ODN1 (Figure 1) are summarized in Table 1. The off-rate (k_{obs}^-) is 5-times lower than that of the non-locked counterpart (probe 1), while the on-rates are not surprisingly similar to each other; the allosteric activation factor (k_{obs}^+/k_{obs}^-) thus increases from 2.8 (probe 1) to 12.3 (probe 2). In order to shed light on the target-dependence, we prepared another set of probes, non-locked 4 and locked 5, which target at an oligonucleotide (ODN2, Figure 1) copying another (10+10) part (184-205 site) of the same 16S rRNA. Probes 5 and 2 (and 4 and 1 as well) differ only in the target-binding sequences. The kinetic data are shown in Figure 3. The change in targets from ODN1 to ODN2 turns out to be slightly inhibitory on non-locked probes with respect to both k^+ and k^- ($k(4)^{ODN2+}/k(1)^{ODN1+}$ = 0.58 and $k(4)^{ODN2-}/k(1)^{ODN1-}$ = 0.67) without any big change in the allosteric factor $k(4)^{ODN2+}/k(4)^{ODN2-}$ = 2.4 as compared with $k(1)^{ODN1+}/k(1)^{ODN1-}$ = 2.8 (Table 1). For locked probes 2 and 5, on the other hand, the effects are again slight but accelerating on k^+ and decelerating on k^- ($k(5)^{ODN2+}$/k(2)$^{ODN1+}$ = 1.7 and $k(5)^{ODN2-}/k(2)^{ODN1-}$ = 0.31); as a consequence, the allosteric factor is significantly enhanced to $k(5)^{ODN2+}/k(5)^{ODN2-}$ = 65. Thus, on one hand, there is little doubt that the present lock strategy (Scheme 1) works in suppressing the off-target reactivity of the probe, while the net sensing accuracy in terms of k^+/k^- ($\geq$ 10) is still delicately dependent on the target sequence. On the other hand, the origin of the non-negligible off-target reactivity of the probe in the neighbourhood of 10^{-4} min^{-1} should be further pursued in terms of lock-unlock equilibrium in reference to the melting temperature of T_m = 45 °C (Figure 2) and intermolecular reactions.

Table 1. Self-cleavage rate constants of 5'-fluorescein labelled TASC probes in the presence (k_{obs}^{+}) and absence (k_{obs}^{-}) of target together with derived allosteric factors (k_{obs}^{+}/k_{obs}^{-})[a]

probe	target	k_{obs}^{-} (min^{-1})[b]	k_{obs}^{+} (min^{-1})[b]	k_{obs}^{+}/k_{obs}^{-}
1	ODN1	1.8×10^{-3}	5.0×10^{-3}	2.8
2	ODN1	3.5×10^{-4}	4.3×10^{-3}	12.3
4	ODN2	1.2×10^{-4}	2.9×10^{-3}	2.4
5	ODN2	1.1×10^{-4}	7.1×10^{-3}	65

[a] [probe] = 1 μM and [target] = 5 μM at 37 °C. [b] First-order rate constants were obtained as the slopes of ln P_f vs t plots where P_f is fraction of uncleaved probe and t = 0-120 min.

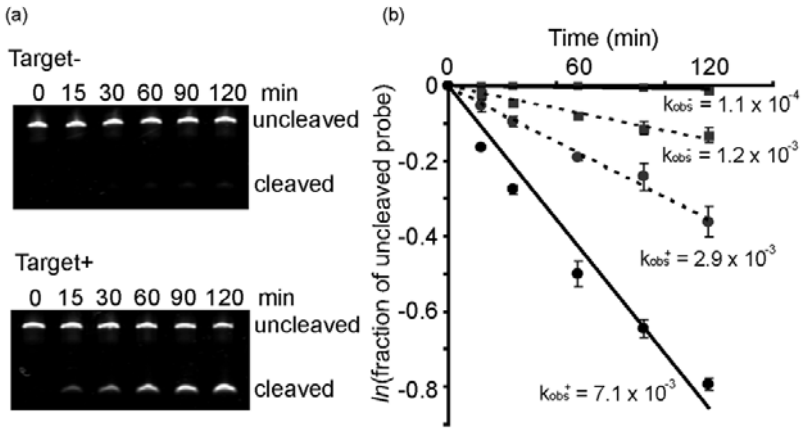


Figure 3. (a) Self-cleavage reactions of locked TASC probe **5** (1 μM) as monitored by gel electrophoresis with fluorescence detection in the absence (Target^{-}) and presence (5 μM) (Target^{+}) of target ODN2. (b) First-order plots and derived rate constants for the self-cleavage reactions of probes **4** (dotted line) and **5** (solid line) (1 μM) in the presence (5 μM) (circle) and absence (square) of target ODN2. The error bar represents a standard deviation of data set obtained from at least three runs.

Single nucleotide discrimination with locked TASC probe. Sequence selectivity of locked probe **2** in terms of single nucleotide discrimination was explored by using a one-base mismatched (T for G shown in bold in Figure 1a) ODN1 as target. Locked TASC probe **2** having a (10+10) target-binding site with two intervening nucleotides turned out to be incapable of discrimination between fully matched (ODN1) and one-base mismatched targets ($k_{obs}^{(full)+} = 4.3 \times 10^{-3}$ min^{-1}, $k_{obs}^{(1\text{-}mis)+} = 3.4 \times 10^{-3}$ min^{-1}, and $k_{obs}^{(full)+}/ k_{obs}^{(1\text{-}mis)+} \cong 1.3$) (Table 2). A one-base selectivity was achieved for target having a shorter (8+8) binding site copying the 328-345 site of 16S rRNA. In the presence of this full-match (8 + 8) target, probe **2** undergoes facilitated cleavage with $k_{obs}^{(full)+}$ in the expected 10^{-3} order (2.5×10^{-3} min^{-1}), while with the mismatch target with T for **G**, the cleavage rate constant ($k_{obs}^{(1\text{-}mis)+} = 2.9 \times 10^{-4}$ min^{-1}) is essentially the same as the off-target rate constant ($k_{obs}^{-} = 3.5 \times 10^{-4}$ min^{-1}); the apparent one-base selectivity of $k_{obs}^{(full)+}/k_{obs}^{(1\text{-}mis)+} \cong 8.6$ may thus be taken as evidence for a high level of single nucleotide discrimination in the present system.

Table 2. Self-cleavage rate constants of locked TASC probe **2** in the presence of fully matched ($k_{obs}^{(full)+}$) and one-base mismatched ($k_{obs}^{(1\text{-}mis)+}$) target oligonucleotides having either a (10 + 10) or (8 + 8) binding site.[a]

target ODN	$k_{obs}^{(full)+}$ (min^{-1})[b]	$k_{obs}^{(1\text{-}mis)+}$ (min^{-1})[b]	$k_{obs}^{(full)+}/k_{obs}^{(1\text{-}mis)+}$
(10 + 10)[c]	4.3×10^{-3}	3.4×10^{-3}	1.3
(8 + 8)[d]	2.5×10^{-3}	2.9×10^{-4}	8.6

[a] [probe] = 1 μM and [target] = 5 μM at 37 °C. [b] First-order rate constants were obtained as the slopes of ln P_f vs t plots where P_f is fraction of uncleaved probe and t = 0-120 min. [c] Fully matched and one-base mismatched targets are 3'-GGC ATC CTC AGA CCT GGC ACA G-5' and 3'-GGC ATC CTC AGA CCT GTC ACA G-5', respectively. [c] Fully matched and one-base mismatched targets are 3'-CAT CCT CAG ACC TGG CAC-5' and 3'-CAT CCT CAG ACC TGT CAC-5', respectively.

Fluorescence imaging of fixed *E. coli* cells. Fluorescence in-cell nucleic acid sensing requires automatic light-up upon target-probe hybridization. We devised a (8 + 8) FRET probe $\mathbf{3}^{\mathbf{FRET}}$ having a dabsyl quencher (Q) at the asterisked T moiety (Figure 1a) in the loop domain of probe 3; Q would form a FRET pair with the fluorescein fluorophore (F) attached to the base-pairing partner A at the 5'-end in the hairpin-locked off-target ground state (**A** in Scheme 1). When the probe is bound to the target (328-345 site of rRNA, Figure 1) with an expected conformational change (**B** in Scheme 1), FRET pair should be unlocked with concomitant emission of fluorescence due to opening of the hairpin-lock domain. For accuracy, the locked TASC probe should keep the off-target locked hairpin conformation in the cells (step a in Figure 4) and should change into an active form only when binding to target RNA (step c). This is not a priori guaranteed, however, in the crowded[18] multi-component in-cell environment, where the structures of the probe and target may be affected by coexisting DNAs, RNAs, and other macromolecules or intramolecularly folded more closely.[19] The present target can be taken as an illustration. The target nucleotide sequence 328-345 of *E. coli* 16S rRNA (shown in blue in Figure 4) is expected to be poorly accessible due to competing self-hybridization, wherein at least 9 (out of 18) nucleotides could form stable base pairs with nearby nucleotides to mask the target site, referring to the "non-accessible" form in Figure 4. In these circumstances, "helper" ODNs can be used; the 3'-helper (3'-CCC TCC GTC GTC ACC CCT TA-5') and the 5'-helper (3'-GTC GGT GTG ACC TTG ACT CT-5') are expected to bind to the flanking sites (308-327 and 346-365, respectively) adjacent to the target site (328-345) to open the latter, i.e., to convert it from the non-accessible form to accessible (step b in Figure 4).[20]

This is indeed the case. *E.coli* cells (K12 strain MG1655) were fixed with paraformaldehyde according to the literature protocols,[15a] washed once with PBS, and incubated with FRET locked TASC probe $\mathbf{3}^{\mathbf{FRET}}$ (1 μM) at 37 °C in a Tris-HCl hybridization buffer (50 mM pH 7.2) containing Mg^{2+} (25 mM) and SDS (0.1%). The cell suspension was directly spotted on a glass slide without post-hybridization washing for fluorescence microscopic observation. In the absence of the helpers, the cells were nonfluorescent (Figure 5a and 6a for confocal and widefield microscopic images, respectively), showing that the probe still keeps a hairpin-locked, off-target conformation in the cells as in solution.[21] In marked contrast, when the helper ODNs were present, the cells became fluorescent (Figure 5b and 6b), indicating that sequence-selective hybridization of unlocked probe $\mathbf{3}^{\mathbf{FRET}}$ and subsequent self-cleavage[22] took place exactly at the target site (step c in Figure 4).[23]

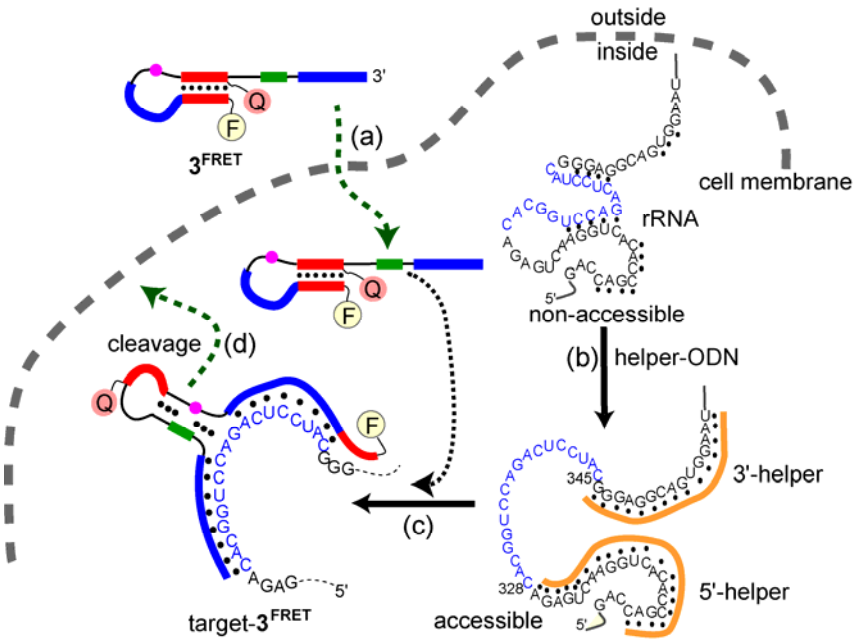


Figure 4. Schematic illustration of the working mechanism of FRET TASC probe **3**$^{\mathbf{FRET}}$ in the *E. coli* cell. Locked probe enters the cell through paraformaldehyde-permeabilized membrane (step a), where it hybridizes (step c) with 16S rRNA, assisted (step b) by helper ODNs in orange, with an open or accessible target site (nucleotides 328-345) in blue to give a cleavage-susceptible and FRET-free target-probe complex, followed by its catalytic cleavage.

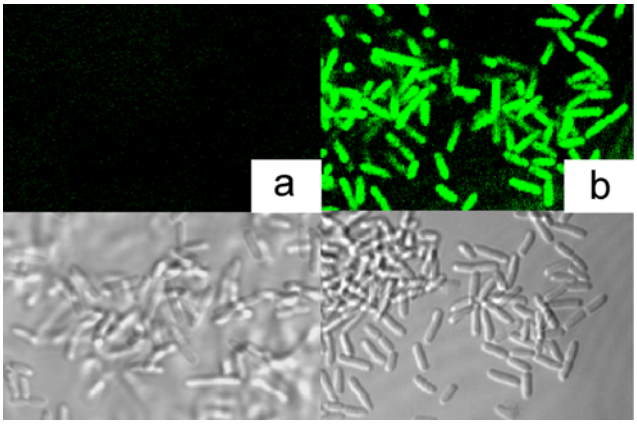


Figure 5. Optical micrographs (bottom) and their confocal fluorescence images (top) of *E. coli* cells incubated for 6 h with FRET locked probe $\mathbf{3}^{\mathbf{FRET}}$ in a Tris-HCl (50 mM, pH 7.2) buffer containing Mg^{2+} (25 mM) and 0.1% SDS (a) without or (b) with helper ODNs (10 μM each).

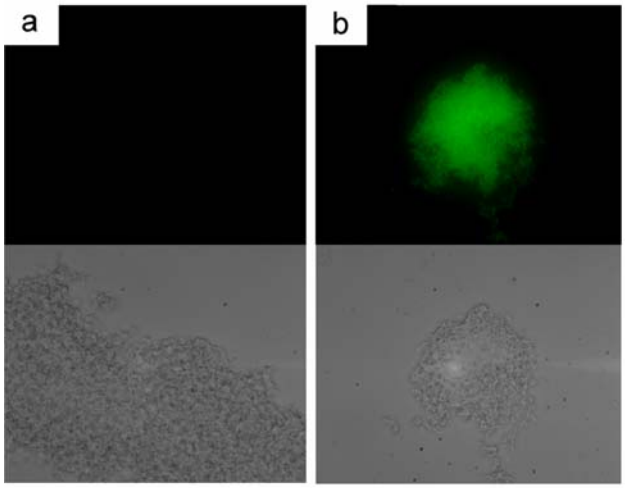


Figure 6. Optical micrographs (bottom) and their widefield fluorescence images (top) of *E. coli* cells incubated for 6 h with FRET locked probe $\mathbf{3}^{\mathbf{FRET}}$ in a Tris-HCl (50 mM, pH 7.2) buffer containing Mg^{2+} (25 mM) and 0.1% SDS (a) without or (b) with helper ODNs (10 μM each).

Conclusion

In summary, the lock strategy gives rise to a sort of generality or versatility of the TASC (target-assisted self-cleavage) approach to the nucleic acid sensing. While it works relatively fine with k_{on}/k_{off} = 65 and a high level of single nucleotide discrimination for targets of particular sequence and length, the general efficiency in terms of allosteric factor of ≥10 still remains to be improved, particularly in view of catalytic or amplified sensing, by stabilizing the off-target lock structure (**A** in Scheme 1) as well as by geometrically manipulating the on-target DNAzyme action on the cleavage site (**B**). The rRNA imaging of *E. coli* cells may be taken as a sign of potentiality of the dissociative TASC probes working in crowded intracellular environments. At the same time, the fate of the sensing-responsible probe-fragment becomes a new concern. With these concerns taken in mind, we step toward further work for genetic sensing/imaging of the cells, which may have a variety of applications such as non-destructive detection of bacteria and viruses and *in vivo* selection of drugs.[24]

Experimental Section

Preparation of TASC probes. TASC probes with fluorescein labelling at the 5'-end were prepared on a DNA/RNA synthesizer (ABI 392, Perkin-Elmer) using phosphoramidites (Glen Research) of Bz-dA, ^{i}Bu-dG, dT, and Ac-dC, fluorescein, Ac-A-TOM as a source of rA, and, in case of FRET TASC probe **3**, dabsyl-dT as a FRET partner. The CPG solid support was incubated in 1.5 mL of a methylamine solution, prepared by mixing 1 volume of 40% aqueous methylamine (Fluka) and 1 volume of 33% ethanolic methylamine (Fluka), for 6 h at 35 °C to liberate the oligonucleotide with concomitant deprotection. The mixture was lyophilized to yield a yellow solid (orange solid in the case of FRET TASC probe), which was dissolved in 1 mL of 1 M tetrabutylammonium fluoride in THF (Sigma) and left at 50 °C for 10 min and at 35 °C for 6 h to be freed from the TOM protecting group. Addition of 1 mL of 1 M Tris buffer was followed by evaporation of most of THF and addition of another 1 mL of Tris buffer. The fully-deprotected TASC probes were purified on NAP-25 desalting column (Amersham), followed by electrophoresis on an 8% polyacrylamide denaturing gel containing 7 M urea. The TASC probes were extracted from the gel with 10 mL of H_2O overnight at 37 °C and then dialyzed using MWCO 1000 Float-A-Lyzer (Spectrum) in a large volume of H_2O. The solution was lyophilized to give TASC probes.

Analysis of self-cleavage reactions. Targets ODN1 and ODN2 were purchased from Qiagen. All of the reactions were carried out at [probe] = 1 μM in 100 μL of Tris-HCl buffer (50 mM, pH 7.2) containing $MgCl_2$ (25 mM) at 37 °C with [target] = 5 or 0 μM. The reaction was quenched at an appropriate time interval by mixing with 800 μL of cold ethanol. The mixture was stored at -80 °C for 30 min and then centrifuged at 9,500 g for 30 min. The ethanol-precipitates were dissolved in a 1:1 mixture of water (5 μL) and loading buffer (1 mM EDTA in formamide, 5 μL) and analyzed by electrophoresis (ATTO PARIDAS AE-6200) on an 8% polyacrylamide denaturing gel containing 7 M urea. The reactions were monitored by fluorescence imaging. The yields of cleavage were assayed by referring to the relative fluorescence intensities of intact TASC probe and its fluorescent fragment after careful calibration with known concentrations of authentic specimen, using an ATTO densitograph AE-6920 equipped with a lane & spot analyzer ver. 6. The first-order rate constants were obtained as slopes of *ln*(fraction of uncleaved probe) vs *t* plots in the conversion range indicated. Each data point was obtained as an average of at least

three independent data sets and the goodness-of-fit of the line to the data points was evaluated by the correlation coefficient of $R > 0.97$ as the criterion for an acceptable fit.

Measurements of melting temperature. Oligonucleotides (2 μM each) were dissolved in a Tris-HCl buffer (50 mM, pH 7.2) containing $MgCl_2$ (25 mM). The mixture was heated for 5 min at 95 °C and cooled down slowly to 4 °C. The thermal denaturation profile was recorded on a UV-Visible Spectrophotometer UV-1650 PC (Shimazu) with a Peltier temperature controller. The absorbance of the sample was monitored at 260 nm from 4 °C to 90 °C with a heating rate of 1 °C/min. The T_m value was determined as the maximum in a plot of $\Delta A_{260}/\Delta T$ vs temperature.

Cell fixation. *E. coli* cells (K12 strain MG1655) were grown in 6 mL of LB broth at 37 °C until the optical density (OD_{600}) of the cell suspension became 0.5. Then, the suspension was quickly chilled on ice for 5 min and aliquots (1 mL) were taken into 1.5 mL eppen-tubes. Cells were harvested by centrifugation at 9,500 g for 5 min at 0 °C. After centrifugation, the supernatant was removed by pipetting and cells were washed once with 1 mL of PBS. For fixation, cells were resuspended in 1 mL of 4% paraformaldehyde (Aldrich) in PBS solution (filter sterilized, pH 8.0 adjusted by 1N NaOH) and the suspension was stored for 1 h at room temperature. After fixation, the suspension was centrifuged at 9,500 g for 5 min at 0 °C and the supernatant was removed by pipetting. Precipitated cells were washed once with 1 mL PBS. The fixed cells were resuspended in 1 mL of 50% ethanol and stored at –20 °C.

Fluorescence imaging of FRET TASC probe in fixed *E.coli* cells. A 167 μL portion of fixed cell suspension stocked in 50% ethanol, prepared according to the method described above, was taken into 1.5 mL eppen-tube and centrifuged at 9,500 g for 5 min at 0 °C. The precipitated fixed cells were washed once with 100 μL PBS and resuspended in 100 μL of hybridization buffer of 50 mM Tris-HCl containing 25 mM $MgCl_2$ and 0.1% SDS with or without 3'/5'-helper oligonucleotide probes (10 μM each, final strand concentration). To the suspension was added FRET TASC probe **3** (1 μM, final strand concentration) and the mixture was incubated at 37 °C. A 3 μL portion of the mixture was taken after gentle centrifugation for optical and fluorescence microghraphy. Confocal and widefield fluorescence images were obtained using Fluoview FV500 (Olympus) and IX70 fluorescence microscope (Olympus), respectively.

References

(1) (a) Schafer, A. J.; Hawkins, J. R. *Nat. Biotechnol.* **1998**, *16,* 33-39. (b) Twyman, R. M. *Curr. Top. Med. Chem.* **2004,** *4,* 1423-1431.

(2) Kwok, P. Y.; Chen, X. *Curr. Issues Mol. Biol.* **2003,** *5,* 43-60.

(3) (a) Mihalcescu, I.; Hsing, W.; Leibier, S. *Nature* **2004,** *430,* 81-85. (b) Johnson, C. H.; *Nature* **2004,** *430,* 23. and references therein.

(4) For recent reports on PCR-free homogeneous nucleic acid sensing, see for example, (a) Storhoff, J. J.; Elghanian, R.; Mucic, R. C.; Mirkin, C. A.; Letsinger, R. L. *J. Am. Chem. Soc*. **1998,** *120,* 1959-1964. (b) Patolsky, F.; Lichtenstein, A.; Willner, I. *J. Am. Chem. Soc.* **2001,** *123,* 5194-5205. (c) Xu, Y.; Karalkar, N. B.; Kool, E. T. *Nat. Biotechnol.* **2001,** *19,* 148-152. (d) Ranasinghe, R. T.; Brown, T.; Browon, L. J. *Chem. Commun.* **2001,** 1480-1481. (e) Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003,** *125,* 1195-1202. (f) Okamoto, A.; Kanatani, K.; Saito, I. *J. Am. Chem. Soc.* **2004,** *126,* 4820-4827. (g) Hwang, G. T.; Seo, Y. J.; Kim, S. J.; Kim, B. H. *Tetrahedron Lett.* **2004,** *45,* 3543-3546. (h) Cai, J.; Li, X.; Yue, X.; Taylor, J. S. *J. Am. Chem. Soc.* **2004,** *126,* 16324-16325. (i) Nakatani, K. *ChemBioChem* **2004,** *5,* 1623-1633. and references therein.

(5) For recent reviews on the FISH technology, see: (a) Moter, A.; Gobel, U. B. *J. Microbiol. Methods* **2000,** *41,* 85-112. (b) Lipski, A.; Friedrich, U.; Altendorf, K. *Appl. Microbiol. Biotechnol.* **2001,** *56,* 40-57. (c) Levsky, J. M.; Singer, R. H. *J. Cell Sci.* **2003,** *116,* 2833-2838.

(6) (a) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002,** *124,* 2096-2097. (b) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002,** *124,* 9686-9687. (c) Sando, S.; Abe, H. Kool, E. T. *J. Am. Chem. Soc.* **2004,** *126,* 1081-1087.

(7) (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996,** *14,* 303-308. (b) Tyagi, S.; Bratu, P.; Kramer, F. R. *Nat. Biotechnol.* **1998,** *16,* 49-53. (c) Bonnet, G.; Tyagi, S.; Libchaber, A. Kramer, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96,* 6171-6176. (d) Bratu, D. P.; Cha, B. J.; Mhlanga, M. M.; Kramer, F. R. Tyagi, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100,* 13308-13313. (e) Tan, W.; Wang, K.; Drake, T. J. *Curr. Opin. Chem. Biol.* **2004,** ***8,*** 547-553. and references therein.

