Cell-free reconstitution reveals the molecular mechanisms for the initiation of secondary siRNA biogenesis in plants

Yuriki Sakurai,a,b,1 Kyungmin Baeg,a,c, Andy Y. W. Lam,a,c, Keisuke Shoji,a,c, Yukihide Tomaria,b,2 and Hiro-oki Iwakawa,a,c,2

aInstitute for Quantitative Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan; bDepartment of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan; and cPrecursory Research for Embryonic Science and Development, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan.

Edited by R. Scott Poethig, University of Pennsylvania, Philadelphia, PA, and approved June 23, 2021 (received for review February 11, 2021)

Secondary small interfering RNA (siRNA) production, triggered by primary small RNA targeting, is critical for proper development and antiviral defense in many organisms. RNA-dependent RNA polymerase (RDR) is a key factor in this pathway. However, how RDR specifically converts the targets of primary small RNAs into double-stranded RNA (dsRNA) intermediates remains unclear. Here, we develop an in vitro system that allows for dissection of the molecular mechanisms underlying the production of trans-acting siRNAs, a class of plant secondary siRNAs that play roles in organ development and stress responses. We find that a combination of the dsRNA-binding protein, SUPPRESSOR OF GENE SILENCING3; the putative nuclear RNA export factor, SILENCING DEFECTIVE5, primary small RNA, and Argonaute is required for physical recruitment of RDR6 to target RNAs. dsRNA synthesis by RDR6 is greatly enhanced by the removal of the poly(A) tail, which can be achieved by the cleavage at a second small RNA-binding site bearing appropriate mismatches. Importantly, when the complementarity of the base pairing at the second target site is too strong, the small RNA–Argonaute complex remains at the cleavage site, thereby blocking the initiation of dsRNA synthesis by RDR6. Our data highlight the light and dark sides of double small RNA targeting in the secondary siRNA biogenesis.

RNASilencing | microRNA | secondary siRNA | trans-acting siRNA | phased siRNA

MicroRNAs (miRNAs) and small interfering RNAs (siRNAs) are critical for the regulation of a broad range of biological functions across the kingdoms of life. Such small RNAs are processed from hairpin RNAs or double-stranded RNAs (dsRNAs) by the RNase III Dicer or Dicer-like proteins (DCLs) (1, 2). Small RNAs cannot function alone and form effector ribonucleoprotein complexes called RNA-induced silencing complexes (RISCs) with Argonaute (AGO) proteins to exert their functions (1, 3). RISCs bind to target RNAs via base complementarity. When the central region of the base pairing between small RNA and the target site is complementary, RISCs cleave the target RNA through the intrinsic slicing activity of AGO. Even in the presence of central mismatches, RISCs can recruit additional regulatory factors, thereby inducing messenger RNA (mRNA) decay and/or translational repression (2, 4–6).

In addition, many organisms, including plants, fungi, and worms, harness a powerful mechanism that amplifies the initial RNA silencing signal, in which target RNAs of primary small RNAs trigger the production of secondary siRNAs (7). RNA-dependent RNA polymerase (RDR), which converts target RNAs into dsRNAs, is central to secondary siRNA production (1, 8–11). RDR-mediated dsRNA production and subsequent Dicer-mediated processing trigger the spreading of silencing from the initial cleavage site targeted by primary small RNA-loaded RISC toward flanking sequences that lie 5′ or 3′ in the target transcript. Production of secondary siRNAs is required for regulation of endogenous gene expression as well as defense against viruses and transposons (1, 2, 12). However, the molecular details of these critical silencing amplification steps are not well understood due to their complexity.

