Diazirine-containing RNA photocrosslinking probes for the study of siRNA–protein interactions†

Satoru Kuboe,‡[a] Mayuko Yoda,‡[cd] Aya Ogata,[a] Yukio Kitade,[ab] Yukihide Tomari*[cd] and Yoshihito Ueno*[ab]

Received 9th July 2010, Accepted 16th August 2010
DOI: 10.1039/c0cc02450c

We here report the synthesis and characterization of small interfering RNAs with aryl trifluoromethyl diazirine moieties in the 3′-overhang regions, which allow sensitive detection of interacting proteins during assembly of the effector ribonucleoprotein complex by irradiation with minimally destructive long-wavelength ultraviolet light.

RNA interference (RNAi) is a process of sequence-specific posttranscriptional gene silencing that is triggered by double-stranded RNAs (dsRNAs).[1] This pathway is initiated by processing of dsRNA into 21- to 23-nucleotide (nt) duplexes by Dicer. These small RNA duplexes, which contain a 2-nt overhang at the 3′-end of each strand, are termed small interfering RNAs (siRNAs). siRNAs are incorporated into the effector ribonucleoprotein complex, termed RNA-induced silencing complex (RISC), which catalyzes the sequence-specific cleavage of target mRNAs.[1,2]

A member of the Argonaute (Ago) family of proteins lies at the heart of RISC, which provides the unique platform for target recognition and cleavage. RISC assembly is not a mere binding of siRNAs to Ago proteins, but rather follows a complex, multi-step pathway.[3] In *Drosophila melanogaster*, it is well established that siRNA duplexes are first incorporated into RISC-loading complex (RLC), which contains Dicer-2 (Dcr-2) and R2D2.[4–6] RLC senses the thermodynamic asymmetry of the siRNA duplex, with R2D2 located at the more stable end and Dcr-2 at the less stable end, and ascertains that the strand with the less stable 5′ end is eventually selected as the guide in mature RISC.[7] Ago2 then receives the siRNA duplex with the predetermined orientation from RLC. This process is called RISC loading and requires ATP and the Hsc70/Hsp90 chaperone machinery.[8] After RISC loading, Ago2 cleaves the center of one strand (the passenger strand) of the siRNA duplex, just like it cleaves target mRNAs.[9–11]

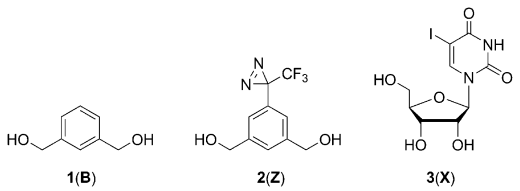
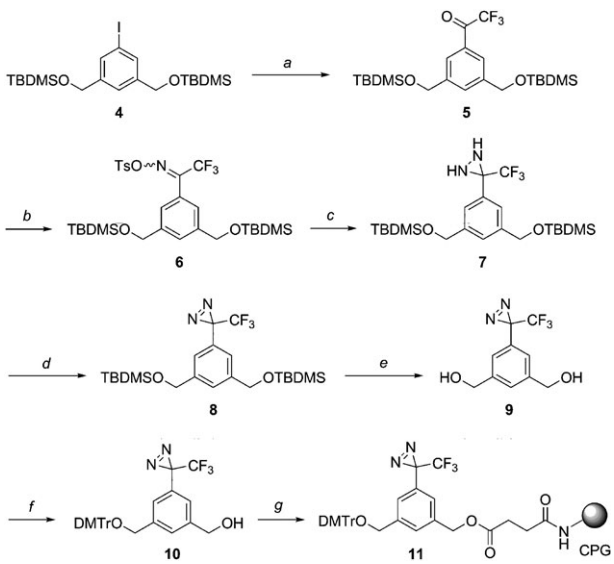


Fig. 1 Structures of functional compounds used in this study.