(8) For recent studies on intracellular trafficking of mRNA using controlled amounts of FISH probes, see for example: (a) Fusco, D.; Bertrand, E.; Singer, R. H. *Prog. Mol. Subcell. Biol.* **2004,** *35,* 135-150. and references therein. (b) Shav-Tal,

Y.; Darzacq, X.; Shenoy, S. M.; Fusco, D.; Janicki, S. M.; Spector, D. L.; Singer, R. H. *Science* **2004,** *304,* 1797-1800.

(9) Stojanovic, M. N.; de Prada, P.; Landry, D. W. *ChemBioChem* **2001,** *2,* 411-415. (b) Stojanovic, M. N.; Mitchell, T. E.; Stefanovic, D. *J. Am. Chem. Soc.* **2002,** *124,* 3555-3561. (c) Stojanovic, M. N.; Stefanovic, D. *Nat. Biotechnol.* **2003,** *21,* 1069-1074.

(10) Hartig, J. S.; Grune, I.; Najafi-Shoushtari, S. H.; Famulok, M. *J. Am. Chem. Soc.* **2004,** *126,* 722-723.

(11) For recent examples of "amplifiable" oligonucleotide reactions on DNA templates, see: (a) Gat, Y.; Lynn, D. G. *Biopolymers* **1998,** *48,* 19-28. (b) Luther, A.; Brandsch, R.; von Kiedrowski, G. *Nature* **1998,** *396,* 245-248. (c) Albagli, D,; Van Atta, R.; Cheng, P.; Huan, B.; Wood, M. L. *J. Am. Chem. Soc.* **1999,** *121,* 6954-6955. (d) Brunner, J.; Mokhir, A.; Kraemer, R. *J. Am. Chem. Soc.* **2003,** *125,* 12410-12411. (e) Abe, H.; Kool, E. T. *J. Am. Chem. Soc.* **2004,** *126,* 13980-13986.

(12) For recent reviews on allosteric DNA/RNAzymes, see: (a) Soukup, G. A.; Breaker, R. R. *Trends Biotechnol.* **1999,** *17,* 469-476. (b) Kuwabara, T.; Warashima, M.; Taira, K. *Curr. Opin. Chem. Biol.* **2000,** *4,* 669-677.

(13) For examples of oligonucleotide control of DNA/RNAzyme activities, see: (a) Porta, H.; Lizardi, P. M. *Biotechnology* **1995,** *13,* 161-164. (b) Wang, D. Y.; Sen, D. *J. Mol. Biol.* **2001,** *310,* 723-734. (c) Wang, D. Y.; Lai, B. H. Y.; Den, D. *J. Mol. Biol.* **2002,** *318,* 33-43. (d) Wang, D. Y.; Lai, B. H. Y.; Feldman, A. R.; Sen, D. *Nucleic Acids Res.* **2002,** *30,* 1735-1742. (e) Refs. 9 and 10.

(14) Sando, S.; Sasaki, T.; Kanatani, K.; Aoyama, Y. *J. Am. Chem. Soc.* **2003,** *125,* 15720-15721.

(15) (a) Amann, R. I.; Krumholtz, L.; Stahl, D. A. *J. Bacteriol.* **1990,** *172,* 762-770. (b) Fuchs, B. M.; Wallner, G.; Beisker, W.; Schwippl, I.; Ludwig, W.; Amann, R. *Appl. Environ. Microbiol.* **1998,** *64,* 4973-4982.

(16) The four stem-forming base pairs adjacent to the GrAT site were varied and so chosen, based on simulation of intrastrand base-pairing, as shown in Figure 1a (CG, TA, CG, and GC) to make the probe least structured and hence most flexible in the absence of the target. This is in order to demonstrate that the lock strategy works even on such systems in lowering the off-target reactivity.

(17) Santoro, S. W.; Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997,** *94,* 4262-4266.

(18) Record, M. T.; Courtenay, E. S.; Cayley, D. S.; Guttman, H. J. *Trends Biochem. Sci.* **1998,** *23,* 190-194. and references therein.

(19) For recent reports on structural analysis of nucleic acids in crowded intracellular environments; see for example: Miyoshi, D.; Matsumura, S.; Nakano, S.; Sugimoto, N. *J. Am. Chem. Soc.* **2004,** *126,* 165-169. and references therein.

(20) Fuchs, B. M.; Glockner, F. O.; Wulf, J.; Amann, R. *Appl. Environ. Microbiol.* **2000,** *66,* 3603-3607.

(21) It is well known that oligonucleotide probes readily enter the paraformaldehyde-fixed *E. coli* cells; see for example Ref. 15a.

(22) In order to shed light on the fate of the target-probe complex target-$\mathbf{3}^{\mathbf{FRET}}$ (Figure 4), we carried out cell sensing in the presence of Na^+ in place of Mg^{2+} which was absolutely essential for cleavage. The idea behind this control run with no Mg^{2+} is that FRET-freed complex target-$\mathbf{3}^{\mathbf{FRET}}$ has no chance of self-cleavage and hence no chance of turnover, thus exhibiting one-equivalent or stoichiometric fluorescence signal at best. Actually, the fixed *E. coli* cells washed with and incubated in hybridization buffer containing probe $\mathbf{3}^{\mathbf{FRET}}$ (1 μM) and Na^+ (900 mM) indeed fluoresce but weakly than they do in the presence of Mg^{2+} (25 mM). On this Mg^{2+}/Na^+ criterion, it is most likely that the enhanced fluorescence in the Mg^{2+} case is due to multiple-turnover catalytic cleavage of the complex target-$\mathbf{3}^{\mathbf{FRET}}$ with signal amplification (step d in Figure 4), although full interpretation of the results is deferred until more information is available as to the uptake/leak-out properties of the full-length probe and its cleavage fragments.

(23) The self-cleavage of the FRET probe $\mathbf{3}^{\mathbf{FRET}}$ in homogenous solution is hardly affected by the helper ODNs.

(24) Kawasaki, H.; Onuki, R.; Suyama, E.; Taira, K. *Nat. Biotechnol.* **2002,** *20,* 376-380.

Chapter 2

Doubly Catalytic Sensing of Nucleic Acid Sequences in Prokaryotic Cell-Free Translation System using Artifitial Riboregulation System

Abstract

A molecular-beacon type riboregulator (mRNA) was applied to multiply catalytic gene-sensing. It consists of a reporter gene for firefly protein luciferase and, upstream thereof, a regulator hairpin domain composed of an RBS/anti-RBS stem (RBS = ribosome binding site) and a loop which is complementary to target. The hairpin and hence the RBS are rendered open upon binding of a target oligonucleotide of the human CC chemokine receptor 5 (CCR5) sequence in a prokaryotic cell-free translation system (10 μL) to ignite ribosomal catalytic translation, or transcription/translation when using a DNA form of the probe, to produce luciferase which is assayed by a catalytic chemiluminescence reaction. The sensing, using an unmodified RNA or even dsDNA as a probe with a chemiluminescence output, is thus doubly catalytic or amplifiable with sensitivity at ≤50 fmol in respect to target with 4.5 fmol (1 ng/μL) of probe and a single-nucleotide resolution.

Introduction

As the number of identified genetic markers continues to grow, there has been much current interest in rapid and simple gene-sensing.[1] One of most successful and general methods of homogeneous nucleic acid sensing so far reported[2] is to use molecular beacons (MBs),[3] which are hairpin-shaped oligonucleotide probes with an essential target-complementary sequence held in a loop domain linked to a FRET pair at the 5' and 3'-ends or, in newer versions, to a ribozyme/DNAzyme sequence[4-8] or a redox center.[9] The recent inhibitor-DNA-enzyme conjugate[10] may also be mechanistically related to MBs. Binding of target results in opening of the hairpin structure to restore fluorescence or switch-on the enzymatic (ribozyme/DNAzyme[4-8] or proteinous[10]) or electrochemical[9] *catalysis*, thus allowing signal amplification.[11] The present work is concerned about linkage of MB motif to a protein-translation regulator for multiply catalytic gene-sensing. We report here that an HIV-related chemokine receptor sequence can be sensed by chemiluminescence using an unmodified RNA or DNA probe with a sensitivity of ≤50 fmol and a single-nucleotide selectivity.

Results and discussion

Regulation of translation with a small RNA is a naturally occurring function.[12] The RpoS mRNA contains a hairpin-shaped structure to sequester the ribosome binding site (RBS) shown in red (Figure 1a).[13] Binding of a small RNA DsrA at the anti-RBS region in yellow leads to opening of the hairpin structure, makes the RBS domain accessible, and hence triggers translation of the RpoS protein (Figure 1a).[13] Recently, Collins et al. have demonstrated that translation of a simplified mRNA (riboregulator) cis- or hairpin-repressed in this manner can be activated by an external, i.e., trans-acting small RNA complementary to the stem-loop region including the anti-RBS.[14] Inspired by these findings, we designed a MB-type riboregulator capable of multiple amplification. Luciferase was a sensing output of choice because of the high sensitivity and good linearity of the chemiluminescence assay based thereupon (Figure 2). Thus, the riboregulator (mRNA) probe[15] (Figure 1b) contains the reporter gene for luciferase and, upstream thereof, a regulatory hairpin domain composed of a 8-bp stem (bp = base-pair) involving a 6-nt RBS in red (nt = nucleotide) and a 19-nt loop, 16-nt of which (Figure 1b, colored in yellow) is complementary to the 620-635 region (Figure 4, in yellow) of the human CC chemokine receptor 5 (CCR5). Binding of 25-mer CCR5 oligodeoxynucleotide (ODN^{full}) as a target with concomitant liberation of RBS would ignite ribosomal *catalytic* production (translation) of luciferase which in turn is capable of *catalytic* production of chemiluminescence. Accidentally, the underlined 5 nucleotides in the 3'-end of the target are complementary, if the central A is bulged, to the anti-RBS region (UCCU). The former may thus serve as an anti-(anti-RBS) to cooperate with the designed 16-bp target-probe hybridization to open the hairpin in competition with the intrastrand 8-bp or (8 + 5)-bp[16] (Figure 3) stem formation.

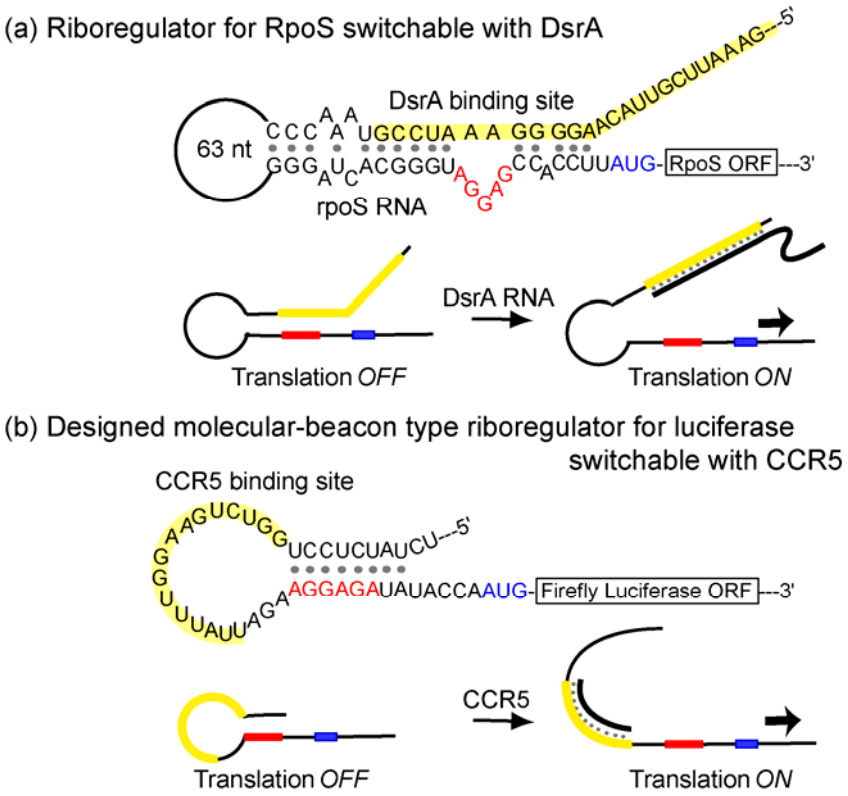


Figure 1. Illustration of (a) natural riboregulator for RpoS and (b) designed molecular-beacon type riboregulator for luciferase switchable with DsrA and CCR5, respectively. The ribosome binding site (RBS), start codon, and target binding site are colored in red, blue, and yellow, respectively.

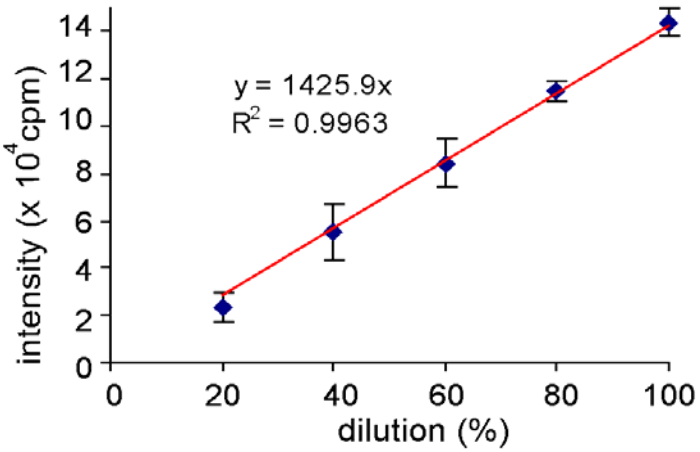


Figure 2. Linearity of chemiluminescence intensity and amount of luciferase under the assay conditions used here. A 2.5 μL aliquot of a serially diluted (20, 40, 60, 80, or 100%) solution (1.0 μg of riboregulator mRNA in 10 μL of a reconstituted cell-free T7-transcription/translation system (Pure System Classic I, Post Genome Institute)[23] incubated at 37 °C for 1 h) was mixed with 100 μL of luciferase assay solution (Promega) in a 96-well plate and luminescence was read 90 s after mixing with a Wallac 1420 microplate reader. Background intensity of ~0.8 × 10^4 cpm with no luciferase was subtracted from each intensity observed. Error bars are standard deviations of three independent measurements.

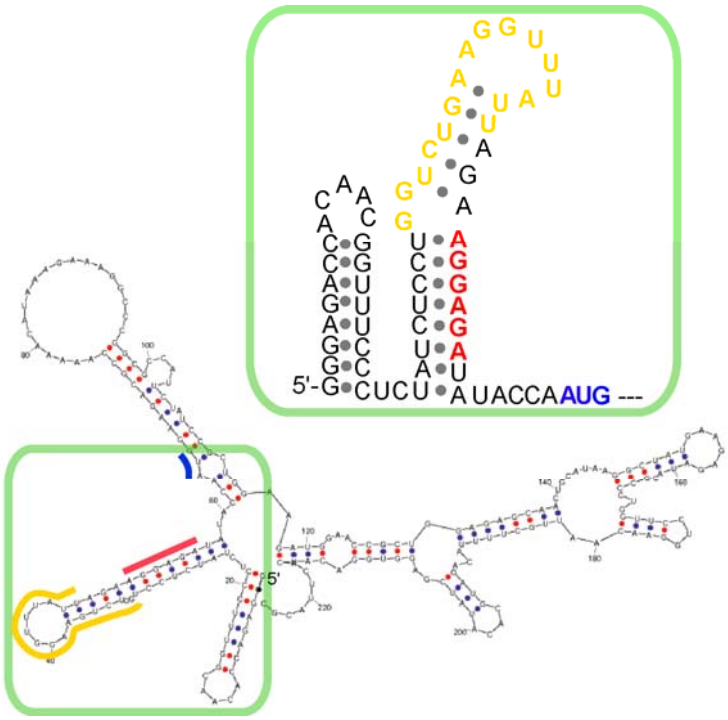


Figure 3. The predicted secondary structure of the 5'-end region of riboregulator mRNA with G-C and A/G-U pairs shown in red and blue dots, respectively, and an expansion of the stem/loop domain in a box. The 6-nt RBS, start codon, and 16-nt target-binding site are shown in red, blue, and yellow, respectively. As shown, the loop is suggested to form another 5-bp stem.

Translation of the template riboregulator mRNA (0.1 μg/μL or ~1.8 pmol in 10 μL of medium) was carried out in a reconstituted prokaryotic cell-free translation system (coupled T7-transcription/translation system, Pure System Classic I)[23] at 37 °C for 1 h in the presence (0.9-21.6 pmol) or absence of target ODN^{full} (Figure 4), followed by chemiluminescence (CL) assay of luciferase. The results are shown in Figure 5a. The off-target translation activity of the present riboregulator is still substantial[17] but the target-induced allosteric activation is rather well-behaved, showing a linear dependence on the amount of target in the range of ≤7 pmol (≤4 equivalents) until saturation is reached thereafter.

The CCR5 is known as an HIV-1 coreceptor and its sequence variations are assumed to be possible haplotype markers of acceleration/retardation of the HIV-1 disease;[18] the sequence selectivity is an importance criterion here. Non-complementary d(ACTG)$_6$AC and 2-base mismatched ODN^{2mis} (T for C and A for T at position 627 and 628, respectively; Figure 4) exhibit no notable enhancement with I_{on}/I_{off} = 112% (Figure 5b, lane 3) and 98%, respectively. One-nucleotide selectivity depends on the identity of mismatch. ODN^{1misA} (A for C at position 627; Figure 4) hardly shows activation in a similar manner (I_{on}/I_{off} =114%; Figure 5b, lane 5). On the other hand, ODN^{1misT} (Figure 4) which would form a relatively stable GT-mismatch in the resulting hetero-duplex is moderately activating (I_{on}/I_{off} = 227%; Figure 5b, lane 4). The sensitivity or detection limit in reference to I_{on}/I_{off} can be enhanced by using a smaller amount (1 ng/μL) of riboregulator probe (to minimize target-independent translation) in combination with a sensitive luminometer (Lumat LB 9507) to 50 fmol with respect to ODN^{full} using 4.5 fmol of the probe for a 2.5 μL aliquot (I_{on} = 6,362 ± 475 cpm, I_{off} = 3,696 ± 42 cpm, I_{on}/I_{off} = 1.7). This detection limit is lower than that of typical fluorescence-based MBs.

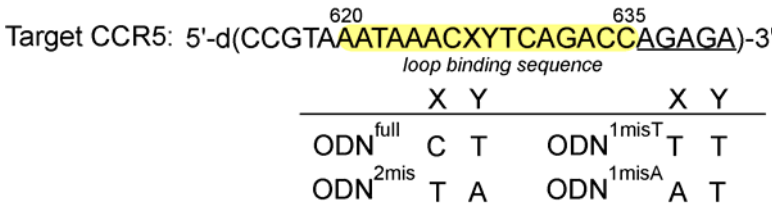


Figure 4. The sequences of CCR5 gene and target ODNs. The target region of CCR5 gene is colored in yellow.

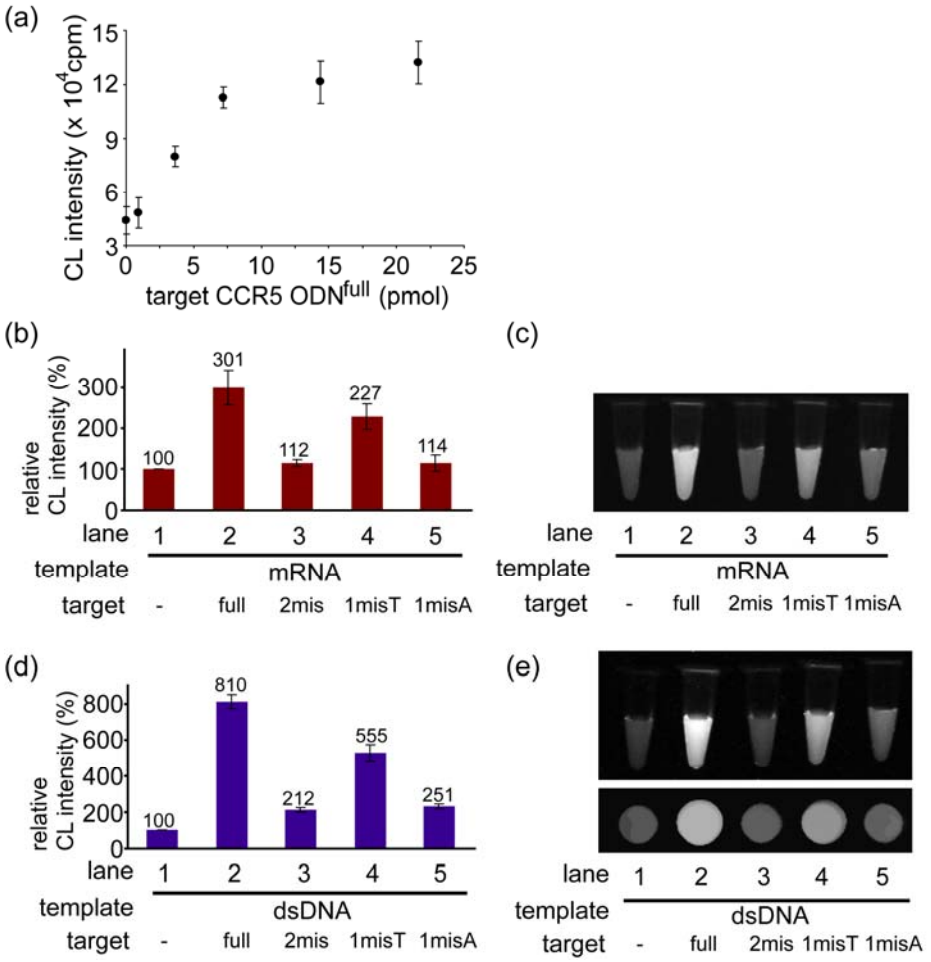


Figure 5. Chemiluminescence (CL) assay of CCR5 sequences for a 2.5 μL aliquot of translation mixture (10 μL) after treatment with 100 μL of luciferase assay solution (Promega). (a) Change in CL intensities as a function of amount of full-match target ODN^{full} with the riboregulator mRNA (1.8 pmol) as a probe. (b), (d) Relative CL intensities in the presence (18 pmol) or absence of full-match and mismatch ODN targets with mRNA (1.8 pmol) (b) or dsDNA (0.2 pmol) (d) as a probe. (c), (e) CL images for samples in Figure 5b and 5d in vials (5c, 5e top) or in a 96-well plate (5e bottom).

As expected, the probe can also be provided in the form of DNA, from which the mRNA (riboregulator) will be *in situ* transcribed. It works better (Figure 5d). Under the coupled-transcription/translation conditions in the same cell-free system as above, the precursor dsDNA[15] in a much smaller amount (0.2 pmol) gives an improved value of I_{on}/I_{off} = 810% (lanes 1 and 2 in Figure 5d) as compared with I_{on}/I_{off} = 301% when using the pre-transcribed/pre-purified mRNA (1.8 pmol) (lanes 1 and 2 in Figure 5b).[19] The overall feature of the sequence selectivity (lanes 2-5 in Figure 5d) is similar, but in a more pronounced manner, to that in the case of mRNA as a probe (lanes 2-5 in Figure 5b). The selectivity can be readily appreciated using an imager (VersaDoc 3000) (Figure 5c for mRNA and 5e for dsDNA) for samples in vials (5c, 5e top) or a 96-well plate (5e bottom). An explanation for the better performance of dsDNA is to assume that the concentration of mRNA *gradually* transcribed could be kept low to make the contribution of the target-dependent translation more important. The transcription process itself is hardly affected, as expected, by the target (Figure 6).