Phased siRNAs (phasiRNAs) are plant secondary siRNAs that regulate development and stress responses. PhasiRNA biogenesis is triggered by the recruitment of primary miRNA-loaded RISCs, the majority of which are 22-nt miRNA-loaded AGO1-RISCs, to the phasiRNA generating (PHAS) precursor transcripts (PHAS transcripts) (13–15). Trans-acting siRNAs (tasiRNAs) are a subtype of phasiRNAs. As the name suggests, tasiRNAs predominantly suppress mRNAs other than the original tasiRNA generating (TAS) precursor transcripts (TAS transcripts) in trans. After RISC-mediated cleavage or binding, RDR6 synthesizes the complementary strand of the precursor transcripts (16, 17). The resulting long dsRNA is then processed by DCLs into 21- or 24-nt secondary siRNAs, with the phase determined by the initial miRNA-guided cleavage site (15, 18, 19). In addition to these factors, two proteins, known as the dsRNA-binding protein, SUPPRESSOR OF GENE SILENCING3 (SGS3) and a putative RNA export protein, SILENCING DEFECTIVE5 (SDE5), are required for tasiRNA biogenesis (20–23). It has been reported that SGS3 forms membrane-associated cytoplasmic foci named siRNA bodies with

Significance

Double-stranded RNA (dsRNA) synthesis by RNA-dependent RNA polymerase (RDR) is a critical step in secondary small interfering RNA (siRNA) biogenesis. However, how RDR specifically converts the targets of primary small RNAs into dsRNA intermediates remains unclear. Here, we developed an in vitro system that recapitulates the production of secondary siRNAs that are physiologically important in plants. Leveraging this system, we showed that a combination of four plant factors promotes physical recruitment of RDR6 to the target RNA. Moreover, we found that dsRNA synthesis by RDR6 is enhanced by the removal of the poly(A) tail, which is achieved by cleavage at another small RNA-binding site bearing appropriate mismatches. Our data elucidate the molecular events necessary for secondary siRNA biogenesis in plants.

Author contributions: Y.S., K.B., and H.-o.I. designed research; Y.S., K.B., and A.Y.W.L. performed research; Y.S., K.B., A.Y.W.L., K.S., Y.T., and H.-o.I. analyzed data; Y.S., K.B., A.Y.W.L., Y.T., and H.-o.I. wrote the paper; and Y.T. and H.-o.I. supervised the project.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1Y.S. and K.B. contributed equally to this work.

2To whom correspondence may be addressed. Email: tomari@iqb.u-tokyo.ac.jp or iwakawa@iqb.u-tokyo.ac.jp.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2102889118/-/DCSupplemental.

Published July 30, 2021.
RDR6 and AGOT7, a special AGO protein able to trigger tasiRNA biogenesis by forming RISC with 21-nt miR390 (24–27). SGS3 is also known to physically interact with RISCs in the presence of precursor transcripts, including miR390-loaded AGOT7 (28) and 22-nt miR173-loaded AGO1-RISC (29). Although no physical interaction has been detected between SDE5 and RISCs, epistasis analysis places SDE5 functions between SGS3 and RDR6 during tasiRNA biogenesis (30). However, due to the lack of a biochemical framework that allows dissection of the complex phasi/tasiRNA pathway, the molecular mechanism of phasi/tasiRNA biogenesis driven by RISCs, SDE5, and SGS3 remains obscure.

PHAS/TAS loci are divided into two classes—dubbed “one hit” and “two hit”—based on the number of primary miRNA binding sites. Most PHAS/TAS transcripts, including TAS1/2 RNA in Arabidopsis, have a single 22-nt miRNA binding site (15). Although a few two-hit precursors have been predicted, only TAS3 is experimentally validated, which is evolutionarily conserved from moss to flowering plants and considered the origin of PHAS/TAS evolutionarily conserved mismatches at the 3′end of pol(U) accumulate in the small RNAs produced in our in vitro system and compared to the extract. (SI Appendix, Fig. S1A). We did not supplement DCLs since previous reports observed endogenous DCL activities in BY-2 lysate (46, 47). Supplementation of tasiRNA factors efficiently triggered tasiRNA production from cleaved TAS3 fragments (Fig. 1C and SI Appendix, Fig. S1C). To check the contribution of each supplemented factor, we systematically omitted miR390 or mRNAs corresponding to each tasiRNA factor, one at a time, from the reaction. Removal of TAS3 mRNA, AtAGOT7 mRNA, miR390 duplex, or AtSDE5 mRNA greatly compromised tasiRNA production (Fig. 1C and D and SI Appendix, Fig. S1 A and D), so we supplemented naïve BY-2 cell extract with all of these factors as we moved forward with further analysis.