The cleaved passenger strand is discarded and rapidly degraded, with only one strand (the guide strand) remaining in Ago2. This process is termed unwinding, and the Ago2 protein containing only the guide strand is called mature RISC, or simply, RISC. Although much is known about RISC assembly of fly Ago2, the protein factors that are required for this entire pathway are not fully understood.[3]

Photocrosslinking has been a tremendously useful technique to study RISC assembly. For example, asymmetric binding of Dcr-2/R2D2 in RLC to the siRNA duplex, and transition from Dcr-2/R2D2 in RLC to Ago2 in RISC can be monitored by using the photocrosslinker 5-iodouridine (**3**) introduced at position 20 of the guide strand (Fig. 1).[7] However, the



Scheme 1 *Reagents and conditions*: (a) CF_3CO_2Et, *n*-BuLi, THF, −78 °C, 1 h, 67%; (b) (1) $HONH_2$·HCl, EtOH/pyridine (1 : 1, v/v), 60 °C, 12 h; (2) TsCl, Et_3N, DMAP, $CHCl_3$, rt, 12 h, 85%; (c) NH_3, THF, rt, 48 h, 60%; (d) I_2, Et_3N, MeOH, rt, 30 min, 74%; (e) TBAF, THF, rt, 2 h, 85%; (f) DMTrCl, pyridine, rt, 5 h, 38%; (g) (1) succinic anhydride, DMAP, pyridine, rt, 24 h; (2) CPG, EDCI, DMF, rt, 48 h, 28 μmol g^{-1} (loading amount).

[a] *Department of Biomolecular Science, Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan. E-mail: uenoy@gifu-u.ac.jp; Fax: +81-58-293-2794; Tel: +81-58-293-2639*
[b] *United Graduate School of Drug Discovery and Medical Information Sciences, Gifu University, Japan*
[c] *Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo, 113-0032, Japan. E-mail: tomari@iam.u-tokyo.ac.jp; Fax: +81-3-5841-8485; Tel: +81-3-5841-7839*
[d] *Department of Medical Genome Sciences, The University of Tokyo, Japan*

† Electronic supplementary information (ESI) available: General experimental section, protocols for synthesis, purification and characterization of aromatic compounds and ONs, and protocols for biological experiments. See DOI: 10.1039/c0cc02450c

‡ These authors contributed equally to this work.

crosslinking efficiency of 5-iodouridine is limited, and irradiation with ultraviolet (UV)-B (280–315 nm) to activate 5-iodouridine poses a risk of damaging biomacromolecules and their interactions. Therefore, an alternative photocrosslinker with higher crosslinking efficiency at a longer wavelength has long been awaited.

Aromatic azides, aromatic diazirines, and benzophenones are representative photoreactive functional groups that generate highly reactive species such as nitrenes, carbenes, and diradicals, respectively, by irradiation with light. Among them, aromatic azides are the most widely used because of their synthetic simplicity.[12] On the other hand, aromatic diazirines are considered the most effective functional group because they generate carbenes, which can react with various chemical groups including the generally inactive C–H bond.[13] Aromatic diazirines are also preferred because they are photoactivatable with long-wavelength UV-A (>315 nm), which is less likely to damage biomacromolecules than UV-B. It has been shown that diazirine–DNA conjugates are useful for the study of DNA–protein interactions.[14]

We have recently reported the synthesis of siRNA possessing an aromatic compound 1,3-bis(hydroxymethyl)benzene (**1**) in the 3′-overhang region (Fig. 1).[15] We found that this modified siRNA was more potent in terms of target gene silencing than the siRNA without 3′-overhang regions; further, the modified siRNA had silencing activity equivalent to that of siRNA possessing natural nucleosides at the 3′-overhang region. On the basis of this information, we report the synthesis and characterization of siRNA with aryl trifluoromethyldiazirine moieties **2** in the 3′-overhang region (Fig. 1); these moieties show robust photocrosslinking to siRNA-interacting proteins on irradiation with minimally destructive long-wavelength UV.