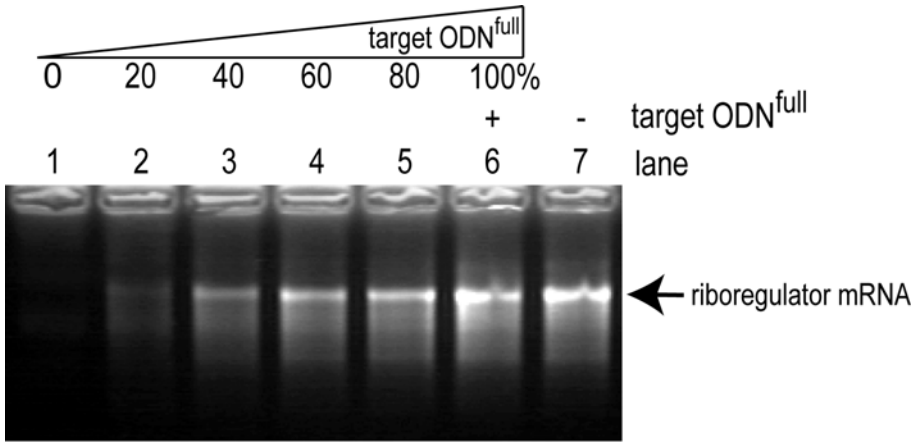


Figure 6. Gel analysis of riboregulator mRNA transcribed from dsDNA template (~0.2 pmol) in the presence (18 pmol) (lane 6) and absence (lane 7) of target ODN^{full} in 10 μL of a reconstituted cell-free T7-transcription/translation system (Classic 1, Post Genome Institute) got rid of ribosome and tRNA at 37 °C for 1 h. Lanes 1-5 are for serially diluted (0, 20, 40, 60, or 80%) solutions of the parent transcription solution in lane 6. There is no notable difference in the amounts of mRNA transcribed in lanes 6 and 7, indicating that target ODN^{full} hardly affects the transcription.

Conclusion

In summary, we, for the first time, applied a luminescence-linked riboregulator system for genotyping. While the allosteric factor (I_{on}/I_{off}) remains to be improved,[20] the present system possesses what are needed for good sensing such as linearity, sensitivity (≤50 fmol of target with 4.5 fmol of mRNA probe),[21] selectivity (1-nt or at least 2-nt), and simplicity in particular. It is indeed remarkable that the sensing by chemiluminescence can be achieved with an unmodified RNA or even stabler and readily prepared/handled dsDNA; all that are needed in addition thereto are an assay kit and a cell-free translation system, both of which are commercially available. In this regard, the present method would be cell-friendly and suited for sensing in the cell,[22] where the intrinsic in-cell translation system could be directly available for a plasmid-based dsDNA probe introduced via transfection. It is also interesting to note that the output of riboregulator can be any type of enzymatic activities, thus allowing not only detection of a particular genotype but also induction of an appropriate phenotype that responds thereto in a pharmaceutical sense. Further work is now under way along these lines.

Experimental Section

Preparation of mRNA probe. The template dsDNA containing the MB domain and the T7 promoter was prepared by two-step PCR amplification using the luciferase gene region of the pGL3 vector (Promega); first PCR (insertion of MB domain) with forward primer 5'-GGT CTG AAG GTT TAT TAG AAG GAG ATA TAC CAA TGG AAG ACG CCA AAA ACA TA-3' and reverse primer 5'-TAT TCA TTA CAC GGC GAT CTT TCC G-3', second PCR (insertion of T7 promoter) with forward primer 5'-GAA ATT AAT ACG ACT CAC TAT AGG GAG ACC ACA ACG GTT TCC CTC TAT CTC CTG GTC TGA AGG TTT A-3' and reverse primer 5'-TAT TCA TTA CAC GGC GAT CTT TCC G-3'. These primers were purchased from Gene Design Inc.. The first- and the second-PCR products were purified by gel electrophoresis. The MB-mRNA (total 1722-nt) was transcribed using T7-MegaShortScript (Ambion) and purified by the RNeasy kit (Qiagen).

Translation of mRNA probe or transcription/translation of dsDNA probe in prokaryotic translation system. Translation or transcription/translation was carried out using T7-coupled reconstituted *E. coli* translation system[23] (Pure System Classic I, Post Genome Institute CO. Ltd.). A mRNA probe (1.8 pmol) or a dsDNA template (0.2 pmol), prepared as above, was incubated in 10 μL of the Pure System translation solution in the presence of a target ODN (0 or 18 pmol) at 37 °C for 1 h, followed by luciferase assay. Target ODNs were purchased from Gene Design Inc..

Luciferase assay. Translation solution (2.5 μL aliquot) was mixed with 100 μL of Luciferase Assay Reagent (100 μL, Promega). Chemiluminescence intensities were observed by Wallac 1420 microplate reader for 96-well plate assay. Error bars are standard deviations of three independent measurements. Chemiluminescence images were taken by VersaDoc 3000.

References

(1) (a) Schafer, A. J.; Hawkins, J. R. *Nat. Biotechnol.* **1998,** *16*, 33-39. (b) Twyman, R. M. *Curr. Top. Med. Chem.* **2004,** *4*, 1423-1431.

(2) Nakatani, K. *ChemBioChem* **2004,** *5*, 1623-1633. and references therein.

(3) (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996,** *14*, 303-308. (b) Tyagi, S.; Bratu, P.; Kramer, F. R. *Nat. Biotechnol.* **1998,** *16*, 49-53. (c) Bonnet, G.; Tyagi, S.; Libchaber, A.; Kramer, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96*, 6171-6176. (d) Bratu, D. P.; Cha, B.-J.; Mhlanga, M. M.; Kramer, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100*, 13308-13313. (e) Brunner, J.; Kraemer, R. *J. Am. Chem. Soc.* **2004,** *126*, 13626-13627.

(4) (a) Stojanovic, M. N.; de Prada, P.; Landry, D. W. *ChemBioChem* **2001,** *2*, 411-415. (b) Stojanovic, M. N.; Mitchell, T. E.; Stefanovic, D. *J. Am. Chem. Soc.* **2002,** *124*, 3555-3561. (c) Stojanovic, M. N.; Stefanovic, D. *Nat. Biotechnol.* **2003,** *21*, 1069-1074.

(5) Hartig, J. S.; Grune, I.; Najafi-Shoushtari, S. H.; Famulok, M. *J. Am. Chem. Soc.* **2004,** *126*, 722-723.

(6) Sando, S.; Sasaki, T.; Kanatani, K.; Aoyama, Y. *J. Am. Chem. Soc.* **2003,** *125*, 15720-15721.

(7) Liu, J. W.; Lu, Y. *J. Am. Chem. Soc.* **2003,** *125*, 6642-6643.

(8) Xiao, Y.; Pavlov, V.; Niazov, T.; Dishon, A.; Kotler, M.; Willner, I. *J. Am. Chem. Soc.* **2004,** *126*, 7430-7431.

(9) Fan, C. H.; Plaxco, K. W.; Heeger, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100*, 9134-9137.

(10) Saghatelian, A.; Guckian, K. M.; Thayer, D. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2003,** *125*, 344-345.

(11) Tan, W.; Wang, K.; Drake, T. J. *Curr. Opin. Chem. Biol.* **2004,** *8*, 547-553.

(12) Brantl, S. *Biochim. Biophys. Acta* **2002,** *1575*, 15-25.

(13) Majdalani, N.; Cunning, C.; Sledjeski, D.; Elliott, T.; Gottesman, S. *Proc. Natl. Acad. Sci. U.S.A.* **1998,** *95*, 12462-12467.

(14) Isaacs, F. J.; Dwyer, D. J.; Ding, C.; Pervouchine, D. D.; Cantor, C. R. J.; Collins, J. *Nat. Biotechnol.* **2004,** *22*, 841-847.

(15) The anti-RBS site in the MB domain is followed by a 23-nt stem-forming sequence often added to improve the stability of the transcribed RNA.

(16) The Zuker Mfold program (Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406-3415.) predicts that the loop domain forms another 5-bp stem (Figure 3).

(17) The off-target activity can be >95% suppressed by using a 28-mer ODN 5'-ATC TCC TTC TAA TAA ACC TTC AGA CCA G-3' trans-acting on the stem/loop domain as an RBS sequester.

(18) Gonzalez, E.; Bamshad, M.; Sato, N.; Mummidi, S.; Dhanda, R.; Catano, G.; Cabrera, S.; McBride, M.; Cao, X. -H.; Merrill, G.; O'Connell, P.; Bowden, D. W.; Freedman, B. I.; Anderson, S. A.; Walter, E. A.; Evans, J. S.; Stephan, K. T.; Clark, R. A.; Tyagi, S.; Ahuja, S. S.; Dolan, M. J.; Ahuja, S. K. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96,* 12004-12009.

(19) The absolute target-off chemiluminescence intensities in lane 1 of Figure 5b and that in 5d were similar to each other.

(20) The hairpin-locked structure would be generally stabilized by a longer stem and destabilized by the presence of an intra-strand anti-(anti-RBS) site and also by helicase activity if any in case of in-cell sensing. Contrary to simple MBs, the present riboregulator is a big molecule, where such factors as mentioned above should be optimized as a whole to give a high level of allosteric (on/off) factor.

(21) Some codons in the template of *eukaryotic* protein luciferase turn out to be rare codons for a *prokaryotic* translation system. Codon optimization as regards degeneracy would further improve the translation efficiency and hence sensitivity.

(22) (a) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002,** *124*, 9686-9687. (b) Sando, S.; Abe, H.; Kool, E. T. *J. Am. Chem. Soc.* **2004,** *126*, 1081-1087. (c) Cai, J.; Li, X.; Yue, X.; Taylor, J. S. *J. Am. Chem. Soc.* **2004,** *126*, 16324-16325.

(23) Shimizu, Y.; Inoue, A.; Tomita, Y.; Suzuki, T.; Yokonaga, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001,** *19*, 751-755.

Chapter 3

Sensing of Nucleic Acid Sequences with Unmodified RNA as a Probe: Visible Sensing of DNA Sequences Using an RNase H Activity Coupled Riboregulation System

Abstract

DNA sensing at a single nucleotide resolution is achieved by using a hairpin-shaped, unmodified (unlabeled) RNA probe or the precursor dsDNA in a prokaryotic cell-free translation medium. The molecular-beacon-like probe (MB-mRNA) consists of a loop region which is complementary to the target sequence and a stem composed of a ribosome-binding site (RBS) and its docking domain; the RBS is followed by the gene for a reporter protein such as luciferase and β-galactosidase. Target binding at the loop region opens the hairpin to make RBS accessible by the ribosome to start translation of the reporter protein. This sensing system is signal-amplifying by virtue of catalytic DNA-to-RNA transcription when using a dsDNA probe, catalytic RNA-to-protein translation, catalytic signal transduction by the enzymatic reaction of the translated reporter protein and, in the presence of RNase H, catalytic or even irreversible translation-activation of the target-probe heteroduplex.

Introduction

As the amount of identified genetic information continues to grow, there has been much current interest in rapid and simple gene sensing with high sequence selectivity.[1] Among various types of homogeneous sensing methods,[1,2] including a representative molecular beacon (MB; Figure 1a),[2a,b] several new strategies have enabled catalytic sensing of a target sequence through the signal-amplifying process over the equivalent stoichiometry of the target molecule through a signal-amplifying process.[3–12] These strategies include DNA/RNAzyme-[3–5] or enzyme-coupled[6,7] approaches, template-directed catalytic chemical reaction probes,[8,9] and a method with electrochemical catalysis[10] or mechanistically controllable magnetic particles.[11] As already stated in chapter 2, we have developed a new strategy of MB-mRNA for sensitive genotyping.[12] The strategy was designed based on the system of naturally occurring[13] or engineered[14] hairpin-shaped RNAs for conformation-induced control of translation frequency, so that the sensing can be conducted using genetically encodable unmodified RNA as a probe in a typical prokaryotic translation system.

Results and discussion

The artificial translation regulation system is composed of *cis*-acting MB-like RNA structure, wherein the loop region (green in Figure 1a) is complementary to the target and the stem is composed of the identical sequence of ribosome binding site (RBS) (red) and anti-RBS or RBS-docking domain complementary thereto (pink).[12] Target binding is designed to result in the opening of the MB structure, thereby making the RBS domain accessible by the ribosome, and hence initiate the translation of a reporter gene such as luciferase. Although the MB-mRNA system actually realized the sensitive detection of nucleic acids through the double signal-amplifying process of catalytic translation and catalytic signal transduction via enzymatic reaction of the translated reporter protein, the method still has two major shortcomings. The first is sequence selectivity. The target-binding domain should be sufficiently long to open the MB (hairpin) structure, so that the probe fails to discriminate single nucleotide differences, for example, the C to T transition that forms a relatively stable GT mismatch in the target/probe heteroduplex. The second is sensitivity. As stoichiometric binding of the target to the probe is required for continuous translation of a reporter protein, the translation-activated mRNA probe is, at best, equimolar to target; this typically leads to low signal intensity, especially when the target is present in a tiny amount.

The present work is concerned with an RNase H activity coupled MB-mRNA system. The RNase H activity coupled approach allows triple catalytic sensing of target nucleic acids with improved selectivity and enhanced sensitivity. In addition, we report herein an application for the visible sensing of nucleic acids with an unmodified RNA probe in this RNase H activity coupled system.

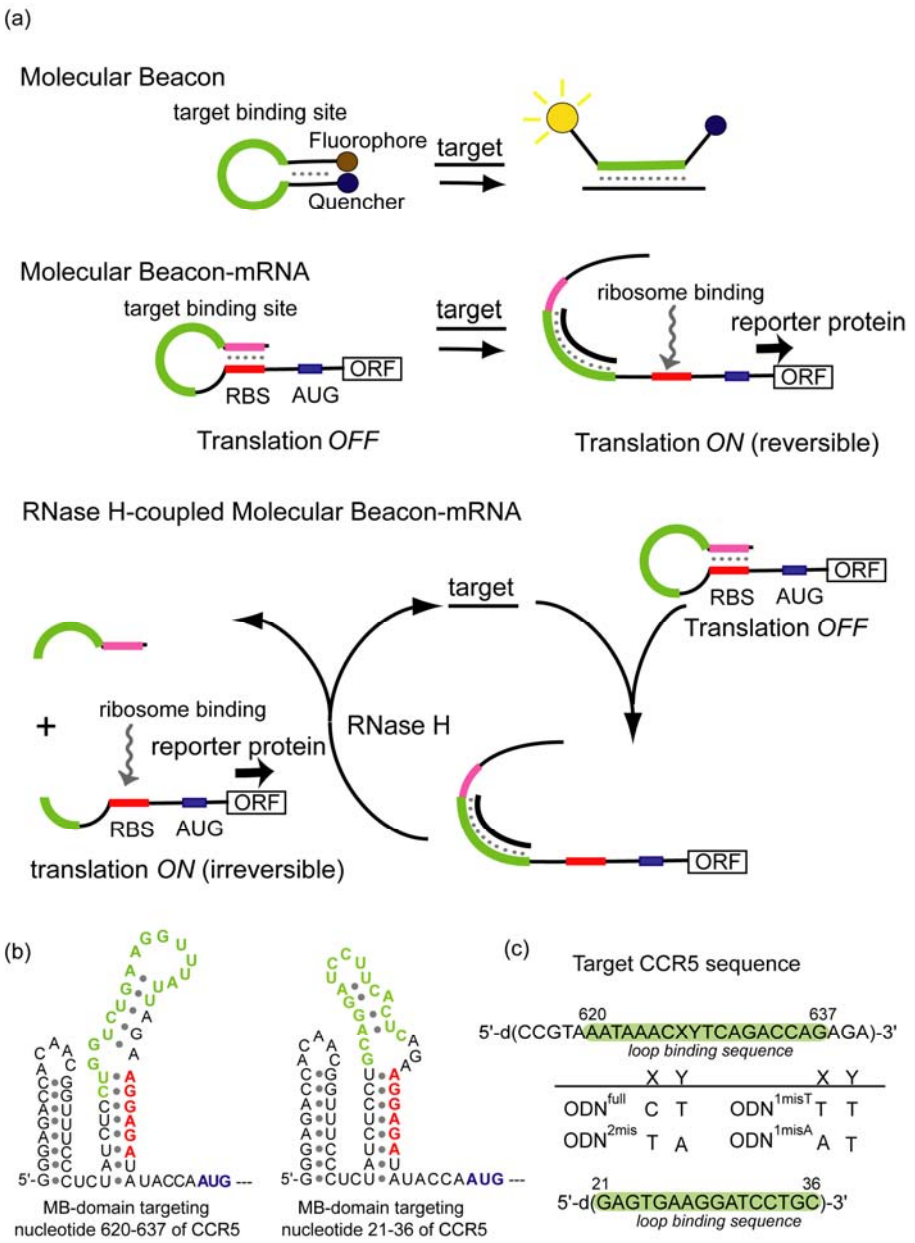


Figure 1. (a) Illustration of molecular beacon, MB-mRNA, and RNase H activity coupled MB-mRNA systems. The RBS, anti-RBS (RBS-locking site), start codon, and target binding domain are colored red, pink, blue, and green, respectively. (b) Predicted secondary structures of the 5′-end region of MB-mRNA targeting nucleotide 620–637 or 21–36 region of the human CCR5 gene sequence. (c) Sequences of the 620–637 and 21–36 regions of target human CCR5 gene.

An RNase H is an endoribonuclease that specifically hydrolyzes the phosphodiester bonds of RNA in the DNA/RNA heteroduplex.[15] Therefore, the MB-mRNA system coupled with the RNase H activity could drive an additional process for catalytic production of a translation-activated mRNA probe, wherein the target-bound loop region (target = oligodeoxynucleotide (ODN)) is digested by coexisting RNase H to release an anti-RBS stem domain away from the mRNA body, thus allowing catalytic use of target DNA (Figure 1a).

To demonstrate the designed triple catalytic gene sensing, we first used an mRNA probe composed of the MB-domain targeting the nucleotide 620–637 site (18-nt) of the human CC chemokine receptor 5 (CCR5) gene sequence (Figure 1b),[16] downstream of which is a codon-optimized luciferase reporter gene, derived from pBESTluc vector (Promega). The sensing of target CCR5 ODNs (Figure 1c) with the MB-mRNA probe (~1.8 pmol in 10 μL solution, 0.1 μg/μL) was carried out at 37 °C for 1 h in a reconstituted prokaryotic translation system (Pure System)[17,18] in the presence or absence of additional *Tth* RNase H (typically 0.01 U/μL); this was followed by chemiluminescence (CL) assay of the luciferase expressed with a 96-well microplate reader (Wallac 1420 system; Figure 2). In the absence of RNase H, an equimolar amount of fully matched target ODN^{full} afforded a moderately CL enhancement (I_{on}/I_{off} = 314%) relative to that in the absence of the target (lane 1 versus 2 in Figure 2a). An RNase H activity coupled (0.01 U/μL) translation system, under otherwise identical conditions, resulted in an enhanced translation activation that reached 615% CL enhancement (lane 3 versus 4).

Control runs indicated that RNase H indeed catalyzed the target-assisted cleavage of the probe[19,20]. The probe can also be provided in the form of dsDNA, from which the MB-mRNA will be transcribed *in situ*. Under the RNase H coupled T7-transcription/translation conditions, the precursor dsDNA (0.2 pmol) gave a further enhanced value of I_{on}/I_{off} = 1,320% for an equimolar amount of target ODN^{full} (lane 1 versus 2 in Figure 2b). Nucleic acid sensing for a different CCR5 sequence was further conducted by using a newly designed MB-type cis-repressing domain targeting nucleotides 21–36 of CCR5 (Figure 1b). As seen in Figure 3, the presence of a fully matched sequence afforded highly enhanced CL under the RNase H activity coupled conditions, with I_{on}/I_{off} = 953% (lane 2 versus 3 in Figure 3), thus partly revealing the generality of the target sequence for this triple catalytic sensing system.

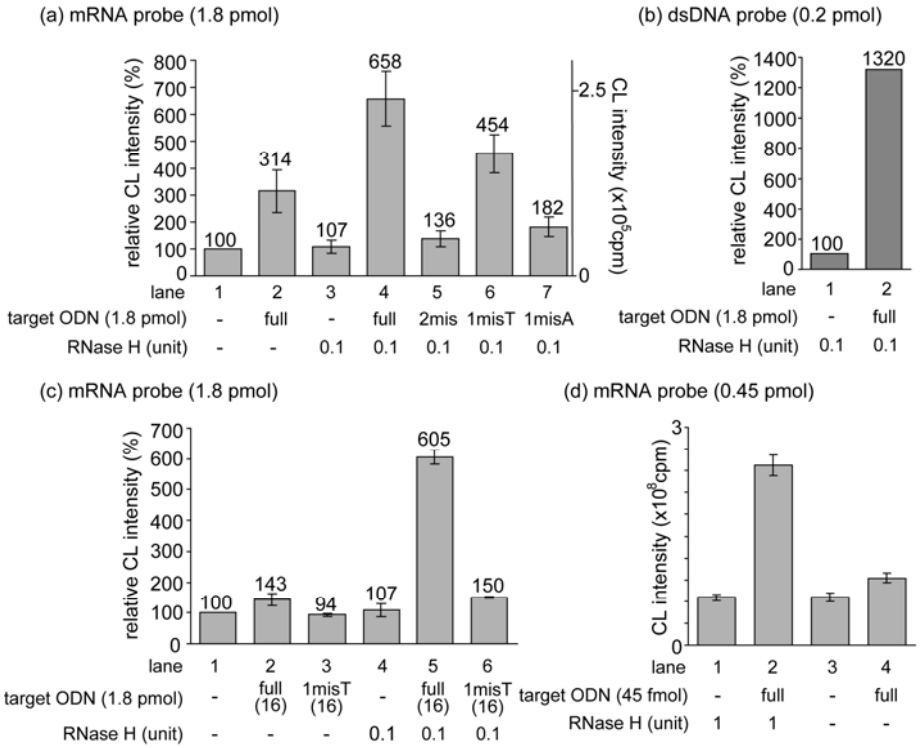


Figure 2. RNase H activity coupled sensing of the human CCR5 sequence by CL assay for a 2.5-μL aliquot of RNase H activity coupled or noncoupled translation system (10 μL) after treatment with luciferase assay solution (100 μL; Promega). (a, b) Relative CL intensities in the presence or absence of target ODNs (1.8 pmol, target site 620–637) with (a) MB-mRNA (1.8 pmol) or (b) dsDNA template (0.2 pmol) as a probe. (c) Relative CL intensities in the presence of a target ODN(16) (1.8 pmol, target site 620–635) with a MB-mRNA probe (1.8 pmol). (d) Relative CL intensities in the presence of target ODNs (45 fmol, target site 620–637) with a MB-mRNA probe (0.45 pmol).

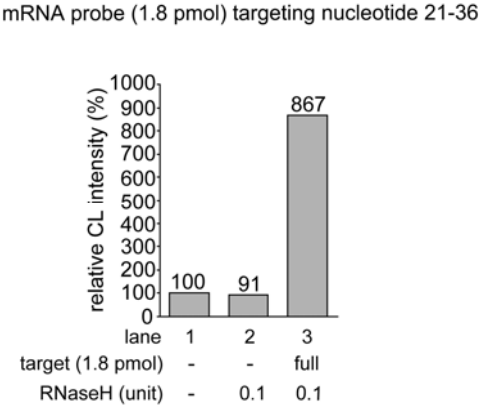


Figure 3. Detection of the nucleotide 21–36 region of the CCR5 sequence using the designed MB-mRNA probe (1.8 pmol) by CL assay.

The sensitivity-enhanced RNase H activity coupled system still retained a low selectivity for single-nucleotide differences in the 18-nt binding site (lanes 5–7 in Figure 2a). Two-base mismatched ODN^{2mis} (T for C and A for T at positions 627 and 628) and one-base mismatched ODN^{1misA} (C to A transversion at position 627) gave no notable enhancement (I_{on}/I_{off} = 127% for lane 3 versus 5 and 170% for lane 3 versus 7, respectively, in Figure 2a). However, nonnegligible CL enhancement was observed for one-base mismatched target ODN^{1misT} (C to T transition at position 627), perhaps due to the resulting stable GT mismatch base pair in the resulting heteroduplex (I_{on}/I_{off} = 424% for lane 3 versus 6 in Figure 2a).