Although we successfully detected two TAS3 tasiRNAs, 5′D7 (+) and 5′D8 (+) (Fig. 1C and SI Appendix, Fig. S1 C and D), it was still unclear whether the pattern of in vitro expressed tasiRNAs was similar to in vivo production. To ascertain this, we sequenced the small RNAs produced in our in vitro system and compared them to TAS3a tasiRNAs produced in wild-type Arabidopsis seedlings (Col-0) but lost in ss3 mutant plants (sgs3-11) (Fig. 1E). Small RNAs generated in vitro mapped between the two miR390 binding sites, but only in the presence of both miR390 duplex and AGOT7 (Fig. 1E), suggesting that the mapped small RNAs were bona fide miR390-AGO7-triggered tasiRNAs. Interestingly, although relatively high levels of 24-nt small RNAs mapped to the TAS3a RNA due to robust DCL3-like activity in BY-2 cell extracts (SI Appendix, Fig. S1E) (46, 47), the position of the most prominent 21-nt tasiRNA hotspot was common between in vivo and in vitro processing (Fig. 1E). Taken altogether, the system we developed accurately recapitulates TAS3 tasiRNA biogenesis in vitro. We also tested that this in vitro system also synthesizes dsRNAs and produces tasiRNAs from the one-hit precursor TAS1a that has a binding site for 22-nt miR173-AGO1 RISC (SI Appendix, Fig. S2).

Results

Reconstitution of the TAS3 tasiRNA Biogenesis Pathway In Vitro. We developed an in vitro system to dissect the plant secondary siRNA biogenesis pathway. We selected the TAS3 pathway as a model because this pathway is well conserved and essential for proper plant development (31, 40). We started with a tobacco Bright Yellow 2 (BY-2) cell lysate, specifically the 17,000 × g supernatant, which has been successfully used to study plant RISC assembly and function (29, 41–45). In order to check whether endogenous tasiRNA biogenesis factors in the naive tobacco cell extract are sufficient for secondary siRNA production, we added TAS3 RNA to the extract (SI Appendix, Fig. S1A). However, 5′D7 (+)—the most abundant TAS3 tasiRNA—was not detected by northern blotting (SI Appendix, Fig. S1A), suggesting a lack of tasiRNA factors in the naïve tobacco cell extract. Indeed, TAS3 mRNA remained intact without cleavage (SI Appendix, Fig. S1A), suggesting that this naïve extract contains insufficient levels of functional AGOT7-RISC. We therefore in vitro translated mRNAs carrying tasiRNA factors fused to Flag tags (AtAGOT7, AtSGS3, AtSDE5, and AtRDR6) in the cell-free system. We added miR390 duplexes, which are bound to AtAGOT7 to form RISC, before incubation with TAS3 RNA (Fig. 1B and C and SI Appendix, Fig. S1B). We did not supplement DCLs since previous reports observed endogenous DCL activities in BY-2 lysate (46, 47). Supplementation of tasiRNA factors efficiently triggered tasiRNA production from cleaved TAS3 fragments (Fig. 1C and SI Appendix, Fig. S1C). To check the contribution of each supplemented factor, we systematically omitted miR390 or mRNAs corresponding to each tasiRNA factor, one at a time, from the reaction. Removal of TAS3 mRNA, AtAGOT7 mRNA, miR390 duplex, or AtSDE5 mRNA greatly compromised tasiRNA production (Fig. 1C and D and SI Appendix, Fig. S1 A and D), so we supplemented naïve BY-2 cell extract with all of these factors as we moved forward with further analysis.