1,3-Bis(*tert*-butyldimethylsilyloxymethyl)-5-iodobenzene (**4**), which was prepared as described previously,[16] was trifluoroacetylated by lithium halogen exchange at −78 °C using *n*-BuLi followed by addition of ethyl trifluoroacetate (Scheme 1). The resulting trifluoroacetophenone **5** was converted to a stereoisomeric (*E*/*Z*) mixture of oximes by treatment with $HONH_2{\cdot}HCl$. The crude mixture of oximes was tosylated to generate a mixture of tosyl-oximes **6**. This mixture was converted to diaziridine **7** upon addition of ammonia, and **7** was oxidized to the corresponding diazirine by using I_2. Deprotection of the silyl ether using tetra-n-butylammonium fluoride (TBAF) afforded the benzylic alcohol **9**. One out of the two hydroxy groups of **9** was protected by a 4,4′-dimethoxytrityl (DMTr) group to produce a mono-DMTr derivative **10**. To enable attachment to the solid support, **10** was succinated to yield the corresponding succinate, which was linked to controlled-pore glass (CPG) to afford the solid support **11** possessing **10**. siRNA sequences used in this study are depicted in Fig. 2.

All the oligonucleotides (ONs) were synthesized by a phosphoramidite method. Fully protected ONs (1.0 μmol each) were treated with concentrated NH_4OH : EtOH (3 : 1, v/v) at room temperature for 12 h and then with 1.0 M TBAF/THF at room temperature for 12 h. The ONs released after the treatment were purified by denaturing 20% polyacrylamide gel electrophoresis (PAGE) to afford deprotected ONs **5** and **9** carrying the aryl diazirine in 14 and 11 OD_{260} absorbance units,

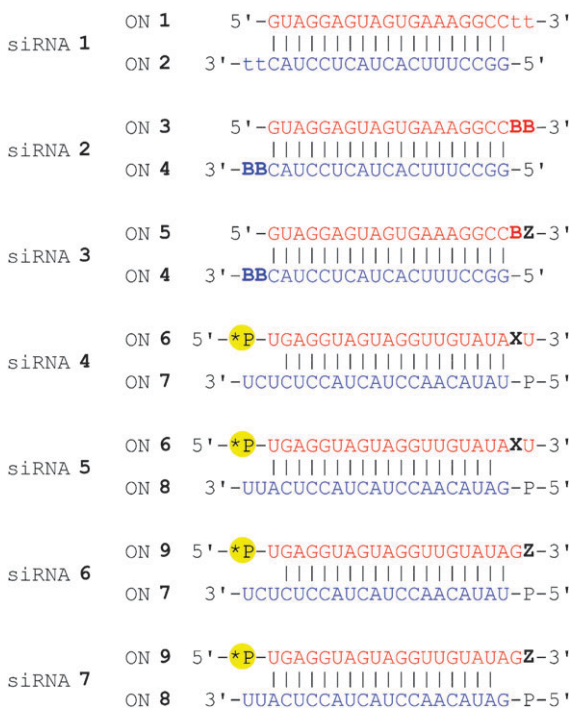


Fig. 2 Sequences of the oligonucleotides (ONs) and siRNA used in this study. Capital letters indicate ribonucleosides, whereas the small italicized letters show 2′-deoxyribonucleosides. **B** denotes 1,3-bis-(hydroxymethyl)benzene. **Z** and **X** indicate aryl diazirine and 5-iodouridine, respectively. In the case of siRNAs **4**–**7**, the upper strands were 5′ radiolabeled with ^{32}P (highlighted in yellow), whereas the lower strands bore nonradioactive 5′ phosphate.

respectively. These ONs were analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/MS), and the observed molecular weights were in agreement with their structures.

We first examined the silencing activity of siRNA carrying the aryl diazirine in its 3′-overhang region by a dual-reporter assay using a psiCHECK-2 vector in HeLa cells. The vector contains the *Renilla* and firefly luciferase genes, and the siRNA sequences were designed to target the *Renilla* luciferase gene. HeLa cells were co-transfected with the vector and indicated amounts of the siRNA, and the signals of *Renilla* luciferase were normalized to those of firefly luciferase. The silencing activity of siRNA with the aryl diazirine (siRNA **3**) was found to be comparable to that of siRNA possessing the natural nucleoside (siRNA **1**) or **1** (siRNA **2**) in the 3′-overhang region (Fig. 3).

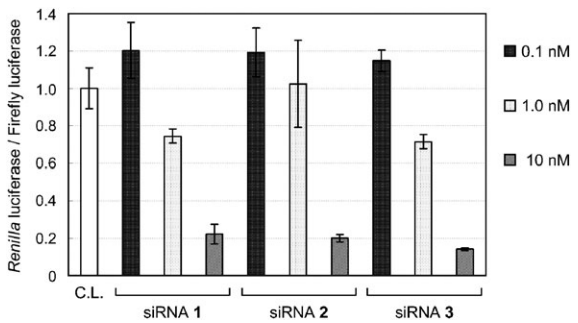


Fig. 3 Dual-luciferase assay.