One-nucleotide selectivity was further investigated by using the shorter targets of 16-nt (the nucleotide 620–635 region of the CCR5 gene sequence in Figure 1c). The shorter target length could lead to higher sequence selectivity; however, the target/probe hybridization of 16 nt seems to be too short to invade into the hairpin domain (a total of 13 internal base pairings) and open the MB structure by the presence of only one equivalent of a target with no external support, as revealed by the design of the conventional MB probe. This was the case. As shown in Figure 2c (lanes 2 and 3), the original system that was not RNase H activity coupled could not detect even a fully matched target of 16-nt (ODN$^{full(16)}$). However, in marked contrast, the RNase H activity coupled system allowed the 16-nt target to activate the translation of MB-riboregulator mRNA (I_{on}/I_{off} = 565% for lane 4 versus 5 in Figure 2c) by concomitant catalytic cleavage of the hybridized MB-mRNA probe. Importantly, as expected, the shorter 16-nt targeting MB-mRNA system, coupled with RNase H activity, succeeded in discriminating even the C to T transition at position 627 (ODN$^{1misT(16)}$, lane 5 versus 6 in Figure 2c).

The advantage of the present RNase H activity coupled system is greatly emphasized when we detect a small amount of target DNA.[19,20] The sensitivity or detection limit related to the I_{on}/I_{off} value was enhanced by the RNase H activity coupled MB-mRNA system, wherein the translation-activated mRNA probe could be catalytically amplified above the equivalent stoichiometry of the target through the third catalytic process. While the RNase H-free system failed to detect 45 fmol (18 nM, 2.5 μL) of target ODN^{full} (lane 3 versus 4 in Figure 2d),[21] the RNase H activity coupled system enabled clear detection of the same amount of target with CL enhancement of about 400% (I_{on} = 24.7 × 10^7 cpm (counts per minute), I_{off} = 6.5 × 10^7 cpm) in combination with a sensitive luminometer (Lumat LB 9507; lane 1 versus 2 in Figure 2d). This triple catalytic system allows detection of ≥ 9 fmol (3.6 nM, 2.5 μL) of

target with repetitively reproducible results in the CL enhancement (I_{on}/I_{off} =1.7; data not shown).

Finally, we achieved visible sensing of the target sequence by an unmodified but functional RNA probe. To visualize the sensing signal, we prepared a color-reporting MB-mRNA probe by exchanging the gene encoding luciferase with β-galactosidase (β-gal) (target: nucleotide 620–637 region of the CCR5 gene sequence). The expressed β-gal could convert the substrate *ortho*-nitrophenyl-β-D-galactopyranoside (ONPG) into *ortho*-nitrophenol (ONP), hence leading to a visible color change to light yellow. Transcription/translation of the dsDNA template for the MB-controlled mRNA for β-gal (0.2 pmol) was carried out in the presence of RNase H (0.1 U/μL) and target ODN^{full} (1.8 pmol). Enzymatic conversion of ONPG into ONP was quantitatively evaluated by measuring the absorbance increase at 405 nm (Figure 4a). In a similar manner to that described above, the remarkably high absorbance increase (I_{off} = 0.04 for lane 3, I_{on} = 0.59 for lane 4) was obtained, as expected, only for the RNase H activity coupled translation system; I_{on}/I_{off} reached 1475% (lane 3 versus 4 in Figure 4a). The detection could be readily achieved with the naked eye. The presence of target ODN^{full} was easily distinguished by the visible yellow color of the solution (lane 2 versus 3 in Figure 4b).

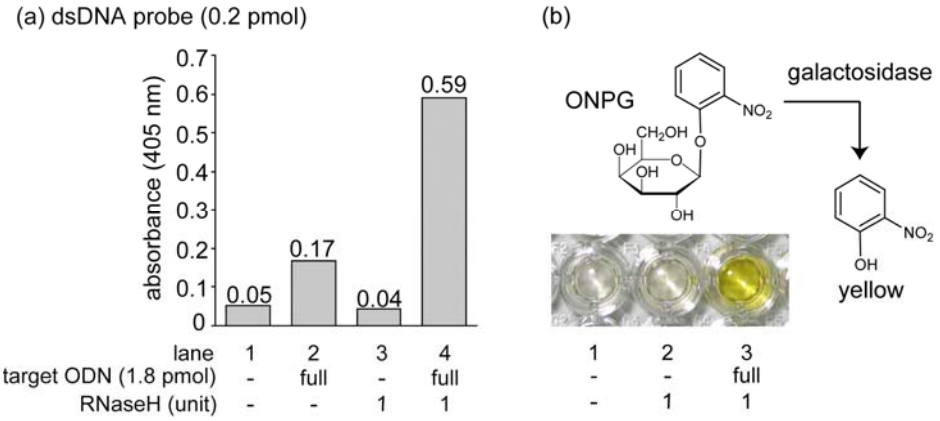


Figure 4. Visible sensing of the CCR5 gene sequence (region 620–637) by a β-gal-encoding MB-mRNA probe. (a) Absorbance (405 nm) of β-gal assay solution with ONPG as a substrate (100 μL) containing an 11-μL aliquot of RNase H (0.1 U/μL) activity coupled MB-mRNA system (0.2 pmol of dsDNA template) translated in the presence of target ODN^{full} (1.8 pmol; target site: 620–637 region of the human CCR5 gene). (b) Reaction scheme and the photographic image of these assay solutions.

Conclusion

In summary, we have demonstrated a catalytic sensing of nucleic acids by using an unmodified RNA probe in an RNase H activity coupled cell-free translation system. The combination with RNase H activity, which induces the third signal-amplifying process of catalytic and irreversible activation of an mRNA probe, achieves an improved sequence selectivity (one-nucleotide selectivity in a target ODN of 16-nt in length) and an enhanced sensitivity ($\geq$ 9 fmol of target). The observed sensitivity is comparable with that of the previously reported isothermal sensing system with an RNase H activity coupled, cleavable, fluorescent resonance energy transfer (FRET) probe,[15c] but should also be compared with much higher sensitivities developed recently.[2f] Nevertheless, the characteristic aspect of the present system that simple/unmodified RNAs or even dsDNAs can be used as probes is something that can not be mimicked readily by other methods.

In particular, the present RNase H activity coupled system also allowed visible sensing of the target sequence under isothermal conditions. It is remarkable that the target sequence can be color-visualized without fluorophore/quencher molecules or inorganic nanoparticles such as quantum dots. What is needed for this color-based sensing is simply a genetically encodable unmodified RNA which works in a prokaryotic translation system. Since, in principle, any type of reporter proteins can be generated, the present method may facilitate high-throughput multiplex screening of genetic sequences (by using fluorescent proteins with different emission properties) or be applicable in electrochemical sensing devices (by using electrochemically sensible redox proteins). Nucleic acid based nanomachines or nanodevices may also be a potential application of this approach.[22]

Experimental Section

Preparation of mRNA probe. The template dsDNA containing the MB domain and the T7 promoter was prepared by two-step PCR amplification using the luciferase gene region of the pBESTluc vector (Promega) or the β-galactosidase gene region of the pSVGal vector (Promega). The first- and the second-PCR products were purified by gel electrophoresis. The MB-mRNA was transcribed using MegaScript (Ambion) and purified by an RNeasy kit (Qiagen).

Primers for template dsDNA of luciferase MB-mRNA targeting nucleotide 620-637: forward primer for first PCR 5'-GGT CTG AAG GTT TAT TAG AAG GAG ATA TAC CAA TGG AAG ACG CCA AAA ACA TA-3', reverse primer for first PCR 5'- TAT TCA TTA CAA TTT GGA CTT TCC GCC -3', forward primer for second PCR 5'-GAA ATT AAT ACG ACT CAC TAT AGG GAG ACC ACA ACG GTT TCC CTC TAT CTC CTG GTC TGA AGG TTT A-3', and reverse primer for second PCR 5'-TAT TCA TTA CAA TTT GGA CTT TCC GCC-3'.

Primers for template dsDNA of luciferase MB-mRNA targeting nucleotide 21-36: forward primer for first PCR 5'-GCA GGA TCC TTC ACT CAG AAG GAG ATA TAC CAA TGG AAG ACG CCA AAA ACA TA-3', reverse primer for first PCR 5'-TAT TCA TTA CAA TTT GGA CTT TCC GCC-3', forward primer for second PCR 5'-GAA ATT AAT ACG ACT CAC TAT AGG GAG ACC ACA ACG GTT TCC CTC TAT CTC CTG CAG GAT CCT TCA CTC AG-3', and reverse primer for second PCR 5'-TAT TCA TTA CAA TTT GGA CTT TCC GCC-3'.

Primers for template dsDNA of β-gal MB-mRNA targeting nucleotide 620-637: forward primer for first PCR 5'-GGT CTG AAG GTT TAT TAG AAG GAG ATA TAC CAA TGA CCA TGA TTA CGG ATT CAC-3', reverse primer for first PCR 5'-GGA TTA GTT ATT CAT TAT TTT TGA CAC CAG ACC AAC-3', forward primer for second PCR 5'-GAA ATT AAT ACG ACT CAC TAT AGG GAG ACC ACA ACG GTT TCC CTC TAT CTC CTG GTC TGA AGG TTT ATT AG-3', and reverse primer for second PCR 5'-GGA TTA GTT ATT CAT TAT TTT TGA CAC CAG ACC AAC-3'.

Translation of mRNA probe or transcription/translation of dsDNA probe in prokaryotic translation system. Translation or transcription/translation was carried out using T7-coupled reconstituted *E. coli* translation system (Pure System, Post Genome Institute CO. Ltd.).[17] mRNA probe (1.8 or 0.45 pmol) or dsDNA template (0.2 pmol), prepared as above, was incubated in 10 μL of the Pure System translation

solution in the presence of target ODN (0-1.8 pmol) and *Tth* RNaseH (1 or 0.1 U, TOYOBO) at 37 °C for 1 h, followed by luciferase or β-galactosidase assay (vide infra).

Luciferase assay. Translation solution (2.5 μL aliquot) was mixed with 100 μL of Luciferase Assay Reagent (100 μL, Promega). Chemiluminescence intensities were observed by Wallac 1420 microplate reader for 96-well plate assay (1.8 pmol of mRNA probe or 0.2 pmol of dsDNA probe) or Lumat LB9507 instrument for single-tube assay (0.45 pmol of mRNA probe). Error bars are standard deviations of three independent measurements.

β-Galactosidase assay. β-Gal assay was performed using β-Gal Enzyme Assay System (Promega) according to manufacturing protocol with the slight modification as follows. Translation solution (11 μL) was mixed with 39 μL of 1× Reporter Lysis Buffer and 50 μL Assay 2× Buffer containing *o*-nitrophenyl-β-D-galactopyranoside (ONPG). The resulting solution was incubated at 37 °C for 60 min and the absorbance at 405 nm was read by Wallac 1420 microplate reader.

Catalytic cleavage of the MB domain of the probe. A MB domain of the probe, having only the 1-92 nucleotide sequence, was prepared as follows. The template dsDNA containing only the MB domain under the control of T7 promoter was prepared by PCR using dsDNA for transcription of luciferase MB-mRNA targeting nucleotide 620-637, prepared above, as a template with a forward primer 5'-GAA ATT AAT ACG ACT CAC TAT AGG-3' and a reverse primer 5'-CCT TTC TTT ATG TTT TTG GCG TC-3'. The MB domain (nucleotide 1-92) was transcribed using T7-MegaShortScript (Ambion) and purified by an RNeasy kit (Qiagen). We checked cleavage of this MB domain (10 pmol/μL) in the presence or absence of RNase H (0.1 U/μL) and target (0.18 pmol/μL). Gel analysis (Figure S1) indicated that the MB domain underwent complete cleavage in 30 min when and only when an RNase H and the target were co-present (lanes 2 and 3).

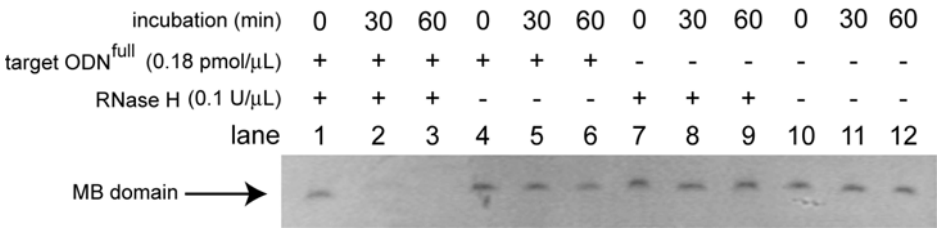


Figure S1. Polyacrylamide gel analysis of the target ODN-induced cleavage of the MB domain by *Tth* RNase H. The MB-domain (10 pmol/μL) was incubated with or without *Tth* RNase H (0.1 U/μL) for 0, 30, or 60 min in the presence (0.18 pmol/μL) or absence of target ODN^{full} in 10 μL of translation buffer (pH 7.3, 50 mM potassium phosphate, 2.5 mM $MgCl_2$, 6.5 mM $Mg(OAc)_2$, 5 mM ammonium acetate). After incubation, 10 μL of loading buffer (1 mM EDTA in formamide) was added to the reaction solution. The resulting mixture was applied and analyzed on 8% polyacrylamide gel containing 7M urea.

Activity of the stem-free reference probe. We prepared a stem-free reference probe (mRNA$^{stem\text{-}free}$) where the antiRBS (RBS docking domain) was changed into a non-complementary sequence (5'-UCCCUUUA-3'). This reference probe (1.8 pmol/10 μL) existed in an open or translation-on state and expressed luciferase constantly in the absence of the target ODN. Under otherwise indentical conditions, a MB-mRNA probe (1.8 pmol/10 μL) expressed luciferase in a [target]-dependent manner (Figure S2) with a yield amounting to ≈70% (target/probe = 0.1), ≈80% (target/probe =0.28), or ≈90% (target/probe = 1.0), respectively, of that of the mRNA$^{stem\text{-}free}$ system which represented the maximal attainable activity of the MB-mRNA probe upon full activation.

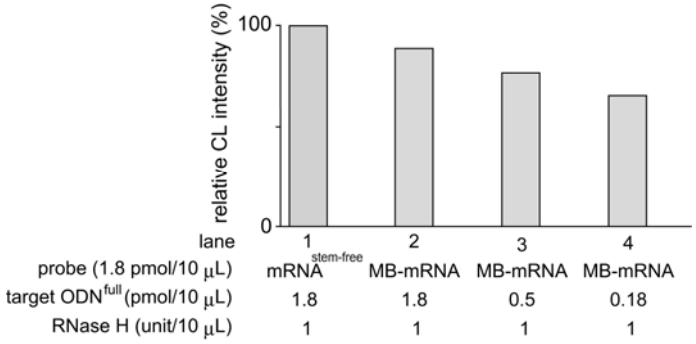


Figure S2. Relative CL intensities arising from a translation system (10 μL) containing 1.8 pmol of mRNA$^{stem\text{-}free}$ or MB-mRNA probe in the presence of RNase H (1 U) and full-match target ODN (1.8 pmol for mRNA$^{stem\text{-}free}$ and 1.8, 0.5, or 0.18 pmol for MB-mRNA probe).

References

(1) Nakatani, K. *ChemBioChem* **2004,** *5*, 1623-1633. and references therein.

(2) For examples of homogeneous gene sensing, see: (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996,** *14*, 303-308. (b) Tyagi, S.; Bratu, P.; Kramer, F. R. *Nat. Biotechnol.* **1998,** *16*, 49-53. (c) Ranasinghe, R. T.; Browon, T.; Browon, L. J. *Chem. Commun.* **2001**, 1480-1481. (d) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002**, *124*, 2096-2097. (e) Brunner, J.; Kraemer, R. *J. Am. Chem. Soc.* **2004,** *126*, 13626-13627. (f) Rosi, N. L.; Mirkin, C. A. *Chem. Rev.* **2005**, *105*, 1547-1562. (g) Yang, C. J.; Lin, H.; Tan, W. *J. Am. Chem. Soc.* **2005**, *127*, 12772-12773.

(3) (a) Stojanovic, M. N.; de Prada, P.; Landry, D. W. *ChemBioChem* **2001,** *2*, 411-415. (b) Stojanovic, M. N.; Stefanovic, D. *Nat. Biotechnol.* **2003,** *21*, 1069-1074. (c) Tian, Y.; Mao, C. *Talanta* **2005**, *67*, 532-537.

(4) (a) Hartig, J. S.; Grune, I.; Najafi-Shoushtari, S. H.; Famulok, M. *J. Am. Chem. Soc.* **2004,** *126*, 722-723. (b) Najafi-Shoushtari, S. H.; Mayer, G.; Famulok, M. *Nucleic Acids Res.* **2004**, *32*, 3212-3219.

(5) Sando, S.; Sasaki, T.; Kanatani, K.; Aoyama, Y. *J. Am. Chem. Soc.* **2003,** *125*, 15720-15721.

(6) Saghatelian, A.; Guckian, K. M.; Thayer, D. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2003,** *125*, 344-345.

(7) Pavlov, V.; Shlyahovsky, B.; Willner, I. *J. Am. Chem. Soc.* **2005**, *127*, 6522-6523.

(8) Abe, H.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 13980-13986.

(9) Cai, J.; Li, X.; Yue, X.; Taylor, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 16324-16325.

(10) Fan, C.; Plaxco, K. W.; Heeger, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9134-9137.

(11) Patolsky, F.; Weizmann, Y.; Katz, E.; Willner, I. *Angew. Chem.* **2003**, 115, 2474-2478; *Angew. Chem. Int. Ed.* **2003**, *42*, 2372-2376.

(12) Sando, S.; Narita, A.; Abe, K.; Aoyama, Y. *J. Am. Chem. Soc.* **2005**, *127*, 5300-5301.

(13) Brantl, S. *Biochim. Biophys. Acta* **2002**, *1575*, 15-25.

(14) Isaacs, F. J.; Dwyer, D. J.; Ding, C.; Pervouchine, D. D.; Cantor, C. R. J.; Collins, J. *Nat. Biotechnol.* **2004,** *22*, 841-847.

(15) For an example of a nucleic acid detection system using RNase H, see: (a) Duck, G.; Alvarado-Urbina, G.; Burdick, B.; Collier, B. *BioTechniques* **1990**, *9*, 142-147. (b) Cloney, L.; Marlowe, C.; Wong, A.; Chow, R.; Bryan, R. *Mol. Cell. Probes*

1999, *13*, 191-197. (c) Harvey, J. J; Lee, S. P.; Chan, E. K.; Kim, J. H.; Hwang, E. S.; Cha, C. Y.; Knutson, J. R.; Han, M. K. *Anal. Biochem.* **2004**, *333*, 246-255.

(16) Gonzalez, E.; Bamshad, M.; Sato, N.; Mummidi, S.; Dhanda, R.; Catano, G.; Cabrera, S.; McBride, M.; Cao, X. -H.; Merrill, G.; O'Connell, P.; Bowden, D. W.; Freedman, B. I.; Anderson, S. A.; Walter, E. A.; Evans, J. S.; Stephan, K. T.; Clark, R. A.; Tyagi, S.; Ahuja, S. S.; Dolan, M. J.; Ahuja, S. K. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96,* 12004-12009.

(17) Shimizu, Y.; Inoue, A.; Tomita, Y.; Suzuki, T.; Yokonaga, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001,** *19*, 751-755.

(18) We independently confirmed that MB-mRNA-based nucleic acid sensing can be carried out also in a normal *E. coli* S30 extract system (RTS HY100, Roche; 10 μL): I_{on}/I_{off} = 400~600% for the MB-mRNA probe (1.8 pmol) with target ODN^{full} (20 pmol).

(19) We carried out RNase H-induced cleavage reactions of a MB domain of the present probe, having the 1-92 nucleotide sequence of the MB-mRNA probe used here. Gel analysis (Figure S1) indicated that cleavage occurred only after incubation in the co-presence of the RNase H and target ODN, indicating that RNase H indeed catalyzed the target-assisted cleavage of the RNA probe with turnover number of ≤ 50 with respect to target ODN.

(20) A-[target]-dependence study (Figure S2) revealed that the RNase H-assisted MB-mRNA probe was ~70% fully activated even in the presence of 0.1 equivalent of target ODN (probe/target = 10), in reference to the activity of a stem-free control probe which existed in an open or translation-on state constantly in the absence of the target. This over-stoichiometric activation of the probe may be most reasonably explained in terms of target-assisted catalytic cleavage of the probe by RNase H.

(21) As described in our previous report (Ref. 12), the original luciferase MB-mRNA probe (4.5 fmol) allowed detection of 50 fmol of target. However, the signal intensity obtained is quite low compared with that of the present RNase H activity coupled system due to the absence of the third signal-amplifying process of catalytic activation of the mRNA probe over the equivalent stoichiometry of the target.

(22) Seeman, N. C. *Trends Biochem. Sci.* **2005**, *30*, 119-125, and references therein

Chapter 4

Sensing of Nucleic Acid Sequences with Unmodified DNA as a Probe: A Strategy to Generate Light-Up Fluorophore/Nucleic Acid Sequence Pair

Abstract

We have designed a strategy to generate a light-up fluorophore/aptamer pair based on a down-modification of a conventional DNA-staining dye to suppress its affinity to the original dsDNA targets, followed by reselection of aptamers that would bind to the modified dye. Following this line, we prepared a micropolarity- sensitive Hoechst derivative possessing two *t*Bu groups with low affinity to the usual AT-rich dsDNA targets. DNA aptamers selected *in vitro* from a random pool worked as triggers to enhance the fluorescence of an otherwise nonfluorescent Hoechst derivative, and the shortened 25-mer sequence showed remarkable enhancement (light-up). The 25-mer sequence was split into binary aptamer probes, thus enabling us to detect a target nucleic acid sequence with a single-nucleotide resolution by use of unmodified DNA as a probe.

Introduction

As the volume of identified genetic information continues to grow, there has recently been much attention on sensing nucleic acid sequences.[1] Among the various types of homogeneous sensing methods so far reported,[2,3] including a representative molecular beacon,[2] several new strategies have enabled the imaging of a target sequence through the use of an unmodified nucleic acid as a probe. For example, Rao and coworkers have reported on a split ribozyme reporter strategy,[4] in which a tetrahymena group I intron was engineered to work as a trans-acting ribozyme, whose splicing activity was switched on by the presence of a target sequence as an allosteric effector. This strategy turned out to be applicable to mRNA sensing, even in mammalian cells.[4] We have also reported on a new strategy involving a molecular beacon-mRNA for sensitive genotyping.[5] This was based on the use of a system of naturally occurring,[6] or engineered,[7] hairpin-shaped RNAs to control the translation frequency, so that the sensing of the nucleic acid sequence, even at single-nucleotide resolution, could be carried out by use of a genetically encodable unmodified RNA as a probe.

Another choice is to use a light-up fluorophore/aptamer pair. Tsien and co-workers first showed that the malachite green (MG) fluorophore dramatically enhances its fluorescence intensity when bound to specific RNA sequences (aptamers).[8] The MG–aptamer is well engineered as a binary probe, and is applicable to the fluorescence detection of nucleic acids.[9] This method is very promising for nucleic acid detection/imaging, owing to its compatibility with in vitro technologies, such as real-time PCR and DNA chips, and also for in-cell applications. In order to make this method a general method, we need to establish a strategy to generate new light-up fluorophore/aptamer pairs, hopefully with multicolor multiaptamer combinations. To this end, Sparano and Koide recently reported on a fluorophore/quencher photoinduced electron transfer (PET) system that exhibits fluorescence enhancement on binding to an RNA aptamer, selected in vitro, that perturbs the PET process.[10] Here, we show a potential approach for generating a bioorthogonal light-up fluorophore/aptamer pair from a single microenvironment-sensitive fluorophore. Our method allows for the conversion of a conventional double-stranded DNA-staining dye into an aptamer-selective light-up fluorophore, which can be applied for nucleic acid sensing with use of an unmodified DNA as a probe.

Results and discussion

Generation of light-up fluorophore/aptamer pairs. Conventional cell-staining fluorescent dyes have already been verified as satisfying the prerequisites as precursors for generating new fluorophore/aptamer pairs: that is, good solubility in aqueous solutions, cell permeability, substantial fluorescence enhancement/shift upon binding to the target, and a high degree of bioorthogonality against other biological components. Our approach is based on an improvement of the known fluorescent dyes used as versatile, and hence less sequence-selective, DNA imagers to a structure-/sequence-specific one by narrowing the range of adaptable target sites by the following steps. The first step is the chemical modification of a conventional DNA-binding fluorophore to suppress its binding to the original DNA target. Because of its original bioorthogonality, the resulting down-modified fluorophore will have no practical target. The second step is reselection, by an in vitro selection method,[11–13] of a new target (aptamer) for the obtained down-modified fluorophore (Figure 1).