Although we successfully detected two TAS3 tasiRNAs, 5′D7 (+) and 5′D8 (+) (Fig. 1C and SI Appendix, Fig. S1 C and D), it was still unclear whether the pattern of in vitro expressed tasiRNAs was similar to in vivo production. To ascertain this, we sequenced the small RNAs produced in our in vitro system and compared them to TAS3a tasiRNAs produced in wild-type Arabidopsis seedlings (Col-0) but lost in ss3 mutant plants (sgs3-11) (Fig. 1E). Small RNAs generated in vitro mapped between the two miR390 binding sites, but only in the presence of both miR390 duplex and AGOT7 (Fig. 1E), suggesting that the mapped small RNAs were bona fide miR390-AGO7-triggered tasiRNAs. Interestingly, although relatively high levels of 24-nt small RNAs mapped to the TAS3a RNA due to robust DCL3-like activity in BY-2 cell extracts (SI Appendix, Fig. S1E) (46, 47), the position of the most prominent 21-nt tasiRNA hotspot was common between in vivo and in vitro processing (Fig. 1E). Taken altogether, the system we developed accurately recapitulates TAS3 tasiRNA biogenesis in vitro. We also tested that this in vitro system also synthesizes dsRNAs and produces tasiRNAs from the one-hit precursor TAS1a that has a binding site for 22-nt miR173-AGO1 RISC (SI Appendix, Fig. S2).

SD5 and SGS3 Function at the dsRNA Conversion Step. We next sought to investigate the exact steps at which SDE5 and SGS3 function in tasiRNA biogenesis. Given that the removal of SGS3 or SDE5 mRNA from the reaction had no effect on AGOT7-mediated target cleavage (Fig. 1C), we hypothesized that both factors function upstream of dsRNA synthesis. However, the antisense strand of TAS3 RNA was barely detected by northern blotting even in the presence of exogenous AGOT7-RISC and SDE5 (Fig. 2A), presumably due to rapid processing of dsRNAs into siRNAs by endogenous DCL activities. To stabilize dsRNA intermediates so that we could detect them and resolve the role of SGS3 and SDE5, we added Flock House virus (FHV) B2 protein, which is known to inhibit dsRNA processing by DCLs (48–50) (Fig. 2B). Addition of FHV B2 protein strongly inhibited tasiRNA processing, leading to accumulation of dsRNA intermediates (Fig. 2A and B), while having no effect on RISC-mediated target cleavage. The size of the antisense strand is in line with a previous report, having the sequence from the third nucleotide of the 3′ cleaved site to the one nucleotide upstream of the 3′ end of the 5′ miR390 binding site (Fig. 2B) (51). FHV B2 exposure in the absence of SDE5 supplementation showed decreased accumulation of dsRNA intermediates (Fig. 2B), indicating that SDE5 functions at the dsRNA conversion step. Since endogenous SGS3 in BY-2 cell extract sufficiently promoted tasiRNA production (Fig. 1 C and D), we immunodepleted endogenous
NiSGS3 using anti-NiSGS3 antibody conjugated beads (SI Appendix, Fig. S3A). In the resulting SGS3-depleted lysate, tasiRNA and dsRNA accumulation was dramatically decreased (Fig. 2C and SI Appendix, Fig. S3A). To further validate the importance of SGS3 protein in the dsRNA conversion step, we added recombinant AtSGS3 protein (28) into SGS3-depleted or mock-depleted lysates (Fig. 2C and SI Appendix, Fig. S3A). Supplementing recombinant SGS3 protein restored tasiRNA biogenesis as well as complementarity strand synthesis in the TAS3-depleted lysate (Fig. 2C and SI Appendix, Fig. S3A), indicating that SGS3 is required for complementarity strand synthesis of TAS3. In summary, we conclude that both SDE5 and SGS3 are required for the dsRNA synthesis step.