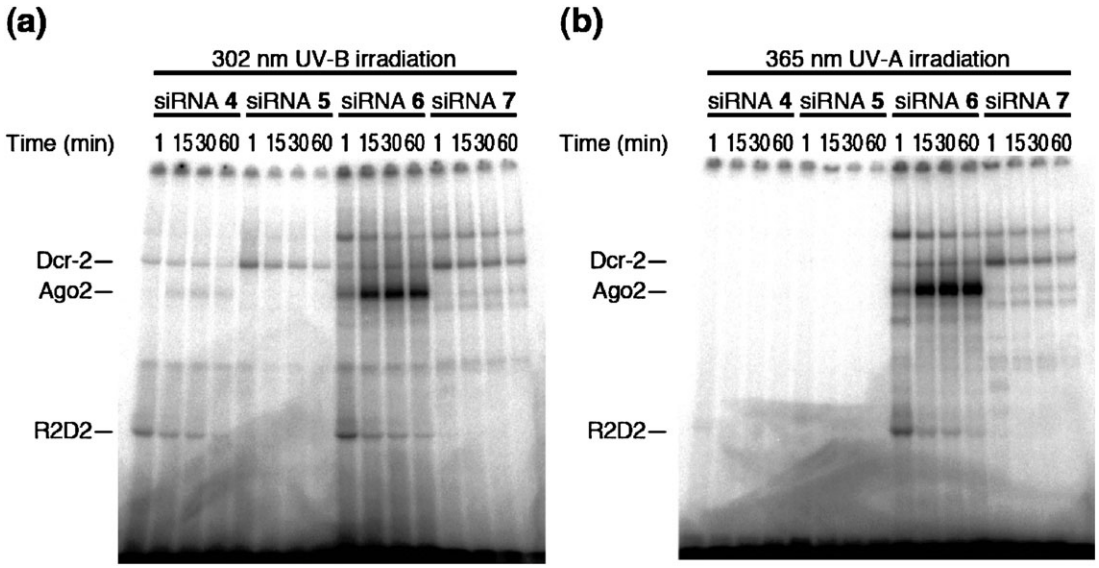


Fig. 4 Superior photocrosslinking efficiency of aryl diazirine compared to 5-iodouridine. (a) Photocrosslinking of the siRNAs to RISC assembly factors with 302 nm UV-B irradiation. (b) Photocrosslinking of the siRNAs to RISC assembly factors with 365 nm UV-A irradiation.

siRNAs **4** and **5** contain the conventional 5-iodouridine at position 20 in the 3′-overhang region of the upper strand as a control, whereas siRNAs **6** and **7** contain the aryl diazirine **2** at position 21 of the upper strand (shown in red) (Fig. 2). siRNAs **4** and **6** have a mismatch at the 5′ end of their respective upper strands; therefore, the crosslinker-containing strand serves as the guide in mature RISC. In contrast, siRNAs **5** and **7** have a mismatch at the 5′ end of their respective lower strands (shown in blue); therefore, the crosslinker-containing strand is discarded during unwinding. In the case of all the siRNAs, we radiolabeled the 5′ ends of the upper strands with ^{32}P, assembled RISC in a *Drosophila* embryo lysate for the indicated time, irradiated the lysate with 302 nm UV-B or 365 nm UV-A, and separated the crosslinked proteins by sodium dodecyl sulfate (SDS)-PAGE.

The results of photocrosslinking are shown in Fig. 4. With 302 nm UV-B irradiation, the crosslinking efficiency was much higher for the siRNAs with aryl diazirine (siRNAs **6** and **7**) than for the siRNAs with 5-iodouridine (siRNAs **4** and **5**) (Fig. 4a). The crosslinking patterns of all the siRNAs faithfully reflected asymmetric binding of the siRNA duplexes to Dcr-2/R2D2 in RLC and transition from Dcr-2/R2D2 in RLC to Ago2 in RISC; siRNAs **4** and **6** were specifically crosslinked initially with R2D2 and subsequently with Ago2, whereas siRNAs **5** and **7** were crosslinked with Dcr-2 but not with Ago2. Most strikingly, with 365 nm UV-A irradiation, we detected robust crosslinking for the siRNA with aryl diazirine, but no crosslinking for the siRNA with 5-iodouridine (Fig. 4b).