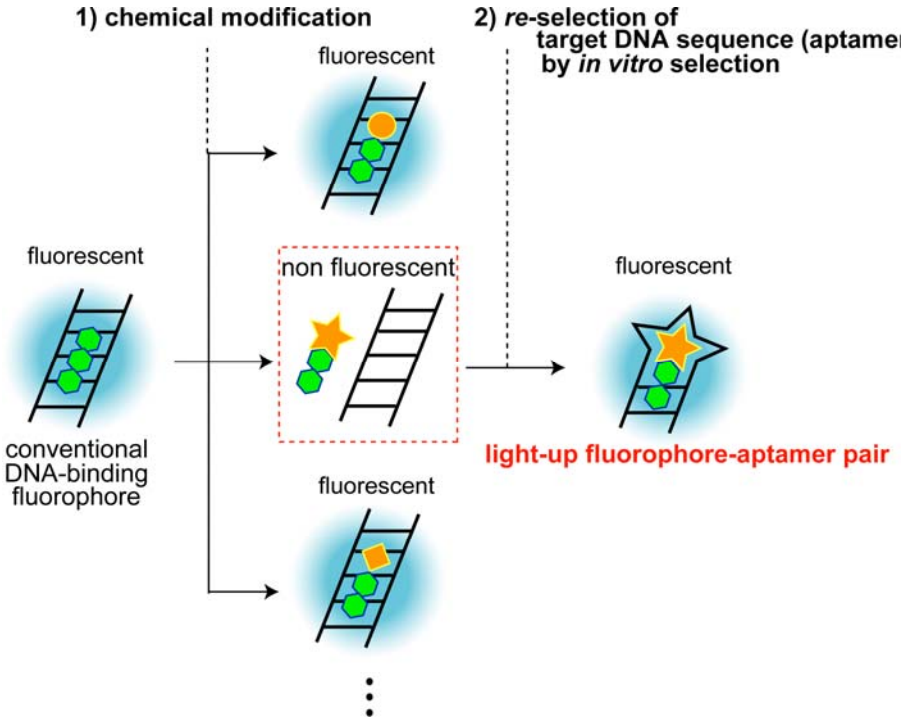
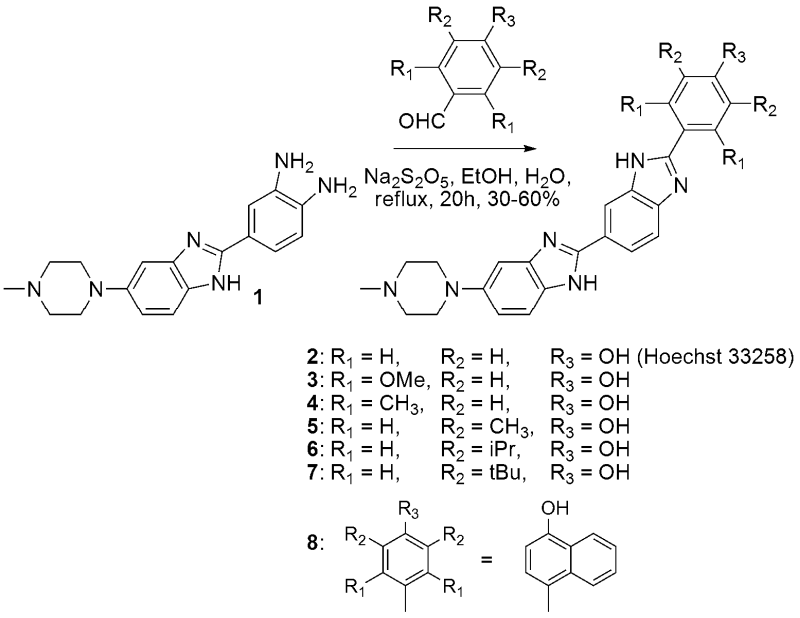


Figure 1. A strategy for converting a conventional DNA-binding fluorophore into an aptamer-selective fluorophore.

Some DNA-binding light-up fluorophores are known to fluoresce more strongly in low dielectric environments. Therefore, we hypothesized that aptamers selected simply in vitro, each having a relatively nonpolar pocket for the binding of a hydrophobic dye, should work as switches to lower the dielectric environment of the bound dye, thus leading to an enhancement of the fluorescence.

As a conventional DNA-imaging probe we focused on the well known Hoechst dye.[14,15] Hoechst dye is a polarity-sensitive fluorophore composed of a bis-benzimidazole chromophore, which preferentially binds to the minor groove of AT-rich dsDNA, and enhances its fluorescence intensity upon binding (dielectric environment of the groove ≈20 D, similar to that of propyl alcohol).[14] Also, the high specificity of Hoechst dye to dsDNAs, among various other components—that is, its high bioorthogonality—has been confirmed in a number of studies on selective genome staining in living cells.[16]

Chemical modification of Hoechst dye. Molecular modeling studies indicate that substitutions on the terminal phenol ring should induce steric repulsion upon binding to DNA. Following this line, we prepared a set of Hoechst derivatives with a variety of substituents on the terminal phenol ring, which might lead to a loss of a binding affinity to AT-rich dsDNAs, the natural targets of the unmodified Hoechst dye. Scheme 1 shows a synthetic scheme for the Hoechst derivatives used in this study. Condensation of *ortho*-diamine **1**[17] with various substituted 4-hydroxybenzaldehydes in the presence of $Na_2S_2O_5$ yielded Hoechst derivatives **2–8** in yields of 30–60%.



Scheme 1. Chemical structures of Hoechst derivatives **2**-**8** and synthetic scheme.

The binding affinities of Hoechst derivatives **2**–**8** to AT- and GC-rich dsDNAs were evaluated by measuring the increases in melting temperature (ΔT_m) due to the presence of these derivatives (Table 1). In the absence of Hoechst derivatives, self-complementary dsDNA(AT) (5'-CGAATTCG-3') and dsDNA(GC) (5'-CGGGCCCG-3') showed T_m = 33.9 ± 0.6 °C and T_m = 51.2 ± 0.2 °C, respectively (double strand concentration = 2.075 μm). The presence of Hoechst 33258 (**2**, 4.15 μm), which is known as an AT-rich dsDNA binder, resulted in a selective increase in the stability of the dsDNA(AT) by ΔT_m = 11.7 °C (ΔT_m = 0.5 °C for dsDNA(GC)). The binding affinities of Hoechst derivatives **5** and **8** to dsDNA(AT) were not weakened compared to that of the parent Hoechst 33258 (ΔT_m > 10 °C). On the other hand, Hoechst derivatives **3**, **4**, **6**, and **7** all showed low binding affinities both to dsDNA(AT) and to dsDNA(GC) (ΔT_m < 2 °C). However, Hoechst derivatives **3** and **4** both

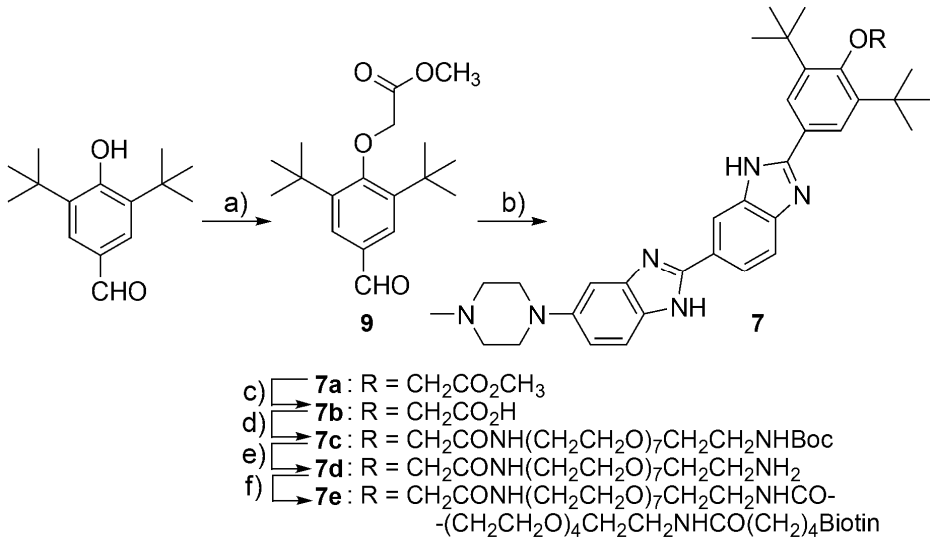
emitted fluorescence in aqueous conditions, even in the absence of DNA, in marked contrast to the nonfluorescent properties of Hoechst derivatives **6** and **7** (data not shown). We selected Hoechst derivative **7** as our candidate for generating a new pair, as it displayed the lowest binding affinity to dsDNA(AT) (ΔT_m = 0.1 °C). There is little doubt that the prevention of effective binding of Hoechst derivative **7** to dsDNA is due to the steric hindrance caused by the two bulky *t*Bu substituents (Table 1).

Table 1. The melting temperatures (T_m) of dsDNA(AT) and dsDNA(GC) in the presence of Hoechst derivatives **2**–**8**.[a]

Hoechst derivatives	dsDNA(AT)			dsDNA(GC)		
	$T_{m(-)}$[b]	$T_{m(+)}$[b]	ΔT_m[c]	$T_{m(-)}$[b]	$T_{m(+)}$[b]	ΔT_m[c]
2	33.9 ± 0.6	45.6	11.7	51.2 ± 0.2	51.7	0.5
3	-	34.3	0.4	-	51.7	0.5
4	-	34.9	1.0	-	51.6	0.4
5	-	44.7	10.8	-	52.3	1.1
6	-	35.5	1.6	-	51.6	0.4
7	-	34.0	0.1	-	51.8	0.6
8	-	45.0	11.1	-	51.3	0.1

[a]The UV melting curve was measured for a self-complementary dsDNA(AT) (5'-CGAATTCG-3') and dsDNA(GC) (5'-CGGGCCCG-3') at a double strand concentration of 2.075 μM in Tris-HCl buffer (pH 7.2, 50 mM) containing NaCl (0.1 M). [b]$T_{m(-)}$ and $T_{m(+)}$ show the melting temperatures in the absence and in the presence, respectively, of Hoechst derivatives (4.15 μM). [c]ΔT_m was calculated from $T_{m(+)}$-$T_{m(-)}$.

In vitro selection of a DNA aptamer targeted towards Hoechst derivative 7. We carried out the in vitro selection of a DNA aptamer that would selectively bind to **7**.[12,13] To immobilize Hoechst derivative **7** on streptavidin-coated magnetic beads, it was conjugated with a biotin moiety through the polyethyleneglycol linker (Scheme 2). 3,5-Di-*tert*-butyl-4-hydroxybenzaldehyde was converted into the ester derivative **9** by coupling with methyl bromoacetate in the presence of K_2CO_3. Condensation of *ortho*-diamine **1** with **9** yielded Hoechst derivative **7a** in a yield of 72%. The terminal methyl ester was then hydrolyzed in aqueous NaOH to give the acid **7b**, which was converted into the amino-PEG derivative **7d** by EDCI coupling with Boc-amino-PEG-amine ($H_2N(CH_2CH_2O)_7CH_2CH_2NHBoc$), followed by removal of the Boc group. The total yield of **7d** in three successive steps was 16%. Finally, **7d** was coupled with NHS-PEO4–biotin ($biotin(CH_2)_4CONHCH_2CH_2(OCH_2CH_2)_4CO_2Su$) to give the terminally biotinylated Hoechst derivative **7e** in a 71% yield.



Scheme 2. Synthesis of Hoechst derivative **7e**. Reagents and conditions: a) $BrCH_2CO_2CH_3$, K_2CO_3, DMF, 10%. b) **1**, $Na_2S_2O_5$, EtOH, H_2O, reflux, 20 h, 72%. c) 1 N NaOH (aq.), 40%. d) 1-ethyl-3-(3'-dimethylaminopropyl)carbodiimide (EDCI), N-hydroxysuccinimide (NHS), Boc-amino-PEG-amine, diisopropylethylamine, DMF. e) 4 N HCl in ethyl acetate, 41% in two steps. f) NHS-PEO_4-biotin, 1× PBS buffer, 71%.

DNA aptamers directed against the immobilized **7e** were selected from a single-stranded DNA (ssDNA) pool with a randomized sequence of 40-nt (total length = 76-nt, approximately 10^{15} different initial ssDNA molecules). The ssDNA pool was mixed with **7e** immobilized on streptavidin-coated magnetic particles in a binding buffer (1 × PBS [10 mM phosphate buffer, pH 7.4, containing NaCl (138 mM) and KCl (2.7 mM)] containing $MgCl_2$ (2.5 mM) and Tween 20 (0.05 %)), and washed several times with the binding buffer. The bound ssDNAs were detached from the magnetic particles by use of an elution buffer [Tris-HCl buffer (pH 8.0, 40 mM, containing urea (3.5M), EDTA (10 mM), and Tween 20 (0.02%)] at 80 °C. The eluted ssDNAs were purified and then subjected to PCR amplification. The recovered ssDNA pool was used for the next round of selection. After the ninth round of the selection, 40 clones were isolated from the pool and sequenced. Thirty of the 40 isolated clones could be classified into three families (Classes I-1, I-2, II, and III in Figure 2A). An estimation of the binding affinity by surface plasmon resonance analysis revealed that all three families were capable of binding to immobilized Hoechst derivative **7e** with dissociation constants of ≤ 10 μM (data not shown). We then measured the enhancements in fluorescence. The fluorescence intensity of the otherwise practically nonfluorescent **7d** (200 nM, excitation at 345 nm and emission at 460 nm) was dramatically increased on addition of all four selected aptamers (200 nM) in the range I_{on}/I_{off} = 57–178 (right-hand side in Figure 2A); this indicated that these aptamers, selected *in vitro* from a random pool, caused an environmental polarity change of the bound Hoechst derivative, as we had initially devised. Aptamer class I showed the highest I_{on}/I_{off} ratio among the families. The most abundant aptamer of class I-1, hereafter denoted aptamer I-1, was chosen for further investigation.

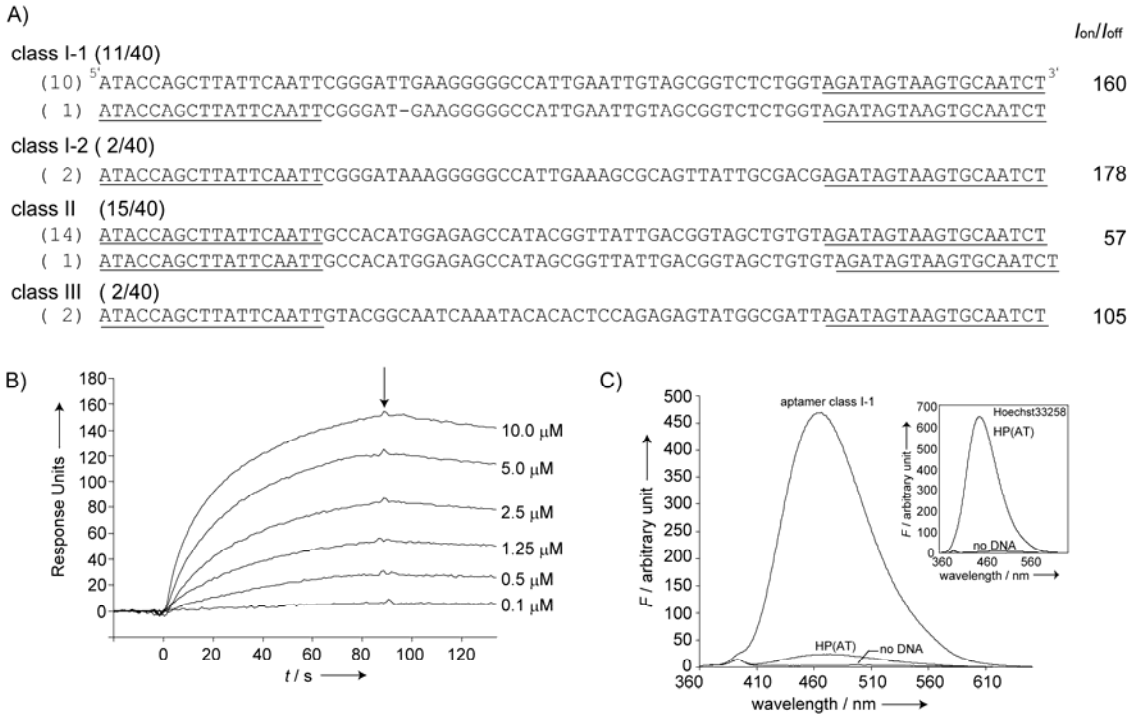


Figure 2. A) The sequences of aptamer classes I–III. The numbers in parentheses indicate the sums of the individual sequences among 40 clones. The primer sites are underlined. The right-hand side shows the enhancement in fluorescence (I_{on}/I_{off} = 460 nm, excitation = 345 nm) of the Hoechst derivative **7d** (200 nM) in the presence of each aptamer (200 nM strand concentration). B) Display of the SPR sensorgrams showing the affinity of Class I-1 aptamer for Hoechst derivative **7e**. Various concentrations of aptamer class I-1 (0.1–10 μm) were injected (total = 15 μL, t = 90 s) at a flow rate of 10 μL/min over a sensor chip SA pre-coated with Hoechst derivative **7e**. The arrow indicates the end of the injection. C) Fluorescence spectra of Hoechst derivative **7d** (200 nM) in the absence and in the presence of Class I-1 aptamer and HP(AT) (200 nM strand concentration). The inset shows the fluorescence spectra of Hoechst 33258 (200 nM) in the absence and in the presence of HP(AT) (200 nM). The fluorescence intensity was measured in the binding buffer [1× PBS containing $MgCl_2$ (2.5 μm)] at 20 °C.

A New Light-Up Fluorophore-DNA aptamer pair. The binding affinity of aptamer class I-1 was assessed by SPR analysis. Biotin-labeled Hoechst **7e** was immobilized on a streptavidin-coated SPR sensor chip. Injection of various concentrations of class I-1 aptamer (0.1–10 μM) over the chip exhibited distinct association and sunsequent dissociation curves (Figure 2B). The binding data were fitted to a one-to-one binding equilibrium, and the dissociation constant (K_d) of aptamer I-1 with the immobilized Hoechst derivatives **7e** was determined to be K_d = 878 nM.

Selective binding was further investigated by measuring the enhancement in fluorescence.Hoechst derivative **7d** (200 nM) showed a 160-fold enhancement in fluorescence in the presence of aptamer I-1 (200 nM; Figure 2 C). In marked contrast, the enhancement in fluorescence was greatly, though not completely, suppressed in the presence of HP(AT) (Figure 2C), a hairpin-structured DNA (5'-CGC GAA TTC GCG TTT TCG CGA ATT CGC G-3' possessing a 5'-AATT-3'/3'-TTAA-5' moiety),[18] which caused a remarkable enhancement in fluorescence in the original (unmodified) Hoechst 33258 (**2** in Scheme 1; inset of Figure 2 C). Thus, the original Hoechst dye (**2** in Scheme 1), which was specific to AT-rich dsDNAs, was rendered aptamer-selective on modification with bulky *t*Bu groups.

We then proceeded to shorten aptamer class I-1. Through measurement of the enhancement in fluorescence for various lengths of truncated aptamers, the 5'-side of aptamer class I-1 was found to be the domain that was responsible for the binding. Systematic investigations of the class I-1 sequence together with other class I family sequences allowed us to minimize the aptamer sequence from a 76-mer to a 25-mer (I-1-mini in Figure 3). This aptamer possesses self-complementary sequences at its ends, and is thus expected to form a hairpin structure. The 25-mer truncated aptamer enhanced the fluorescence of Hoechst derivative **7d** (fluorescence quantum yield (Φ_F) = 0.371),[19] with I_{on}/I_{off} = 191.

To shed more light on the aptamer sequence, six additional aptamers (loop-mod1–2 and stem-mod1–4) were prepared and subjected to fluorescence enhancement measurements. The aptamers loop-mod1 and -2 possess simplified T_7 and T_{17} loop domains and enhance the fluorescence of **7d** far less effectively, with $(I_{on}/I_{off})_{relative}$ = 0.08 and 0.01, respectively, than I- 1-mini, where $(I_{on}/I_{off})_{relative}$ was set at 1.0. Thus, the loop region is important for the enhancement of fluorescence in **7d**.

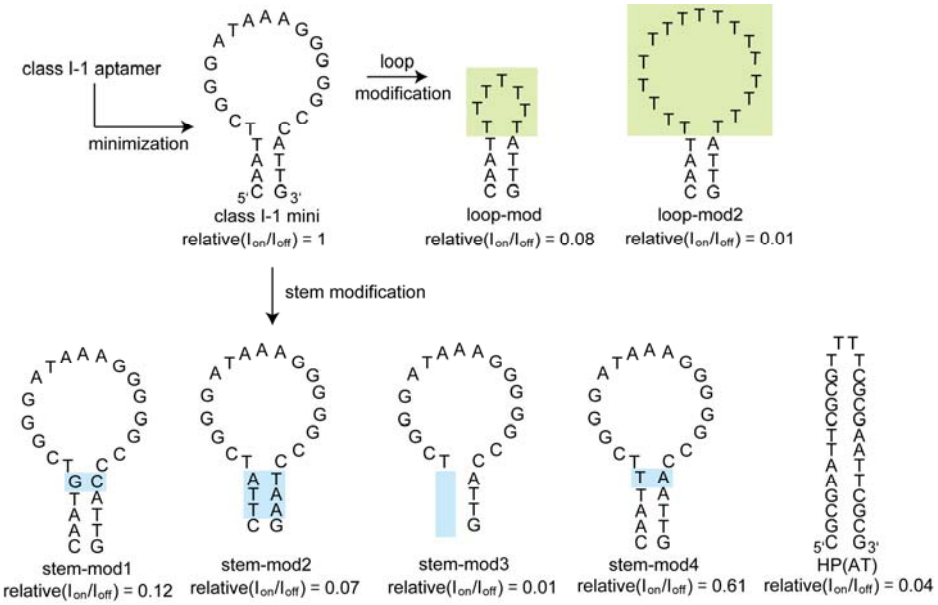


Figure 3. Aptamer class I-1-mini and modified aptamers. The relative enhancement in fluorescence relative to that of the Hoechst derivative **7d**/aptamer class I-1-mini pair [relative (I_{on}/I_{off})] is shown at the bottom of each structure. The fluorescence intensity (excitation = 345 nm and emission = 460 nm) of Hoechst derivative **7d** (200 nM)/aptamer (200 nM) was measured in the binding buffer [1× PBS containing $MgCl_2$ (2.5 mM)] at 20 °C.

The stem region is also important. Exchange of the three complementary bases (stem-mod2), deletion of a stem-forming sequence (stem-mod3), and the addition of a GC pair at the junction of the stem and loop domains (stem-mod1) all resulted in decreases in fluorescence intensity, with $(I_{on}/I_{off})_{relative}$ = 0.01–0.12, although stem-mod4, with the addition of a TA pair at the junction, retained more than half of the fluorescence intensity ($(I_{on}/I_{off})_{relative}$ = 0.61). These results suggest that the key for good enhancement of fluorescence is a combination of an AT-rich stem to provide the primary binding site for the aromatic region of dye **7d**, and a loop to accommodate the bulky *t*Bu groups.

Binary I-1-mini aptamer probes for sensing nucleic acid sequences. With the new light-up fluorophore/DNA-aptamer pair to hand, we applied the pair in sensing nucleic acid sequences. We engineered aptamer class I-1-mini into a nucleic acid sensing probe by dividing it into binary aptamers (5'-BinApt and 3'-BinApt, Figure 4A), according to Kolpashchikov's design principle.[9] The 5'-BinApt and 3'-BinApt have 9- and 20-nt target-binding sequences (black in Figure 4A), respectively. The binary probes were designed to restore the core structure of the I-1-mini aptamer (red in Figure 4A) only in the presence of a target (blue in Figure 4A), so that the presence or absence of the target sequence could be detected by measuring the enenhancement in fluorescence of Hoechst derivative **7d** that was also present. Figure 4B shows the fluorescence spectra of Hoechst derivative **7d** (200 nM) with binary aptamers 5'-BinApt and 3'-BinApt (200 nM each). In the presence of a fully matched target ODN (200 nM, X = G), Hoechst derivative **7d** emitted 143 times the fluorescence at 460 nm of Hoechst derivative **7d** alone (lane 1 versus 3 in Figure 4C). Interestingly, the enhancement in fluorescence induced by the binary aptamers in the absence of the target was only twofold (lower spectrum in Figure 4B, and lane 1 versus 2 in Figure 4C). The fully matched target can therefore be easily imaged by its fluorescence (inset of Figure 4C). These results clearly indicate that the presence of a complementary target is essential for the binary aptamers to construct the hairpin structure responsible for the light-up of Hoechst derivative **7d**. The full loop regions in the binary aptamers are also essential for sensing, since the simplified binary aptamers 5'-BinApt2 and 3'-BinApt2, whose loop-forming sequences had been deleted, were barely capable of achieving a light-up, with only fivefold enhancement (lane 1 versus 5 in Figure 4C). We also tested the selectivity of the binary aptamers. When single-mismatched ODN (X = C) was used as the target, the fluorescence enhancement was only ninefold (lane 1 versus 4 in Figure 4 C), indicating that these binary DNA aptamer probes were capable of discrimination of a single nucleotide.