SGS3 and SDE5 Promote Physical Association between RDR6 and the Target RNA Bound to AGO7-RISC. The requirement for SGS3 and SDE5 at the dsRNA conversion step led us to hypothesize that SGS3 and SDE5, together with AGO7-RISC, form a large complex that recruits RDR6 to the target mRNA. To test this hypothesis, we first investigated physical interactions among SGS3, SDE5, AGO7-RISC, and RDR6. We in vitro translated the mRNAs carrying AtSDE5 with an N-terminal triple influenza hemagglutinin (HA) epitope tag (3×HA-AtSDE5) and 3×FLAG-AtAGO7 in the cell-free system and then performed communoprecipitation experiments using the anti-FLAG antibody (Fig. 2D). In the presence of both miR390 and TAS3 RNA, endogenous NiSGS3 and NiRDR6 were specifically communoprecipitated with AtAGO7 (Fig. 2E) but not with the negative control, 3×FLAG-β-galactosidase (LacZ) (SI Appendix, Fig. S3B). Similar results were obtained when NiSGS3 was used as the bait protein (Fig. 2F). Thus, AtAGO7, SGS3, and RDR6 are physically associated with each other when miR390 base pairs with TAS3 RNA, in line with previous in vivo data showing that SGS3, AGO7, and RDR6 were colocalized in siRNA bodies (26, 27).
served miR390 target sites (31) (Fig. 1) | SG3 and SDE5 are required for RDR6-mediated dsRNA conversion. (Fig. 2. RDR6 and AGO7-RISC. The 5′ spliceosomal RNA was used as a loading control. In the presence of SDE5, addition of FHV B2 leads to accumulation of the antisense strand of TAS3 mRNA. In contrast, TAS3 antisense strand was barely observed in the absence of SDE5. (b) Schematic illustrating FHV B2 function. FHV B2 blocks the dicing reaction, leading to accumulation of dsRNA intermediates generated by RDR6. (c) In vitro synthesis of the complementary strand of TAS3 RNA with SGS3-immunodepleted lysate in the presence of FHV B2 protein. TAS3 sense and antisense RNAs were detected by northern blotting with specific probes. U6 spliceosomal RNA was used as a loading control. SGS3 depletion leads to loss of dsRNA accumulation, suggesting that SGS3 is required for RDR6-mediated dsRNA conversion. See also SI Appendix, Fig. S3A. (d) Schematic illustrating coimmunoprecipitation assay using the anti-FLAG antibody. (E) Coimmunoprecipitation experiments with 3′-FLAG-AtAGO7 (F-AtAGO7). NtRDR6 and NtSGS3 were coimmunoprecipitated with F-AtAGO7, but only in the presence of both TAS3 mRNA and miR390 duplex. F-AtAGO7, 3′-HA-AtSDE5 (HA-AtSDE5), NtSGS3, and NtRDR6 were detected by Western blotting. α-tubulin was used as a loading control as well as a negative control for coimmunoprecipitation experiments. The asterisk indicates nonspecific bands. (F) Coimmunoprecipitation experiments with NtSGS3. NtRDR6 and F-AtAGO7 were coimmunoprecipitated with NtSGS3 but only in the presence of both TAS3 mRNA and miR390 duplex. F-AtAGO7, 3′-FLAG-AtSDE5 (F-AtSDE5), NtSGS3, and NtRDR6 were detected by Western blotting. α-tubulin was used as a loading control as well as a negative control for coimmunoprecipitation experiments. (G) Coimmunoprecipitation experiments with F-AtAGO7 in the NtSGS3 or mock-depleted lysate. NtRDR6 was coimmunoprecipitated by F-AtAGO7 but only in the presence of both TAS3 mRNA and miR390 duplex. F-AtAGO7, 3′-HA-AtSDE5, NtSGS3, and NtRDR6 were detected by Western blotting. α-tubulin was used as a loading control as well as a negative control for coimmunoprecipitation experiments. Asterisks indicate nonspecific bands.