In conclusion, we have demonstrated the synthesis of siRNAs carrying the photoactivatable residues in their 3′-overhang regions. It was revealed that, compared to the conventional 5-iodouridine photocrosslinker, diazirine-containing siRNAs allowed much more sensitive detection of interacting proteins by minimally destructive long-wavelength UV-A irradiation. Therefore diazirine-containing siRNAs will be a useful tool to investigate the mechanism of RISC assembly of not only fly Ago2 but also other Ago proteins, which is still poorly understood.

This work was in part supported by a grant from Precursory Research for Embryonic Science and Technology (PRESTO) of the Japan Science and Technology Agency.

Notes and references

1 A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello, *Nature*, 1998, **391**, 806–811.
2 S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl, *Nature*, 2001, **411**, 494–498.
3 T. Kawamata and Y. Tomari, *Trends Biochem. Sci.*, 2010, **35**, 368–376.
4 Q. Liu, T. A. Rand, S. Kalidas, F. Du, H. E. Kim, D. P. Smith and X. Wang, *Science*, 2003, **301**, 1921–1925.
5 J. W. Pham, J. L. Pellino, Y. S. Lee, R. W. Carthew and E. J. Sontheimer, *Cell*, 2004, **117**, 83–94.
6 Y. Tomari, T. Du, B. Haley, D. S. Schwarz, R. Bennett, H. A. Cook, B. S. Koppetsch, W. E. Theurkauf and P. D. Zamore, *Cell*, 2004, **116**, 831–841.
7 Y. Tomari, C. Matranga, B. Haley, N. Martinez and P. D. Zamore, *Science*, 2004, **306**, 1377–1380.
8 S. Iwasaki, M. Kobayashi, M. Yoda, Y. Sakaguchi, S. Katsuma, T. Suzuki and Y. Tomari, *Mol. Cell*, 2010, **39**, 292–299.
9 C. Matranga, Y. Tomari, C. Shin, D. P. Bartel and P. D. Zamore, *Cell*, 2005, **123**, 607–620.
10 T. A. Rand, S. Petersen, F. Du and X. Wang, *Cell*, 2005, **123**, 621–629.
11 K. Miyoshi, H. Tsukumo, T. Nagami, H. Siomi and M. C. Siomi, *Genes Dev.*, 2005, **19**, 2837–2848.
12 E. F. V. Scriven and K. Turnbull, *Chem. Rev.*, 1988, **88**, 297–368.
13 (*a*) R. A. G. Smith and J. R. Knowles, *J. Am. Chem. Soc.*, 1973, **95**, 5072–5073; (*b*) T. Yamaguchi, K. Suyama, K. Narita, S. Kohgo, A. Tomikawa and M. Saneyoshi, *Nucleic Acids Res.*, 1997, **25**, 2352–2358; (*c*) T. Hiramatsu, Y. Guo and T. Hosoya, *Org. Biomol. Chem.*, 2007, **5**, 2916–2919; (*d*) U. K. Shigdel, J. Zhang and C. He, *Angew. Chem., Int. Ed.*, 2008, **47**, 90–93; (*e*) Z. Qiu, L. Lu and C. He, *J. Am. Chem. Soc.*, 2008, **130**, 14398–14399.
14 M. Winnacker, S. Breeger, R. Strasser and T. Carell, *ChemBioChem*, 2009, **10**, 109–118.
15 Y. Ueno, Y. Watanabe, A. Shibata, K. Yoshikawa, T. Takano, M. Kohara and Y. Kitade, *Bioorg. Med. Chem.*, 2009, **17**, 1974–1981.
16 Y. Ueno, A. Kawamura, K. Takasu, S. Komatsuzaki, T. Kato, S. Kuboe, Y. Kitamura and Y. Kitade, *Org. Biomol. Chem.*, 2009, **7**, 2761–2769.