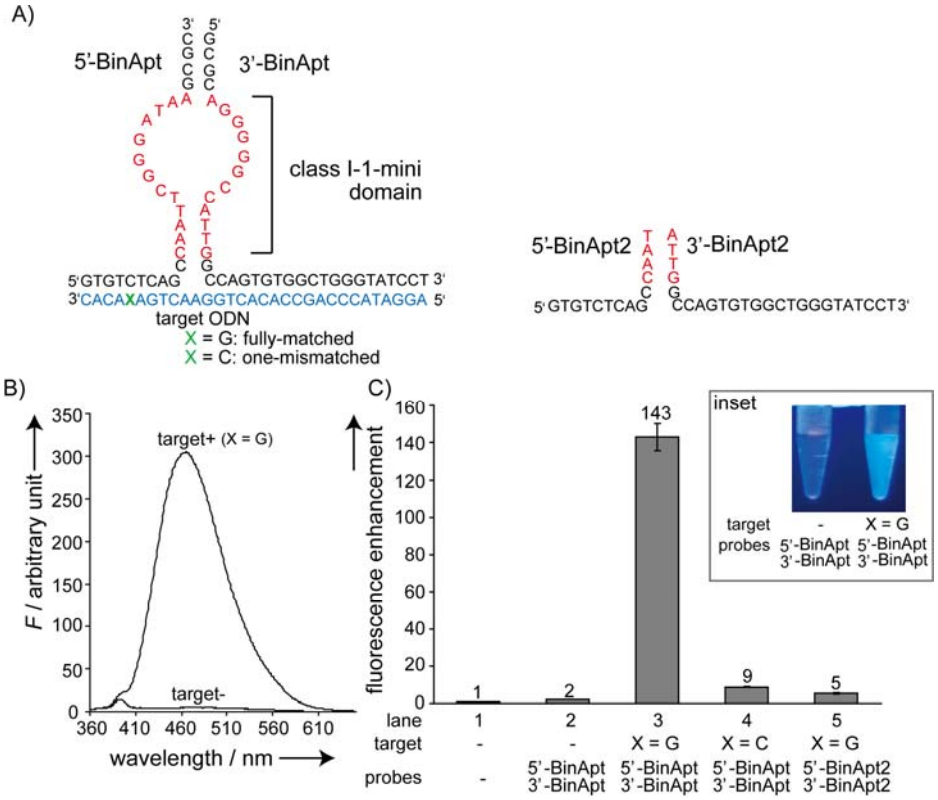


Figure 4. A) Binary DNA aptamer probes for nucleic acid sensing, based on Kolpashchikov's binary RNA aptamers.[9] Domains for the Class I-1-mini-aptamer and target ODN are shown in red and blue, respectively. Fully matched and single-mismatched ODNs possess G and C bases at nucleotide X (green), respectively. B) Fluorescence spectra of Hoechst derivative **7d** (200 nM) with binary probes 5'-BinApt and 3'-BinApt (200 nM each) in the absence and presence of a fully matched target ODN (200 nM). The fluorescence spectra were measured in 1× PBS buffer containing $MgCl_2$ (2.5 mM) at 35 °C. C) The enhancement in fluorescence (excitation = 345 nm, emission = 460 nm at 35 °C) of Hoechst derivative **7d** (200 nM) in the presence of fully matched (200 nM, X = G, Lanes 3 and 5) and single-base mismatched target ODN (200 nM, X = C, lane 4) with binary probes (200 nM each, 5'-BinApt and 3'-BinApt for Lanes 2-4, 5'-BinApt2 and 3'-BinApt2 for lane 5). Error bars are standard deviations of three independent measurements. The inset shows the fluorescence image (excited by a 366 nm transilluminator) of a Hoechst derivative **7d** solution (2 μM) with binary probes (2 μM each, 5'-BinApt and 3'-BinApt) in the absence (left-hand side) and in the presence (right-hand side) of fully matched target ODN (2 μM, X = G).

Conclusion

In summary, we have developed a strategy to generate new light-up fluorophore/aptamer pairs. The relevant observations are as follows: 1) Hoechst dye used as a conventional dsDNA-staining dye loses its affinity to the intrinsic target (dsDNA) on simple modification with bulky alkyl groups. 2) Nevertheless, we can select particular aptamers that will still bind to the modified Hoechst derivatives (**7d**), with the light-up of the latter being achieved by a simple *in vitro* selection method. 3) The aptamer can be shortened to a 25-mer without damaging the binding and light-up capability. 4) The shortened aptamer (I-1-mini) was split into binary DNA probes, thus allowing us to sense a target ODN sequence at single-nucleotide resolution.

The significance of these findings is twofold. Firstly there is the utility of unmodified DNA, free from covalent linkages to fluorophore and quencher dyes, acting as a probe. Molecular beacons, while widely used, can generate false signals owing to unexpected cleavage or conformational changes in the probe. This is not the case in this system, for which unexpected cleavage, if it happens at all, would not lead to a permanent signal being switched on, since only the reconstruction of the hairpin structure is responsible for the light-up of the dye. Secondly, there is the significance of the versatility of this fluorophore/aptamer pair-generating strategy. We have demonstrated that a microenvironment-sensitive fluorophore can be converted into an aptamer-selective fluorophore by using SELEX technology. Though the orthogonality of the pair should be further improved, it is interesting that simple *in vitro* selection was able to generate active aptamers that function as triggers for lighting up otherwise nonfluorescent fluorophores. Owing to its simplicity, our strategy could be applied to other living-cell-staining fluorophores, if necessary, taking into account factors other than polarity such as molecular conformation and/or rigidity of the bound dye, which might also come into play. Unmodified RNA,[8–10,12,13] in addition to DNA, could, in principle, also be used as a binding aptamer; this would allow the possibility of a systematic generation of living-cell-compatible multicolor light-up fluorophore/aptamer pairs. Further work along this line, including the improvement of the strategy, is underway in our laboratory.

Experimental Section

General. The reagents and solvents used were purchased from commercial suppliers and used without any further purification. The ^{1}H NMR spectra were taken with a JEOL (Japan) JNM-A500 (500 MHz) NMR. The coupling constants (*J* values) are reported in Hertz. The FAB mass spectra were recorded on a JEOL (Japan) JMS HX110A spectrometer. The DNA oligomers used were purchased from Gene Design, Inc. (Japan). The fluorescence spectra were obtained with a Shimadzu (Japan) RF-5300PC spectrofluorometer. The thermal denature profiles were recorded with a Shimadzu (Japan) UV1650PC spectrophotometer equipped with a temperature controller. The SPR measurements were performed with a BIAcore X system (GE Healthcare).

General procedure for the synthesis of the Hoechst derivatives. Freshly prepared *ortho*-diamine **1**[17] (93 mg, 0.29 mmol) was dissolved in EtOH (12 mL). The 4-hydroxybenzaldehyde derivative (2.5 equivalents) and sodium pyrosulfate (46 mg, 0.24 mmol in 270 μL H2O) were added to this solution. The resulting solution was heated at reflux for a period of 20 h, and the solvent was then removed by evaporation. The crude products were recovered by column chromatography (silica gel), and further purified by GPC (MeOH) to give the pure Hoechst derivative in a yield of 30–60%.

4-[5-(4-Methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo[*d*]imidazol-2'-yl]phenol (Hoechst derivative 2). Hoechst derivative **2** was synthesized by the general procedure described above with use of 4-hydroxybenzaldehyde. Yield = 55%, ^{1}H NMR (500 MHz, CD_3OD): δ = 6.92–8.22 (10H), 3.19 (t, *J* = 5.0 Hz, 4H), 2.65 (t, *J* = 5.0 Hz, 4H), 2.35 ppm (s, 3H); HRMS (FAB): *m*/*z* calcd for $C_{25}H_{24}N_6O_1$ $[M]^+$: 24.2012; found: 424.2010.

3,5-Dimethoxy-4-[5-(4-methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo[*d*]imidazol-2'-yl]phenol (Hoechst derivative 3). Hoechst derivative **3** was synthesized by the general procedure described above with use of 2,6-dimethoxy-4-hydroxybenzaldehyde. Yield = 38%; ^{1}H NMR (500 MHz, CD_3OD): δ = 7.01–8.25 (6H), 6.20 (s, 2H), 3.76 (s, 6H), 3.20 (t, *J* = 4.5 Hz, 4H), 2.65 (t, *J* = 4.5 Hz, 4H), 2.35 ppm (s, 3H); HRMS (FAB): *m*/*z* calcd for $C_{27}H_{28}N_6O_3$ $[M]^+$: 484.2223; found: 484.2228.

3,5-Dimethyl-4-[5-(4-methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo-[*d*]imidazol-2'-yl]

phenol (Hoechst Derivative 4). Hoechst derivative **4** was synthesized by the general procedure described above with use of 4-hydroxy-2,6-dimethylbenzaldehyde. Yield = 39%; ^{1}H NMR (500 MHz, CD_3OD): δ = 7.03–8.30 (6H), 6.61 (s, 2H), 3.22 (t, *J* = 4.5 Hz, 4H), 2.67 (t, *J* = 4.5 Hz, 4H), 2.37 (s, 3H), 2.10 ppm (s, 6H); HRMS (FAB): *m*/*z* calcd for $C_{27}H_{29}N_6O_1$ $[M+H]^+$: 453.2403; found: 453.2399.

2,6-Dimethyl-4-[5-(4-methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo-[*d*]imidazol-2'-yl] phenol (Hoechst derivative 5). Hoechst derivative **5** was synthesized by the general procedure described above with use of 4-hydroxy-3,5-dimethylbenzaldehyde. Yield = 39%; ^{1}H NMR (500 MHz, CD_3OD): δ = 7.00–8.20 (8H), 3.22 (4H), 2.73 (4H), 2.42 (s, 3H), 2.30 ppm (s, 6H); HRMS (FAB) *m*/*z* calcd for $C_{27}H_{29}N_6O_1$ $[M+H]^+$: 453.2403; found: 453.2404.

2,6-Diisopropyl-4-(5-(4-methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo[*d*]imidazol-2'-yl)phenol (Hoechst derivative 6). Hoechst derivative **6** was synthesized by the general procedure described above with use of 3,5-diisopropyl-4-hydroxybenzaldehyde. Yield = 58%; ^{1}H NMR (500 MHz, CD_3OD): δ = 6.92–8.20 (8H), 3.16 (2H), 3.11 (4H), 2.55 (4H), 2.29 (s, 3H), 1.23 ppm (12H); HRMS (FAB): *m*/*z* calcd for $C_{31}H_{37}N_6O_1$ $[M+H]^+$: 509.3029; found: 509.3028.

2,6-Di-tert-butyl-4-[5-(4-methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo[*d*]imidazol-2'-yl]phenol (Hoechst derivative 7). Hoechst derivative **7** was synthesized by the general procedure described above with use of 3,5-di-tert-butyl-4-hydroxybenzaldehyde. Yield = 53%; ^{1}H NMR (500 MHz, CD_3OD): δ = 7.00–8.25 (8H), 3.19 (4H), 2.68 (4H), 2.38 (s, 3H), 1.50 ppm (s, 18H); HRMS (FAB): *m*/*z* calcd for $C_{33}H_{40}N_6O_1$ $[M]^+$: 536.3264; found: 536.3265.

4-[5-(4-Methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bibenzo[*d*]imidazol-2'-yl]naphthalen-1-ol (Hoechst derivative 8). Hoechst derivative **8** was synthesized by the general procedure described above with use of 4-hydroxynaphthaldehyde. Yield = 60%; ^{1}H NMR (500 MHz, CD_3OD): δ = 6.96–8.52 (12H), 3.23 (t, *J* = 4.5 Hz, 4H), 2.69 (t, *J* = 4.5 Hz, 4H), 2.38 ppm (s, 3H); HRMS (FAB): *m*/*z* calcd for $C_{29}H_{27}N_6O_1$ $[M+H]^+$: 475.2246; found: 475.2231.

Methyl 2-(2,6-di-tert-butyl-4-formylphenoxy)acetate (9). Methyl bromoacetate (1.0 mL, 10.6 mmol) and K_2CO_3 (2.19 g, 15.9 mmol) were added to a solution of

3,5-di-*tert*-butyl-4-hydroxybenzaldehyde (1.5 g, 6.41 mmol) in dry DMF (20 mL). The solution was heated at reflux overnight under N_2, and the solvent was then removed by evaporation. The resulting mixture was dissolved in ethyl acetate, washed with H_2O and brine, and evaporated to dryness. The product was roughly purified by column chromatography (silica gel) to give a 3:1 mixture of compound **9** and 3,5-di-*tert*-butyl-4-hydroxybenzaldehyde in a total yield of approximately 10%. The mixture was used for the ensuing coupling step without further purification. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 9.90 (s, 1H), 7.78 (s, 2H), 4.36 (s, 2H), 3.83 (s, 3H), 1.43 ppm (s, 18H) for **9**.

Hoechst derivative 7b. Hoechst derivative **7a** was synthesized by the general procedure described above with use of methyl 2-(2,6-di-*tert*-butyl-4-formylphenoxy)acetate (**9** as a 3:1 mixture with 3,5-di-*tert*-butyl-4-hydroxybenzaldehyde). Yield = 72%. ^{1}H NMR (500 MHz, CD_3OD): δ = 6.82–8.20 (8H), 4.35 (s, 2H), 3.74 (s, 3H), 3.15 (4H), 2.60 (4H), 2.30 (s, 3H), 1.42 ppm (s, 18H). Hoechst derivative **7a** (20 mg) was dissolved in NaOH (1 N, 660 μL), and the mixture was stirred for a period of 3 h at room temperature and then for an additional 1 h at 50 °C. The mixture was then neutralized with aqueous HCl (1 N, ~600 μL). TEAA buffer (100 mM, pH 7.0, 1200 μL) was added to this solution, the resulting solution was centrifuged for 5 min at 13000 rpm, and the supernatant was removed by pipetting. TEAA buffer (100 mM, pH 7.0, 1200 μL) was added to the resulting solid, which was vortexed, and centrifuged for 5 min at 13000 rpm. After removal of the supernatant, the resulting wet product was lyophilized to dryness to give the TEAA salt of Hoechst derivative **7b** in a yield of approximately 40%. ^{1}H NMR (500 MHz, CD_3OD): δ = 7.07–8.30 (8H), 4.25 (s, 2H), 3.42 (4H), 3.32 (4H), 2.86 (s, 3H), 1.53 ppm (s, 18H); HRMS (FAB): *m*/*z* calcd for $C_{35}H_{43}N_6O_3$ $[M+H]^+$: 595.3397; found: 595.3391.

Hoechst derivative 7d. 1-Ethyl-3-(3'-dimethylaminopropyl)carbodiimide (EDCI, 6.4 mg, 0.033 mmol), N-hydroxysuccinimide (NHS, 3.9 mg, 0.034 mmol), and diisopropylethylamine (4.4 mg, 0.034 mmol) were added to a solution of Hoechst derivative 7b·TEAA salt (2 mg) in dry DMF (800 μL). The resulting mixture was stirred under N_2 at room temperature. After the mixture had been stirred for a period of 1 h, Boc-amino-PEG-amine (15 μL, Polypure, Norway, $H_2N(CH_2CH_2O)_7CH_2CH_2NHBoc$) was added, and the mixture was stirred for 10 h at room temperature. The crude product was purified by HPLC and lyophilized to give

Hoechst derivative **7c**. Compound **7c** was then dissolved in CH_3CN (400 μL), and mixed with HCl (4 N) in ethyl acetate (500 μL). The mixture was stored for 1 h at room temperature and the organic solvent was removed. The crude product was purified by HPLC and lyophilized to give **7d** in a yield of 41%. The concentration of Hoechst derivative **7d**, when needed, was determined by use of the known extinction coefficient of Hoechst 33342. MALDI-TOF mass: *m*/*z* calcd = 946.21 for $[M+H]^+$; found: 946.15.

Hoechst derivative 7e. PBS buffer (10×, 60 μL) and NHS-PEO_4-biotin (2 mg, Pierce, USA, biotin$(CH_2)_4CONHCH_2CH_2(OCH_2CH_2)_4$-$CO_2Su$) in H_2O (200 μL) were added to a solution of Hoechst derivative **7d** (200 μL, 2.83 mM) in H_2O. The mixture was stored for a period of 1 h at room temperature. The crude product was purified by HPLC and lyophilized to dryness to give Hoechst derivative **7e**. The concentration of Hoechst derivative **7e** was determined by the same method as described for **7d**. Yield = 71%; MALDI-TOF mass: *m*/*z* calc. for $[M+H]^+$: 1419.79; found: 1419.75.

Preparation of Hoechst derivative 7e-immobilized magnetic beads. Streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen, USA, 1 mg) were transferred into a 1.5 mL tube, and washed once with B&W buffer [500 μL, Tris-HCl (5 mM), pH 7.5, containing EDTA (0.5 mM) and NaCl (1 M)] and then twice with 1× PBS buffer [500 μL, phosphate buffer (pH 7.4, 10 mM), containing NaCl (138 mM) and KCl (2.7 mM)]. The washed beads were resuspended in 1× PBS buffer (100 μL) containing Hoechst derivative **7e** (500 pmol), and the resulting mixture was stored at room temperature with mild shaking. After incubation for a period of 30 min, the beads were washed three times with PBST [500 μL, 1× PBS buffer containing Tween 20 (0.05%)], resuspended in 1× PBS (500 μL), and stored at 4 °C .

***In vitro* selection.** *In vitro* selection was carried out by the protocols of Strehlitz[20] and Doyle[21] with slight modifications. A ssDNA pool of 5'-ATA CCA GCT TAT TCA ATT-$(N)_{40}$-AGA TAG TAA GTG CAA TCT-3' containing 40 randomized nucleotides (3 nmol for the initial round of selection, ~100 pmol for the second and subsequent rounds) in binding buffer (500 μL, 1× PBS [10 mM phosphate buffer, pH 7.4, containing NaCl (138 mM) and KCl (2.7 mM)] containing $MgCl_2$ (2.5 mM) and Tween 20 (0.05%)) was denatured at 90 °C for a period of 10 min. It was then immediately cooled to 4 °C for 15 min, and left at room temperature for a further 8 min.

Negative selection was performed before each positive selection round. In the negative selection, the pre-annealed ssDNA library was applied to streptavidin-coated magnetic beads [1.0 mg, prewashed twice with binding buffer (500 μL)] and incubated for 15 min at 37 °C with mild rotation. The recovered DNA, which was not bound to the streptavidin-coated magnetic beads, was then subjected to positive selection. The recovered DNA solution was incubated with Hoechst derivative **7e**-immobilized magnetic beads [0.1 mg, prewashed twice with binding buffer (500 μL)] for 30 min at 37 °C with mild rotation. After the beads had been washed three times with binding buffer (1000 μL), the bound DNA was detached from the magnetic beads with elution buffer [200 μL, Tris-HCl (40 mM) containing Tween 20 (0.02%), EDTA (10 mM), and urea (3.5 M), pH 8.0] by incubation for 7 min at 80 °C. The elution step was repeated twice. The recovered DNA was purified by filtration through a Microcon YM-10 column (Millipore, USA), and was then amplified by PCR [5 min at 94 °C, 15 cycles of (1 min at 94 °C, 1 min at 47 °C, 1 min at 72 °C), and 10 min at 72 °C] with use of the 5'-biotin-labeled primer 5'-biotin-A_{20}-HEG-AGA TTG CAC TTA CTA TCT-3' and the 5'-fluorescein-labeled primer 5'-FAM-ATA CCA GCT TAT TCA ATT-3'. The PCR solution contained each primer (1.25 μM) in a 1× Taq PCR Master Mix (Qiagen, USA). The PCR solution (180 μL) was mixed with NaCl (5 M, 46 μL). This mixture was applied to the streptavidin-coated magnetic beads (1 mg, prewashed twice with 1000 μL of B&W buffer), incubated for 10 min at room temperature, and washed three times with PBST (0.05%, 1000 μL). The washed beads were resuspended in aqueous NaOH (100 mM, 100 μL), and incubated for 5 min at room temperature. The supernatant was transferred to a new 1.5 mL tube containing monobasic potassium phosphatase buffer (100 mM, 100 μL). PBST (0.05%, 800 μL) was added to this solution, to give the 5'-fluorescein-labeled ssDNA pool. All the recovered ssDNA solutions were combined in a single tube and stored at -20 °C. Fluorescein labeling was used to determine the amount of ssDNA present. After the selection, the enriched ssDNA was converted into dsDNA by PCR amplification, and cloned into a pDrive cloning vector (Qiagen, USA), and sequenced by the usual technique.

Fluorescence measurements. Fluorescence spectra were obtained at 20 °C. The Hoechst derivative (200 nM) was dissolved in binding buffer [1× PBS containing Mg^{2+} (2.5 mM)] in the presence (200 nM), or absence of ODNs. The solutions were excited at 345 nm, and the emission was monitored in the 360–650 nm wavelength range. The enhancement in fluorescence (I_{on}/I_{off}) was determined by comparing the

intensity of the fluorescence emissions at 460 nm. Nucleic acid sensing experiments using binary probes were carried out at 35 °C.

***T*$_m$ measurements.** Self-complementary dsDNA(AT) (5'-CGAATTCG-3') or dsDNA(GC) (5'-CGGGCCCG-3') (2.075 μM double strand concentration, pre-annealed) was dissolved in Tris-HCl buffer (pH 7.2, 50 mM) containing NaCl (100 mM) in the presence (4.15 μM) or absence of a Hoechst derivative. The solution was degassed for a period of 10 min. The thermal denature profiles were obtained by monitoring the absorbance of the sample at 260 nm in the range of 15–95 °C with use of a heating ratio of 1 °C/min. The melting temperature (T_m) was determined from plots of the absorbance versus temperature, and defined as the midpoint of the transition.

Surface plasmon resonance analysis. SPR measurements were performed with a BIAcore X system (GE Healthcare). The Hoechst derivative **7e** was immobilized on a sensor chip SA (streptavidin) in a continuous flow of HBS-N buffer [pH 7.4, HEPES (10 mM) containing NaCl (150 mM)] at a flow rate of 10 μL/min. A solution (70 μL) of Hoechst derivative **7e** (5.0 μM) was injected onto the streptavidin-coated gold surface of the chip, and nonspecifically bound materials were washed off with NaOH (100 mM). Binding experiments with various concentrations of aptamer class I-1 (15 μL) were performed at 25 °C in a continuous flow of running buffer [1× PBS containing Mg^{2+} (2.5 mM)] at a flow rate of 10 μL/min. Kinetic analyses were performed with the BIAevaluation 3.1 software (GE Healthcare).

References

(1) Nakatani, K. *ChemBioChem* **2004,** *5*, 1623-1633.

(2) (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996,** *14*, 303-308. (b) Tyagi, S.; Bratu, P.; Kramer, F. R. *Nat. Biotechnol.* **1998,** *16*, 49-53.

(3) For examples of homogeneous gene sensing, see: (a) Stojanovic, M. N.; de Prada, P.; Landry, D. W. *ChemBioChem* **2001,** *2*, 411-415. (b) Ranasinghe, R. T.; Brown, L. J.; Brown, T. *Chem. Commun.* **2001,** 1480-1481. (c) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002,** *124,* 2096-2097. (d) Saghatelian, A.; Guckian, K. M.; Thayer, D. A.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2003,** *125,* 344-345. (e) Brunner, J.; Kraemer, R. *J. Am. Chem. Soc.* **2004,** *126*, 13626-13627. (f) Hartig, J. S.; Grune, I.; Najafi-Shoushtari, S. H.; Famulok, M. *J. Am. Chem. Soc.* **2004,** *126*, 722-723. (g) Pavlov, V.; Shlyahovsky, B.; Willner, I. *J. Am. Chem. Soc.* **2005,** *127,* 6522-6523. (h) Okamoto, A.; Tainaka, K.; Ochi, Y.; Kanatani, K.; Saito, I. *Mol. Biosyst.* **2006,** *2,* 122-127.