Although the signals were faint, SDE5 also coimmunoprecipitated with SGS3 (Fig. 2F). To confirm the interaction between SDE5 and other tasiRNA factors, we used SDE5 as the bait protein. AGO7, SGS3, and RDR6 coimmunoprecipitated with SDE5, the efficiency of which was the highest in the presence of both miR390 and TAS3 RNA (SI Appendix, Fig. S3C). Thus, although the interactions are weak, SDE5 also physically associated with RDR6 and the TAS3miR390-AGO7-SGS3 complex. Importantly, when either SGS3 or SDE5 was depleted from the lysate, RDR6 did not coimmunoprecipitate with AGO7 even in the presence of both miR390 and TAS3 RNA, suggesting that both SGS3 and SDE5 are required for promoting the physical association between RDR6 and AGO7-RISC-bound TAS3 RNA (Fig. 2G).

The 5′ miR390-Binding Site of TAS3 Facilitates the Interaction between RDR6 and AGO7-RISC. TAS3 RNA carries two evolutionarily conserved miR390 target sites (31) (Fig. L4). As multiple miRNA binding sites often augment silencing activity (5), we investigated the contribution of the two miR390-binding sites on TAS3 to RDR6 recruitment. We prepared three TAS3 RNA variants with mutations in one or both miR390 binding sites (5m3W, 5W3m, and 5m3m) (Fig. 3A). These mutations were introduced into target sequence complementary to the seed sequence of miR390 (2 to 8 nt position from the 5′ end), a region that is critical for target binding (Fig. 3A). As expected, a TAS3 variant carrying mutations in both miR390 binding sites (5m3m) impaired interaction between NtRDR6 and AGO7-RISC (Fig. 3B). Similarly, no interaction was observed between NtRDR6 and AGO7-RISC when the 5′ miR390 binding site mutant (5m3W) was incubated with AGO7-RISC. In contrast, wild-type TAS3 (5W3W) and the 3′ miR390 binding site mutant (5W3m) retained interaction between NtRDR6 and AGO7-RISC (Fig. 3B). Thus, the 5′ miR390 binding site is specifically required for the interaction between RDR6 and AGO7-RISC.
Poly(A) Removal by the 3′ Site Cleavage Facilitates TAS3 Complementary Strand Synthesis. Since the 5′ but not the 3′ miR390 binding site, was necessary for RDR6 recruitment to TAS3 mRNA (5W3W and 5W3m), we asked if the 5′ miR390 binding site alone is sufficient for the production of long dsRNAs. Surprisingly, antisense TAS3 was detected only when both miR390 binding sites were wild type (5W3W) (Fig. 3C). This result indicates that the 3′ miR390 binding site has a crucial role for RDR6 activity after its recruitment to the AGO7-bound TAS3 mRNA. We wondered why synthesis of the antisense strand of 5W3m is impaired, even though this mutant still recruits RDR6. We previously reported that the poly(A) tail blocks recombinant RDR6-mediated complementary strand RNA synthesis in vitro (17). We therefore hypothesized that removal of the poly(A) tail through cleavage of the 3′ miR390 binding site enhances RDR6 activity. If so, replacement of the poly(A)-tail of 5W3m with a non-poly(A) sequence should support complementary strand synthesis (Fig. 3D). Indeed, replacing the poly(A) tail with 140-nt of non-poly(A) sequence promoted complementary strand synthesis (Fig. 3D). Since RDR6 recruitment was similar between polyadenylated and non-polyadenylated 5W3m (SI Appendix, Fig. S4), the poly(A) tail does not impair RDR6 recruitment but rather blocks the initiation of complementary strand synthesis. Taken together, we conclude that the 3′ miR390 binding site is important for not only determination of the starting position of tasiRNA phasing (34) but also the removal of the poly(A) tail to enhance dsRNA synthesis by RDR6. We also observed a similar enhancement of complementary strand synthesis by poly(A) removal in the TAS1 pathway (SI Appendix, Fig. S2 B and D).