(4) Hasegawa, S.; Gowrishankar, G.; Rao, J. *ChemBioChem* **2006,** *7,* 925-928.

(5) (a) Sando, S.; Narita, A.; Abe, K.; Aoyama, Y. *J. Am. Chem. Soc.* **2005,** *127,* 5300-5301. (b) Narita, A.; Ogawa, K.; Sando, S.; Aoyama, Y. *Angew. Chem. Int. Ed. Engl.* **2006,** *45,* 2879-2883.

(6) Brantl, S. *Biochim. Biophys. Acta* **2002,** *1575,* 15-25.

(7) Isaacs, F. J.; Dwyer, D. J.; Ding, C.; Pervouchine, D. D.; Cantor, C. R. J.; Collins, J. *Nat. Biotechnol.* **2004,** *22*, 841-847.

(8) Babendure, J. R.; Adams, S. R.; Tsien, R. Y. *J. Am. Chem. Soc.* **2003,** *125,* 14716-14717.

(9) Kolpashchikov, D. M. *J. Am. Chem. Soc.* **2005,** *127,* 12442-12443.

(10) (a) Sparano, B. A.; Koide, K. *J. Am. Chem. Soc.* **2005,** *127,* 14954-14955. (b) Sparano, B. A.; Koide, K. *J. Am. Chem. Soc.* **2007,** *129,* 4785-4794.

(11) (a) Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346*, 818-822. (b) Robertson, D. L.; Joyce, G. F. *Nature* **1990,** *344,* 467-468. (c) Tuerk, C.; Gold, L. *Science* **1990,** *249*, 505-510.

(12) Werstuck, G.; Green, M. R. *Science* **1998,** *282*, 296-298.

(13) For examples of in vitro selection of aptamers against fluorophores, see: (a) Wilson, C.; Szostak, J. W. *Chem. Biol.* **1998,** *5,* 609-617. (b) Holeman, L. A.; Robinson, S. L.; Szostak, J. W.; Wilson, C. *Folding Des.* **1998,** *3,* 423-431.

(14) Jin, R.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988,** *85,* 8939-8942.

(15) Latt, S. A.; Stetten, G.; *J. Histochem. Cytochem.* **1976,** *24,* 24-33.

(16) Martin, R. M.; Leonhardt, H.; Cardoso, M. C. *Cytometry* **2005,** *67A,* 45-52.

(17) Argentini, M.; dos Santos, D. F.; Weinreich, R.; Hansen, H. J. *Inorg. Chem.* **1998,** *37,* 6018-6022.

(18) Breusegem, S. Y.; Clegg, R. M.; Loontiens, F. G. *J. Mol. Biol.* **2002,** 315, 1049-1061.

(19) The quantum yield of Hoechst **7d** (bound) was measured in the presence of a 3× molar excess of aptamer I-1-mini in 1× PBS containing 2.5 mM $MgCl_2$ with 9,10-diphenylanthracene in EtOH (Φ_F = 0.95) as a standard.

(20) Stoltenburg, R.; Reinemann, C.; Strehlitz, B.; *Anal. Bioanal. Chem.* **2005,** *383,* 83-91.

(21) Murphy, M. B.; Fuller, S. T.; Richardson, P. M.; Doyle, S. A. *Nucleic Acids Res. 2003,* 31, *e110.*

Chapter 5

RNA-Aptamer as a Tag Sequence to Enhance the Blue Fluorescence of Modified Hoechst Dye: A Strategy to Develop "Fluorescent RNA"

Abstract

We created a new light-up fluorophore/aptamer pair by a down-modification of Hoechst dye and re-selection of RNA-aptamer that binds to the modified Hoechst derivative. All of RNA aptamers, *in vitro*-selected against the immobilized Hoechst dye, induced a fluorescent enhancement of the Hoechst dye. The selected RNA sequences were optimized/minimized into a short 29-nt RNA-tag sequence. The short RNA aptamer sequence enhanced the fluorescence of the structurally optimized-Hoechst dye by approximately 50-folds. It was demonstrated that the optimized RNA-tag/Hoechst derivative pair can be used as a dynamic fluorescent label of mRNA transcription by fusing the tag sequence at 3'-side of mRNA.

Introduction

Green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* has proven widely useful as a method to fluorescently label proteins. Generally, the gene encoding a GFP is fused with the gene encoding the endogenous protein so that the resulting protein is tagged with the fluorescent module in the cells or tissues,[1] allowing monitoring of the location and fate of the target protein in real time by optical imaging (Figure 1a). There is a continuing effort to develop new fluorescent proteins (FPs) and these techniques of protein imaging have enhanced our understanding of cell biology.[2]

Although such protein imaging methods have been established so far, they can not simply be applied for RNA imaging. Current methods of fluorescent RNA imaging rely on the use of fluorescently labeled-molecules that can bind to RNA of interest.[3-6] However, these probes, wherein fluorophore or quenchers are covalently attached onto the probes, sometimes suffer from a low signal-to-noise ratio because of false-positive fluorescence of the probes. Now, there is necessary to develop a new method for real-time/dynamic monitoring of RNA such as mRNA and ncRNA (non-coding RNA) in living cells. One possible approach is to use intelligent fluorophores that light-up only when bind to specific RNA-tag sequences. Sparano and Koide recently reported light-up fluorophore/RNA sequence pair, based on fluorophore-quencher PET (photoinduced electron transfer) system that exhibits fluorescence enhancement upon binding to an RNA aptamer via perturbing PET.[7] Here, we report a new bio-orthogonal light-up fluorophore/RNA-aptamer pair that can be used as a dynamic blue fluorescent tag by fusing with mRNA of interest (Figure 1b).

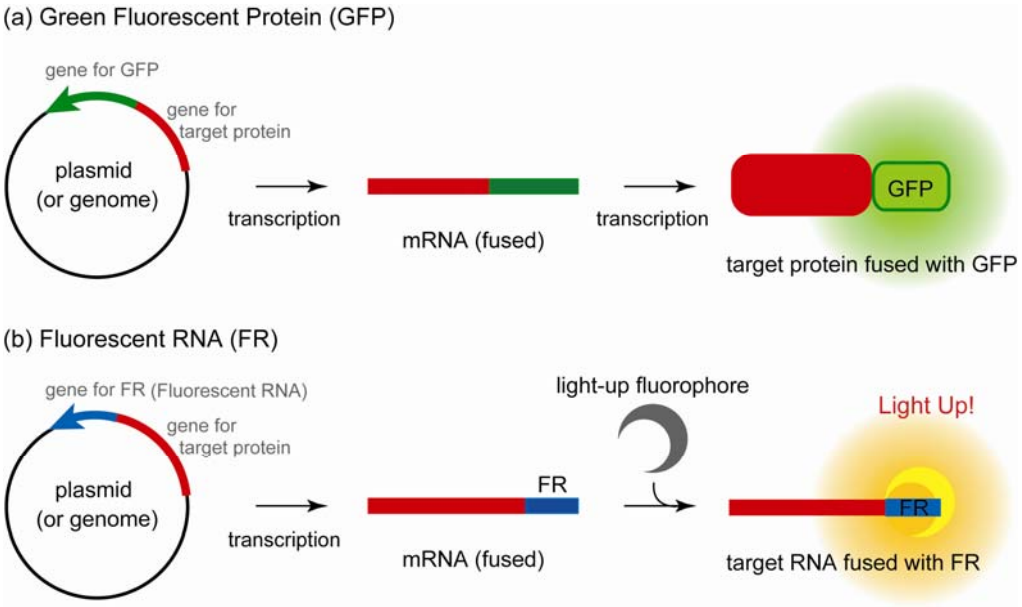


Figure 1. Concept of (a) Green Fluorescent Protein and (b) Fluorescent RNA.

Results and discussion

Recently, Stojanovic and Kolpashchikov reported aptameric sensors constructed from natural RNA components to detect small organic molecules by fluorescence,[8] which were termed as "green fluorescent RNA (GFR)" by Famlok.[9] Inspired by these systems, we have established a strategy to generate an aptamer-selective fluorophore from a conventional DNA-staining dye, Hoechst (chapter 5 of this thesis).[10] It is interesting that a simple *in vitro* selection could generate active aptamers, which work as a trigger for lighting up otherwise nonfluorescent fluorophores, without taking into account the molecular conformation or rigidity of the bound dye. The original Hoechst dye is a well-known fluorescent imager targeting the minor-groove of AT-rich dsDNA (double-stranded DNA), so that it is reasonable that the selected DNA-aptamer has AT-rich stem structure in aptamer core domain and successfully works as a trigger for lighting up Hoechst derivative **1a** (Figure 2a). Although the original Hoechst dye is specific to DNA structures, the present strategy can be, in principle, applicable to RNA. In this chapter, we applied our pair-creating strategy to generate a light-up Hoechst derivative/RNA-aptamer pair.

We carried out an *in vitro* selection[11] of an RNA-aptamer for Hoechst **1**, which is modified to suppress its binding to the original dsDNA target (Figure 2a). To immobilize Hoechst **1** on streptavidin-coated magnetic beads, Hoechst **1a** was conjugated with a biotin moiety through the polyethyleneglycol linker (for detail procedure, see Chapter 5 or our previous study[10]). RNA-aptamers directed against the immobilized Hoechst **1b** were selected from ssRNA (single-stranded RNA) pool with randomized sequence of 31-nt (total length = 71-nt, approximately 10^{15} different initial ssRNA molecules).[12] After the 8th round of the selection, 32 clones were isolated from the pool and sequenced. Twenty-five of the 32 isolated clones could be classified into three families (Classes I, II, and III in Figure 2b).

We first measured the enhancement in fluorescence. The fluorescence intensity of the otherwise practically non-fluorescent Hoechst **1a** (200 nM, excitation at 345 nm and emission at 470 nm) was enhanced by addition of all the selected aptamers, which indicated that the selected RNA-aptamer causes an environmental polarity change of the bound Hoechst derivative and works as a trigger for lighting up as well as the DNA-aptamer (Figure 2b). The aptamer of class II (200 nM), hereafter denoted as Aptamer II, showed the highest I_{on}/I_{off} ratio of 13.1 among the families (Figure 2c), so it was chosen for further investigation.

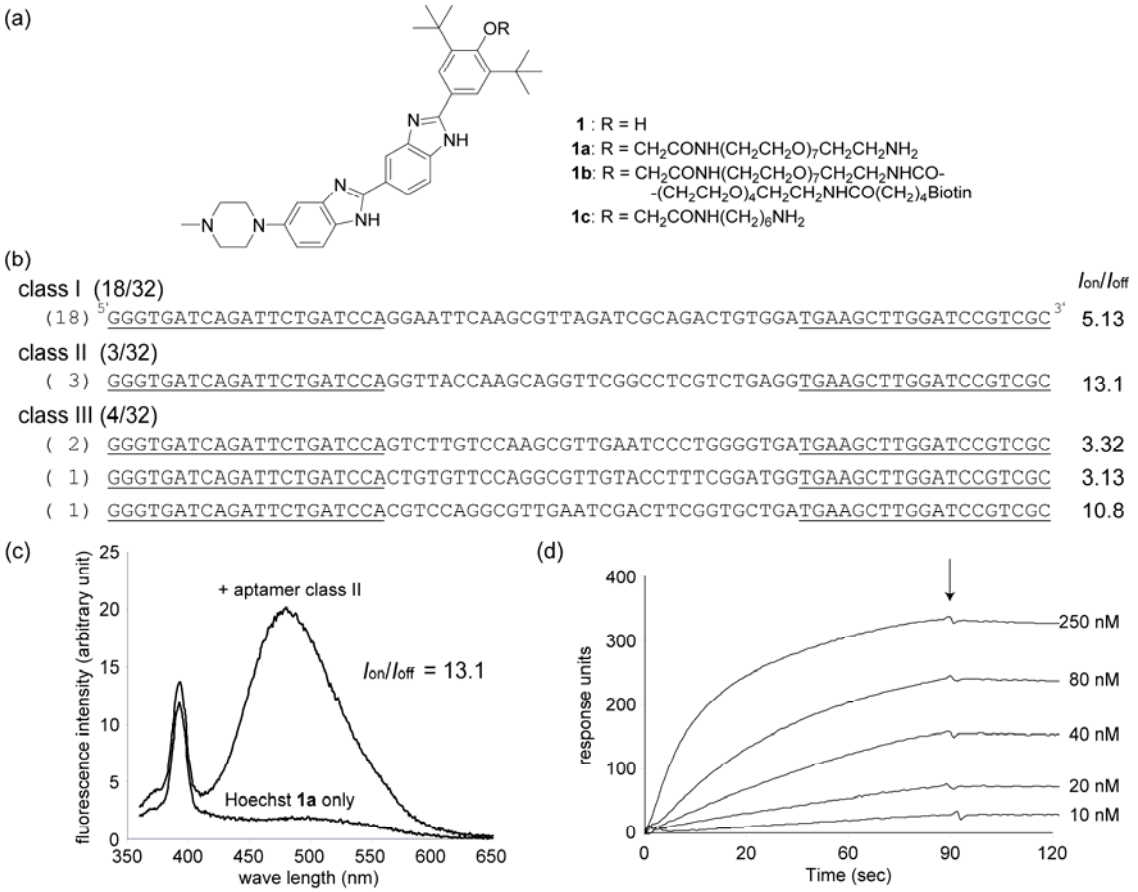


Figure 2. (a) Chemical structures of Hoechst derivatives **1**-**1c**. (b) The sequences of classes I-III aptamers. The numbers in parentheses indicate the sum of the individual sequences among 32 clones. The primer sites are underlined. On the right side, the fluorescence enhancements (I_{on}/I_{off} at 470 nm, ex. = 345 nm) of Hoechst **1a** (200 nM) with each aptamers (200 nM) are shown. (c) Fluorescence spectra of Hoechst **1a** (200 nM) in the absence or presence of Aptamer II (200 nM). The fluorescence intensity was measured in the binding buffer (1× PBS containing 2.5 mM $MgCl_2$) at 25 °C. (d) Display of the SPR sensorgrams showing the affinity of the Aptamer II for Hoechst **1b**. Various concentrations of Aptamer II (10-250 nM) were injected (total = 30 μL, t = 90 s) at a flow rate of 20 μL/min over a sensor chip SA pre-coated with Hoechst **1b**. The arrow indicates the end of injection.

The binding affinity of Aptamer II was assessed using SPR analysis. Biotin-labeled Hoechst **1b** was immobilized on a streptavidin-coated SPR sensor chip. Injection of various concentrations of Aptamer II (10-250 nM) over the chip exhibited distinct association and subsequent dissociation curves (Figure 2d). The binding data were fitted to a one-to-one binding equilibrium, and the dissociation constant (K_d) of Aptamer II against the immobilized **1b** was determined to be K_d = 3.2 nM.

We then proceeded to shorten Aptamer II to suppress the formation of various secondary structures that might destabilize the active aptamer-structure.[13] Measuring the enhancement in fluorescence for various shortened aptamers (Figure 3a) allowed us to minimize the aptamer sequence from 71-nt to 28-nt (Aptamer II–mini3 in Figure 3a). Aptamer II–mini3 enhanced the fluorescence of Hoechst **1a,** with I_{on}/I_{off} = 10.4. To shed more light on the aptamer sequence, additional modified aptamers were prepared and subjected to fluorescence enhancement measurements (Figure 3b). Among them, Aptamer II–mini3-4, whose G-U pair was substituted by G-C pair, most enhanced the fluorescence of **1a** with I_{on}/I_{off} = 28.2. These results, obtained from a series of modified aptamers, suggest that the sequence of the head-side stem and bulge domain are particularly important for the enhancement of fluorescence in **1a**. Unlike the DNA-aptamer sequence, the stem region of the RNA-aptamer does not contain significant AT-pairs. To obtain more efficient fluorescence enhancement, four Hoechst **1** derivatives were prepared (Scheme 1) and subjected to measurements of enhancement in fluorescence (Figure 4). Hoechst **1c** showed the most efficient fluorescence enhancement (I_{on}/I_{off} = 56.2) by the addition of aptamer II–mini3-4 (Figure 4d).

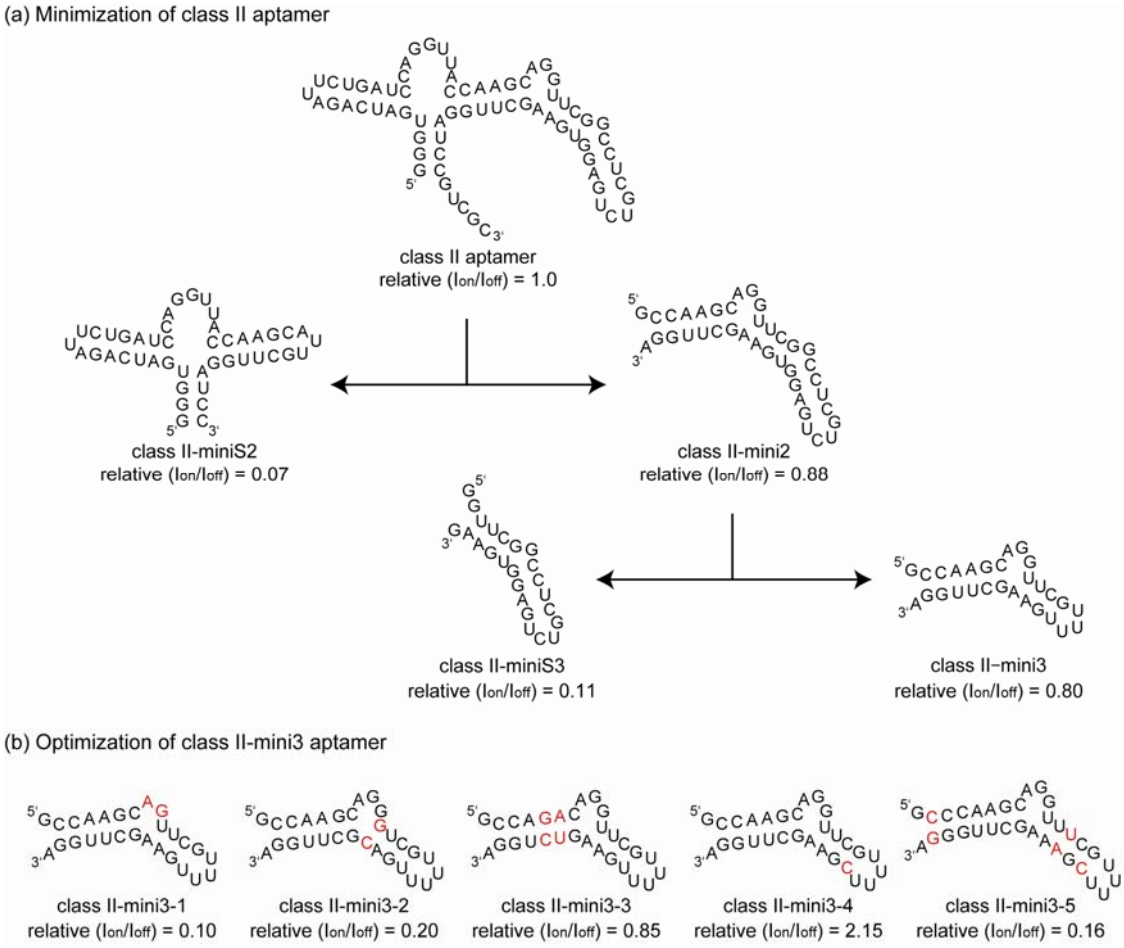
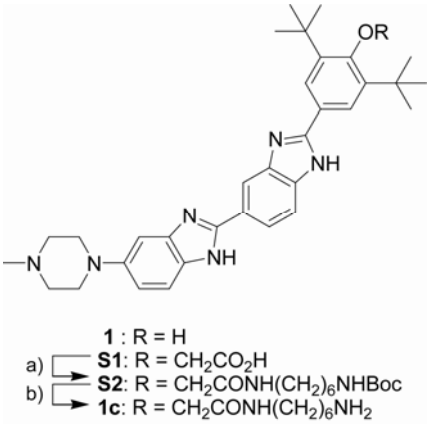


Figure 3. Predicted secondary structures of (a) class II aptamer, shortened aptamers and (b) optimized aptamers, generated by an RNAstructure program[14]. The relative fluorescence enhancements compared to that of Hoechst **1a**/Aptamer II pair [relative (I_{on}/I_{off})] are shown at the bottom of each structures.



Scheme 1. Synthesis of Hoechst derivative **1c**. Reagents and conditions: a) 1-ethyl-3-(3'-dimethylaminopropyl)carbodiimide (EDCI), *N*-hydroxysuccinimide (NHS), Boc-amino-alkyl-amine, diisopropylethylamine (DIPEA), DMF. b) 4 N HCl in ethyl acetate, 38% in two steps.

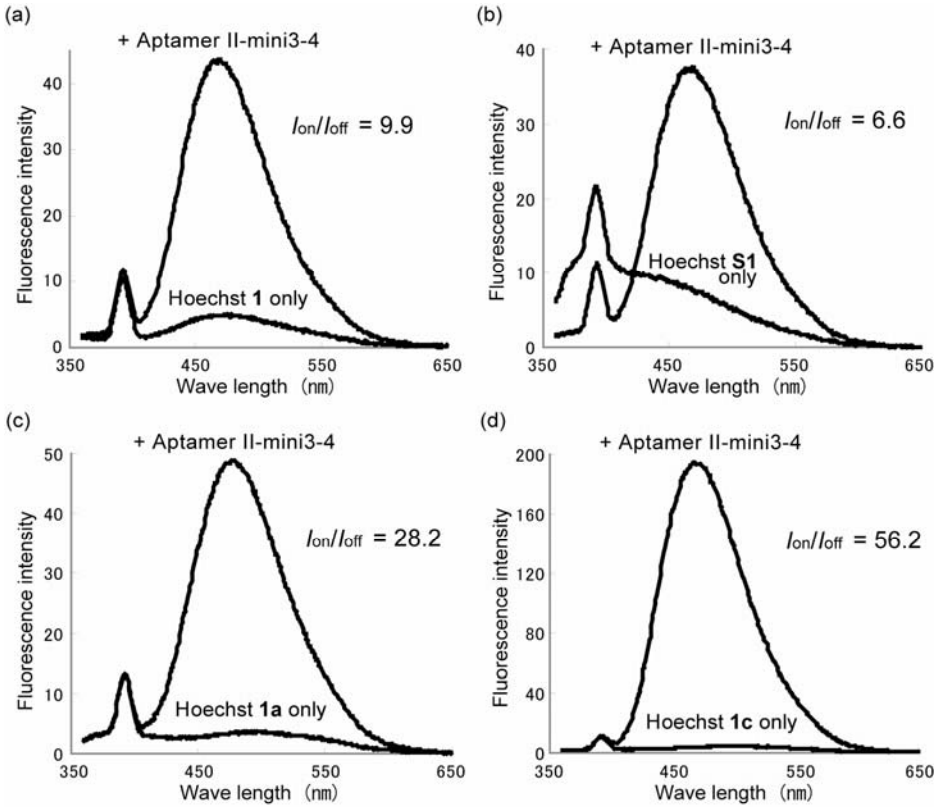


Figure 4. Fluorescence spectra of (a) Hoechst **1**, (b) Hoechst **S1**, (c) Hoechst **1a** and (d) Hoechst **1c** (200 nM) in the absence and in the presence of Aptamer II-mini3-4 (200 nM). The fluorescence intensity was measured in the binding buffer (1× PBS containing 2.5 mM $MgCl_2$) at 25 °C.