The Mismatches in the 3′ miR390-Binding Site Are Essential to Initiate dsRNA Synthesis by Promoting Rapid Release of AGO7-RISC from the Cleavage Site. Arabidopsis TAS3 RNA has mismatches in both 5′ and 3′ miR390-binding sites. The 5′ miR390-binding site has central...
mismatches in the base pairing with miR390, while the 3’ miR390-binding site has conserved mismatches in the base pairing with the 3’ end of miR390 (Fig. 4A). To examine the importance of these mismatches, we introduced compensatory mutations in the 5’ and 3’ miR390 binding sites that allow full Watson–Crick base pairing with the central and the 3’ end region of miR390, respectively (5P3W and 5W3P; Fig. 4A). Although the introduction of compensatory mutations into the 5’ miR390 binding site induced cleavage at the 5’ site in the sense strand, both antisense RNAs with the sequence between the two binding sites and tasiRNAs were efficiently generated from 5P3W (Fig. 4B). Thus, the central mismatches in the 5’ miR390 binding site are dispensable for the initiation of dsRNA synthesis. These results are in line with previous transient expression assays with Nicotiana benthamiana and the fact that some plant species, including moss Physcomitrella patens, have the cleavable 5’ miR390 binding site (33, 34). Note that an additional band was observed above the ~234 nt antisense strand in the reaction mixture containing 5P3W (Fig. 4B). This may be the antisense RNA that is produced when RDR6 reaches the 5’ end of the cleavage fragment after AGO7-RISC leaves the cleaved 5’ site in 5P3W.

In contrast to the central mismatches in the 5’ site, the conserved mismatches in the base pairing between the 3’ end of miR390 and the 5’ miR390-binding site was critical for both antisense strand synthesis and tasiRNA biosynthesis (5W3P; Fig. 4C). Given that the 3’ site of 5W3P was efficiently cleaved by the miR390-AGO7-RISC just like wild-type TAS3 RNA (5W3W) (Fig. 4C), we speculate that the additional base pairings compromise a downstream step(s). To test if the RDR6 recruitment is impaired by this mutation, we performed communoprecipitation with AGO7 as a bait. This experiment showed that RDR6 can associate with 5W3P in the presence of miR390-AGO7 RISC (Fig. 4D), suggesting that the extensive base pairing at the 3’ miR390-binding site compromises the dsRNA conversion step (rather than recruitment of RDR6).