Finally, we moved on mRNA transcription monitoring using selected fluorophore/aptamer pair as a fluorescent RNA (FR). We prepared a dsDNA-template for T7-transcription of luciferase mRNA fused with the five-successive aptamer sequences, which based on the selected Aptamer II-mini3-4, at its 3'-terminus as a trigger domain for fluorescent signal (Figure 5a). The stem regions of the aptamers (shown in green), which were expected not to have considerable influence on the binding affinity, were partly modified to suppress unexpected self-hybridization. The RNA tag-fused or non-fused mRNAs were transcribed from dsDNA templates using T7 RNA polymerase and fluorescence enhancements of the co-existing Hoechst **1c** were monitored by microplate reader. As shown in Figure 5b, apparent fluorescence enhancement was observed during the course of transcription only for tag-fused mRNA. These results suggested that the selected RNA-aptamer for the Hoechst derivative can be used as a tag sequence to enhance the fluorescence and, furthermore, has the potential to work as a FR.

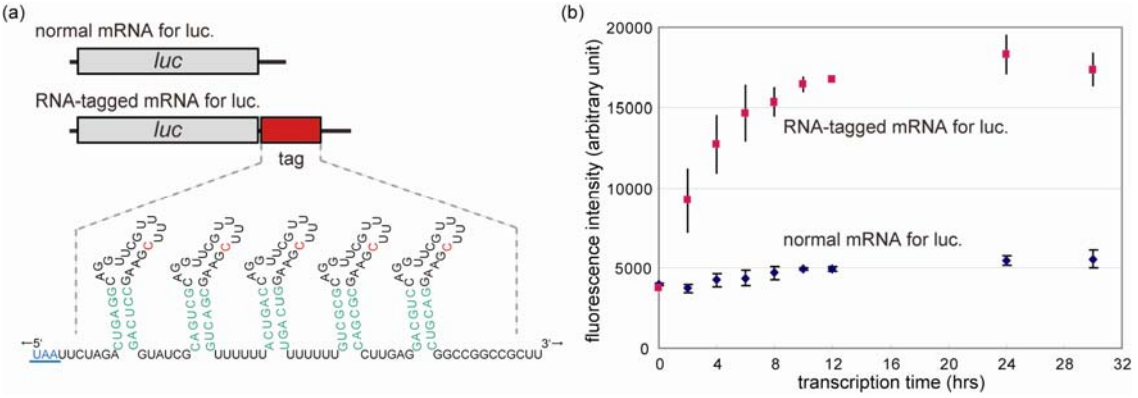


Figure 5. (a) Aptamer domain conjugated downstream of luciferase mRNA. The RNA-tag domain contains five-successive aptamer sequences based on the selected Aptamer II-mini3-4. The stem region (shown in green) is partly modified to suppress unexpected self-hybridization. Sequence underlined (blue) is stop codon of mRNA. (b) Transcription monitoring using light-up Hoechst/aptamer pair. mRNAs were transcribed from dsDNA-templates (6.8 ng/μL) in T7-buffer (Ambion: T7-MEGAscript) containing 7.5 mM NTPs, and 2.5 μM Hoechst **1c**. After incubation for various times, 5 μL of the reaction solution (50 μL) was mixed with 145 μL of binding buffer (pre-heated at 37 °C). The fluorescent intensities were measured by microplate reader. Error bars are standard deviations of the mean.

Conclusion

We developed a light-up Hoechst/RNA-aptamer pair, which has the potential to be used as a fluorescent RNA (FR, Figure 1b). The selected RNA-aptamer provided an environmental polarity change of the bound Hoechst derivative and actually worked as a trigger for lighting up of the dye. The RNA-aptamer, which consisted of unmodified RNA, was demonstrated to be applicable as a tag module to enhance the fluorescence of the Hoechst by fusing with mRNA of interest. In principle, the present strategy is applicable for any conventional fluorophores, giving a possibility for systematic generation of multicolor light-up fluorophore/aptamer pairs.

Experimental Section

General. Reagents and solvents were purchased from standard suppliers and used without further purification. The DNA oligomers were purchased from Gene Design Inc. (Japan). PCR was carried with an *i*Cycler thermalcycler Dual Block (Bio-Rad). Sequencing was carried out using a BigDye Terminator v3.1 Cycle Sequencing and Applied Biosystems 3130/3130xl Genetic Analyzer (Applied Biosystems). The fluorescence spectra were measured with a Shimadzu (Japan) RF-5300PC spectrofluorometer. The SPR measurements were performed with a BIAcore X system (GE Healthcare). The FAB mass spectra were recorded on a JEOL (Japan) JMS DX-300 spectrometer. The fluorescent intensities in transcription monitoring were measured with a 96-well Assay Plate with black bottom (Costar) and a Multilabel counter 1420 (Wallac).

Pool Construction. A ssDNA pool containing 31 randomized nucleotides (5'-GGT GAT CAG ATT CTG ATC CAN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNT GAA GCT TGG ATC CGT CGC-3') was purchased.[15] The pool was PCR amprified using primers (5'-GTA ATA CGA CTC ACT ATA GGG TGA TCA GAT TCT GAT CCA-3' and 5'-GCG ACG GAT CCA AGC TTC A-3'), where the underlined residues are part of the T7 RNA polymerase promoter and are not transcribed.

***In vitro* selection.** *In vitro* selection was carried out by referring conventional protocols with slight modifications. An ssRNA pool that containing 31 randomized nucleotides was transcribed from the dsDNA template using a T7 MEGAshortscript (Ambion). Transcribed RNA (60 nmol for the initial round of selection, 10 nmol for the second round, 1 nmol for the third and subsequent round) was dissolved in 150 μL of binding buffer [1× PBS (pH 7.4, 10 mM phosphate buffer containing 138 mM NaCl and 2.7 mM KCl) containing 2.5 mM $MgCl_2$] and then heated at 60 °C for 3 min, cooled to RT (room temprature) and left for 10 min. To exclude mis-binding RNA sequences from the pool, negative selection was performed before each positive selection round. In the negative selection, the pre-annealed RNA library was applied to streptavidin-coated magnetic beads (1 mg, pre-washed with binding buffer) and incubated for 15 min at RT under mild shake. The recovered RNA, which was not bound to the streptavidin-coated magnetic beads,

was then subjected to positive selection. The recovered RNA solution was incubated with Hoechst **1b**-immobilized magnetic beads (0.1 mg, pre-washed with binding buffer) for 15 min at RT under mild shake (for immobilization procedure, see previous study[10]). After the beads were washed three times with 1000 μL of wash buffer [1× PBS (pH 7.4, 10 mM phosphate buffer containing 138 mM NaCl and 2.7 mM KCl) containing 2.5 mM $MgCl_2$ and 0.05% Tween20], they were detached from the magnetic beads with 400 μL of elution buffer (40 mM Tris-HCl pH 8.0 containing 0.02% Tween20, 10 mM EDTA, and 3.5 M urea) by 30 min incubation at RT. The recovered RNA was purified by 2-propanol precipitation and ethanol precipitation. Collected RNA was reverse transcrobed and amplified using a PrimeScript One Step RT-PCR Kit (Takara). Amplified dsDNA was used for the next round.

Cloning and Sequencing. The RT-PCR products of the 8th round pool were ligated into a pDrive Cloning Vector (Qiagen) and cloned into EZ Competent Cell (Qiagen) according to the manufacture's manual. Plasmid DNAs were isolated and sequenced using a BigDye Terminator v3.1 Cycle Sequencing and Applied Biosystems 3130/3130xl Genetic Analyzer (Applied Biosystems).

Fluorescence measurements. Fluorescence spectra were obtained at 25 °C. The Hoechst derivatives (200 nM) were dissolved in binding buffer in the presence (200 nM), or absence of aptamers. The solutions were excited at 345 nm, and the emissions were monitored in the 360-650 nm wavelength range. The Fluorescent enhancement (I_{on}/I_{off}) was determined by comparing the intensity of the fluorescence emissions at 470 nm.

Surface plasmon resonance analysis. SPR measurements were performed with the BIAcore X system (GE Healthcare). The Hoechst **1b** was immobilized on a sensor chip SA (streptavidin) in a continuous flow of HBS-N buffer (pH 7.4, 10 mM HEPES containing 150 mM NaCl) at a flow rate of 10 μL/min. A solution (60 μL) of Hoechst **1b** (1.0 μM) was injected onto the streptavidin-coated gold surface of the chip, and nonspecifically bound materials were washed off with 100 mM NaOH. Binding experiments with various concentrations of Aptamer class II (30 μL) were performed at 25 °C in a continuous flow of running buffer (1× PBS containing 2.5 mM Mg^{2+}) at a flow rate of 20 μL/min. Kinetic analyses were performed with BIAevaluation 3.1 software (GE Healthcare).

Preparation of various shortened RNA-aptamers. The dsDNA templates for RNA-aptamers were PCR amplified using forward and reverse primers (for Aptamer II-mini2 5'-GTA ATA CGA CTC ACT ATA GCC AAG CAG GTT CGG CCT CGT CTG-3' and 5'-TCC AAG CTT CAC CTC AGA CGA GGC-3', for Aptamer II-mini3 5'-GTA ATA CGA CTC ACT ATA GCC AAG CAG G-3' and 5'-TCC AAG CTT CAA AAC GAA CCT GCT TGG C-3', for Aptamer II-mini3-1 5'-GTA ATA CGA CTC ACT ATA GCC AAG CAG-3' and 5'-TCC AAG CTT CAA AAC GAA CTG CTT GGC-3', for Aptamer II-mini3-2 5'-GTA ATA CGA CTC ACT ATA GCC AAG CAG G-3' and 5'-TCC AAG CGT CGA AAA CGA CCC TGC TTG GC-3', for Aptamer II-mini3-3 5'-GTA ATA CGA CTC ACT ATA GCC AGA CAG G-3' and 5'-TCC AGA CTT CAA AAA CGA ACC TGT CTG GC-3', for Aptamer II-mini3-4 5'-GTA ATA CGA CTC ACT ATA GCC AAG CAG G-3' and 5'-TCC AAG CTT CGA AAA CGA ACC TGC TTG GC-3', for Aptamer II-mini3-5 5'-GTA ATA CGA CTC ACT ATA GCC CAA GCA GG-3' and 5'-TCC CAA GCT TTC GAA AAC GAA ACC TGC TTG GGC-3', for Aptamer II-miniS2 5'-GTA ATA CGA CTC ACT ATA GGG TGA TCA GAT TCT GAT CCA GGT TAC-3' and 5'-GGA TCC AAG CAA TGC TTG GTA ACC TGG-3', for Aptamer II-miniS3 5'-GTA ATA CGA CTC ACT ATA GGT TCG G-3' and 5'-CTT CAC CTC AGA CGA GGC CGA ACC TAT AGT G-3'), where the underlined residues are part of the T7 RNA polymerase promoter and are not transcribed. RNA-aptamers were transcribed from the dsDNA templates using a T7 MEGAshortscript (Ambion), and purified by 8% PAGE (polyacrylamide gel electrophoresis) containing 7 M urea.

***N*-(6-aminohexyl)-2-(2,6-di-*tert*-butyl-4-(5-(4-methylpiperazin-1-yl)-1*H*,1'*H*-2,5'-bi benzo[*d*]imidazol-2'-yl)phenoxy)acetamide (Hoechst derivative 1c).** 1-ethyl-3-(3'-dimethylaminopropyl)carbodiimide (EDCI, 6.4 mg, 0.034 mmol), *N*-hydroxysuccinimide (NHS, 3.9 mg, 0.034 mmol), and diisopropylethylamine (DIPEA, 4.4 mg, 0.034 mmol) were added to a solution of Hoechst **S1**·TEAA salt (2 mg, for synthetic procedure of **S1**, see previous study[10]) in dry DMF (800 μL). The resulting mixture was stirred under N_2 at RT. After stirring for a period of 1 h, Boc-amino-alkyl-amine (Polypure, $H_2N(CH_2)_6NHBoc$) was added, and the mixture was stirred under N_2 for 10 h at RT. The crude product was purified by HPLC and lyophilized to give Hoechst **S2**. **S2** was then dissolved in CH_3OH (480 μL), and mixed with 4N HCl in ethyl acetate (480 μL). The mixture was stored for 2 h at RT and the organic solvent was removed. The crude product was purified by HPLC and lyophilized to give Hoechst **1c** in a yield of 38%. The concentration of **1c**, when

needed, was determined by use of the known extinction coefficient of Hoechst 33258. MALDI-TOF Mass: *m*/*z* calcd = 693.95 for $[M + H]^+$; found: 693.28.

Preparation of the template for luciferase mRNA fused with an aptamer domain. A dsDNA encoding the aptamer domain, which contains five aptamer sequences, was PCR amplified using template 5'-GTA TCG CAG TCC CAG GTT CGT TTT CGA AGG GAC TGT TTT TTA CTG ACC AGG TTC GTT TTC GAA GGT CAG TTT TTT TGT CGG GCG GTT CGT TTT CGA AGC CCG ACC TTG AGG ACG GAC AGG-3', forward primer 5'-ATC GCC GTG TAA TTC TAG ACT GAG GCA GGT TCG TTT TCG AAG CCT CAG GTA TCG CAG TCC CAG G-3' and reverse primer 5'-GTC TGC TCG AAG CGG CCG GCC GAC GGA CTT CGA AAA CGA ACC TGT CCG TCC TCA AG-3'. The amplified dsDNA was inserted into the Xba I-Fse I site immediately downstream of the stop codon in the pGL3-Control Vector (Promega). The template for T7-transcription of mRNA which has aptamer domain was PCR amplified using this vector for template, forward primer 5'-<u>GTA ATA CGA CTC ACT ATA</u> GGC ATT CCG GTA CTG-3' and reverse primer 5'-GTT GTT AAC TTG TTT ATT GCA GCT TAT AAT GG-3', where the underlined residues are part of the T7 RNA polymerase promoter and are not transcribed. The amplified dsDNA was purified using a QIAquick PCR Purification Kit (Qiagen) according to the manufacture's manual.

Transcription monitoring using light-up Hoechst 1c/aptamer pair. The mRNAs were transcribed from the dsDNA-templates (6.8 ng/μL) in 50 μL of T7-buffer (Ambion: T7-MEGAscript) containing 7.5 mM NTPs, and 2.5 μM Hoechst **1c**. After incubation for various times, 5 uL of the reaction solution was mixed with 145 uL of binding buffer [1× PBS (pH 7.4, 10 mM phosphate buffer containing 138 mM NaCl and 2.7 mM KCl) containing 2.5 mM $MgCl_2$, pre-heated at 37 °C]. The fluorescent intensities were measured by microplate reader.

References

(1) Tsien, R. Y. *Annu. Rev. Biochem.* **1998,** *67*, 509-544.

(2) (a) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006,** *312*, 217-224. (b) Chudakov, D. M.; Lukyanov S.; Lukyanov K. A. *Trends Biotechnol* **2005,** *23*, 605-613. (c) Zhang, S.; Ma, C.; Chalfie, M. *Cell* **2004,** *119*, 137-144.

(3) (a) Ekong, R.; Wolfe, J. *Curr. Opin. Biotechnol.* **1998,** *9*, 19-24. (b) Levsky, J. M.; Singer, R. H. *J. Cell Sci.* **2003,** *116*, 2833-2838.

(4) (a) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol.* **1996,** *14*, 303-308. (b) Santangero, P.; Nitin, N.; Bao, G. *Ann. Biomed. Eng.* **2006,** *34*, 39-50.

(5) (a) Bertrand, E.; Chartrand, P.; Schaefer, M.; Shenoy, S. M.; Singer, R. H.; Long, R. M. *Mol. Cell* **1998,** *2*, 437-445. (b) Janicki, S. M.; Tsukamoto, T.; Salghetti, S. E.; Tansey, W. P.; Sachidanandam, R.; Prasanth, K. V.; Ried, T.; Shav-Tal, Y.; Bertrand, E.; Singer, R. H.; Spector, D. L. *Cell* **2004,** *116*, 683-698.

(6) (a) Sando, S.; Kool, E. T. *J. Am. Chem. Soc.* **2002,** *124*, 9686-9687. (b) Sando, S.; Narita, A.; Sasaki, T.; Aoyama, Y. *Org. Biomol. Chem.* **2005,** *3*, 1002-1007.

(7) Sparano, B. A.; Koide, K. *J. Am. Chem. Soc.* **2007,** *129*, 4785-4794.

(8) Stojanovic, M. N.; Kolpashchikov, D. M. *J. Am. Chem. Soc.* **2004,** *126*, 9266-9270.

(9) Famulok, M. *Nature* **2004,** *430*, 976-977.

(10) Sando, S.; Narita, A.; Aoyama, Y. *ChemBioChem* **2007,** *8*, 1795-1803.

(11) (a) Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346*, 818-822. (b) Robertson, D. L.; Joyce, G. F. *Nature* **1990,** *344,* 467-468. (c) Tuerk, C.; Gold, L. *Science* **1990,** *249*, 505-510.

(12) Werstuck, G.; Green, M. R. *Science* **1998,** *282*, 296-298.

(13) Shangguan, D.; Tang, Z.; Mallikaratchy, P.; Xiao, Z.; Tan, W. *ChemBioChem* **2007,** *8*, 603-606.

(14) Predictions of RNA secondary structures were carried out by an RNAstructure Version 4.4 program (Mathews, D. H.; Zuker, M.; Turner, D. H.)

(15) Singh, R.; Valcarcel, J.; Green, M. R. *Science* **1995,** *268*, 1173-1176.

List of Publications

Chapter 1 Locked TASC Probes for Homogeneous Sensing of Nucleic Acids and Imaging of Fixed *E. coli* Cells
Shinsuke Sando, Atsushi Narita, Toshinori Sasaki, and Yasuhiro Aoyama
Org. Biomol. Chem. **2005,** *3,* 1002-1007.

Chapter 2 Doubly Catalytic Sensing of HIV-1-Related CCR5 Sequence in Prokaryotic Cell-Free Translation System using Riboregulator-Controlled Luciferase Activity
Shinsuke Sando, Atsushi Narita, Kenji Abe, and Yasuhiro Aoyama
J. Am. Chem. Soc. **2005,** *127,* 5300-5301.

Highly Sensitive Genotyping using Artificial Riboregulator System
Atsushi Narita, Kazumasa Ogawa, Shinsuke Sando, and Yasuhiro Aoyama
Nucleic Acids Res. Suppl. **2005,** *49,* 271-272.

Chapter 3 Visible Sensing of Nucleic Acid Sequences with a Genetically Encodable Unmodified RNA Probe
Atsushi Narita, Kazumasa Ogawa, Shinsuke Sando, and Yasuhiro Aoyama
Angew. Chem. Int. Ed. **2006,** *45,* 2879-2883.

Cis-Regulatory Hairpin-Shaped mRNA Encoding a Reporter Protein: Catalytic Sensing of Nucleic Acid Sequence at Single Nucleotide Resolution
Atsushi Narita, Kazumasa Ogawa, Shinsuke Sando, and Yasuhiro Aoyama
Nat. Protoc. **2007,** *2,* 1105-1116.

Visible Sensitive of Nucleic Acid Sequences using a Genetically Encodable Unmodified mRNA Probe
Atsushi Narita, Kazumasa Ogawa, Shinsuke Sando, and Yasuhiro Aoyama
Nucleic Acids Res. Suppl. **2006,** *50,* 283-284.

Chapter 4 Light-Up Hoechst–DNA Aptamer Pair : Generation of an Aptamer-Selective Fluorophore from a Conventional DNA-Staining Dye
Shinsuke Sando, Atsushi Narita, and Yasuhiro Aoyama
ChemBioChem **2007,** *8,* 1795-1803.

Sensing of Nucleic Acid Sequences using Unmodified Nucleic Acid as a Probe
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
Nucleic Acids Res. Suppl. **2007,** *51,* 281-282.

Chapter 5 RNA-Aptamer as a Tag Sequence to Enhance the Blue Fluorescence of Hoechst Dye

Atsushi Narita, Masayoshi Hayami, Shinsuke Sando, and Yasuhiro Aoyama

To be submitted.

Other Publications

1. A Facile Route to Dynamic Glycopeptide Libraries based on Disulfide-Linked Sugar–Peptide Coupling
Shinsuke Sando, Atsushi Narita, and Yasuhiro Aoyama
Bioorg. Med. Chem. Lett. **2004,** *14*, 2835-2838.

2. Tuning the Hoechst Dye into Color-Changing Fluorescent pH Indicator in an Acidic Range
Shinsuke Sando, Atsushi Narita, Masayoshi Hayami, and Yasuhiro Aoyama
Chem. Lett. To be submitted.

List of Presentations

International Symposium

1. Highly Sensitive Genotyping using Artificial Riboregulator System
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
4th International Symposium on Nucleic Acids Chemistry, Kyusyu University, Hakata, Japan, September 2005.

2. Doubly Catalytic Sensing of HIV-1-Related CCR5 Sequence in Prokaryotic Cell-free Translation System using Riboregulator-Controlled Lluciferase Activity
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
PACIFICHEM 2005, Mid Pacific Conference Center, Honolulu, U.S.A., December 2005.

3. Sensing of Nucleic Acid Sequence using Unmodified DNA/RNA as a Probe
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
234th ACS National Meeting, Boston Convention and Exhibition Center, Boston, U.S.A., August 2007.

4. Sensing of Nucleic Acid Sequences using Unmodified Nucleic Acid as a Probe
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
5th International Symposium on Nucleic Acids Chemistry, Tokyo University, Tokyo, Japan, November 2007.

Domestic Symposium

1. Bottom-Up Construction of Glycopeptide Mimetic Libraries by Disulfide Exchange Reaction
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
1st Joint Symposium on 18th Biofunctional Chemistry and 7th Biotechnology, Kumamoto University, Kumamoto, Japan, October 2003.

2. Construction of Self-evolving Glycopeptide Libraries
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
84th Annual Meeting of Chemistry Society of Japan, Kwansei Gakuin University, Nishinomiya, Japan, March 2004.

3. A new Genotyping Method using a Conformational Change of *Cis*-Acting Nucleic Acid
Atsushi Narita, Kenji Abe, Shinsuke Sando, and Yasuhiro Aoyama
21st COE (Center of Excellence) 4th Chemistry Coursework on Bio-related Material Chemistry, Kyoto University, Kyoto, Japan, February 2005.

4. A new Genotyping Method using a Conformational Change of *Cis*-Acting Nucleic Acid
Atsushi Narita, Kenji Abe, Shinsuke Sando, and Yasuhiro Aoyama
85th Annual Meeting of Chemistry Society of Japan, Kanagawa University, Yokohama, Japan, March 2005.

5. Highly Sensitive Genotyping using Artificial Riboregulator System
Atsushi Narita, Kazumasa Ogawa, Shinsuke Sando, and Yasuhiro Aoyama
20th Symposium on Biofunctional Chemistry, Nagoya City University, Nagoya, Japan, September 2005.

6. Molecular Beacon-mRNA: Sensing of Nucleic Acid Sequences using Unmodified RNA as a Probe
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
21st COE (Center of Excellence) 4th Chemistry Coursework on Bio-related Material Chemistry, Kyoto University, Uji, Japan, February 2006.

7. Genotyping with Unmodified RNA Probe: Signal Amplification using RNase H-Activity Coupled Cell-Free Translation System
Kazumasa ogawa, Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
86th Annual Meeting of Chemistry Society of Japan, Nihon University, Funabashi, Japan, March 2006.

8. Visible Sensing of Nucleic Acid Sequences using a Genetically Encodable Unmodified mRNA Probe
Atsushi Narita, Kazumasa Ogawa, Shinsuke Sando, and Yasuhiro Aoyama
33rd Symposium on Nucleic Acids Chemistry, Osaka University, Suita, Japan, November 2006.

9. Efforts toward Cell-Compatible Genotyping System
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
2nd Annual Meeting of Japanese Society for Chemical Biology, Kyoto University, Kyoto, Japan, March 2007.

10. Efforts toward Cell-Compatible Genotyping System
Atsushi Narita, Kazumasa ogawa, Shinsuke Sando, and Yasuhiro Aoyama
87th Annual Meeting of Chemistry Society of Japan, Kansai University, Suita, Japan, March 2007.

11. Approach toward Cell-Compatible Nucleic Acid Imaging: Sensing of Nucleic Acid Sequences using Unmodified Nucleic Acid as a Probe
Atsushi Narita, Shinsuke Sando, and Yasuhiro Aoyama
22th Symposium on Biofunctional Chemistry, Tohoku University, Sendai, Japan, September 2007.

List of Honors

SAFC-Proligo Award
33rd Symposium on Nucleic Acids Chemistry, Osaka University, Suita, Japan, November 2006.

Nucleic Acids Research Award
5th International Symposium on Nucleic Acids Chemistry, Tokyo University, Tokyo, Japan, November 2007.