Interestingly, SG35 was more efficiently pulled down in the presence of 5W3P than 5W3W (Fig. 4D). This suggests that SG35 is recruited not only to the 5’ miR390 binding site but also to the 3’ miR390 binding site in 5W3P. Indeed, our recent study showed that the extensive base pairing between the 3’ end region of miR390 and the 5’ binding site promotes stable binding of SG3G to the target site (28). Thus, we hypothesized that the SG35-miR390-AGO7 complex can unmask the 3’ end of the cleavage site of 5W3P, thereby blocking the initiation of dsRNA synthesis by RDR6. To test this idea, we sought to observe the interaction between the SG35-miR390-AGO7 complex and the 3’ miR390-binding sites of 5W3P by RNA immunoprecipitation analysis (Fig. 5A). As the 5’ miR390 binding site of TAS3 by itself can bind to SG35 and AGO7-RISC, we introduced mutations in the seed region of the 5’ site of 5W3P (5m3P; Fig. 5A). After AGO7-RISC assembly in BY-2 lysate, cap-labeled TAS3 mutants (5m3W, 5m3m, or 5m3P) were incubated in the reaction. Then, SG35 and FLAG-tagged AGO7 were pulled down with anti-NiSG35 and anti-FLAG antibodies, respectively, and the communoprecipitated cap-labeled TAS3 fragments were detected (Fig. 5B). As expected, only a negligible amount of 5m3m was pulled down after immunoprecipitation (Fig. 5C). Importantly, the 5’ cleaved fragment of 5m3P was more efficiently communoprecipitated than that of 5m3W by both anti-NiSG35 and anti-FLAG antibodies (Fig. 5C). Thus, the additional 4-nucleotide base pairings promote the association between the 3’ miR390 cleavage site and the SGS3-miR390-AGO7 complex. Although a previous study showed that SG35 promotes association of miR173-AGO1-RISC with the cleavage fragments of TAS2 RNA (29), it is unclear if the same is true for the binding between miR390-AGO7-RISC and the cleavage fragments of 5m3P. When we predicted the free energies of the heterodimers between the 3’ half of miRNAs and the 5’ half of the binding sites (ΔGbinding), we found that the base pairing between miR390 and the 3’ end of the cleavage fragment of 5m3P is more stable (ΔGbinding = −24.19 kcal/mol; Fig. 5A) than the corresponding base pairings between miR173 and TAS1/2 (ΔGbinding = −16.1 kcal/mol; SI Appendix, Fig. S5) or between miR390 and the wild-type 3’ binding site (ΔGbinding = −12.77 kcal/mol; Fig. 5A). Thus, it is possible that the guanine-cytosine (GC)-rich base pairing in the 3’ binding site in 5m3P is sufficient for the stable binding of AGO7-RISC to the 3’ end of the cleavage fragment without the help of SG35 protein. To investigate whether SG35 recruitment enhances AGO7-RISC binding to the 3’ miR390 cleavage site of 5m3P, we performed coimmunoprecipitation experiments using the NtSG35-depleted lysate with supplementation of different concentrations of the recombinant AtSG35 proteins. The binding between AGO7-RISC and 5m3P was not promoted by the addition of AtSG35 protein (Fig. 5D). Thus, the GC-rich base pairing can retain AGO7-RISC at the 3’ end of the cleavage fragment for extended periods of time even without SG35. Taken all together, the evolutionarily conserved mismatches in the 3’ miR390 binding site are essential to initiate dsRNA synthesis by promoting destabilization and rapid release of AGO7-RISC from the cleavage site.

**Discussion**

Here, we developed an in vitro system that successfully recapitulates the widely conserved plant endogenous secondary siRNA biogenesis pathway and allows its mechanistic dissection. Using this system, we demonstrated that a dsRNA-binding protein, SG3S, and a putative homolog of a human mRNA export factor, SDE5, facilitate recruitment of RDR6 to the primary mRNA bound precursor RNA. Furthermore, we found that removal of the poly(A)-tail enhances initiation of RDR6-mediated dsRNA synthesis but not the recruitment of RDR6. Moreover, we revealed that the evolutionarily conserved mismatches between the 3’ end of miR390 and the 3’ miR390-binding site are essential to initiate dsRNA synthesis by promoting rapid release of AGO7-RISC from the cleavage site.

Our reconstituted tasiRNA biogenesis system uses the 17,000 × g supernatant of tobacco cell extract (Fig. 1), in which most of the membrane fraction has been removed. This suggests that strict subcellular localization of tasiRNA factors is not absolutely required for small RNA amplification and that this reaction cascade, including targeting by AGO71-RISC, dsRNA conversion by RDR6, and dicing by DCLs, can occur in soluble fractions. However, AGO71 is still present at the 3’ end of the cleavage fragment even if some microsomes remain in the 17,000 × g supernatant and act as scaffolds for the formation of the siRNA body-like structures in vitro. Future studies are warranted to examine the interplay between membranes, siRNA body formation, and tasiRNA production in our in vitro system.

Prior work shows that SG3S plays a role in stabilizing the 5’ and 3’ TAS1/2 fragments generated by miR173-directed cleavage (29). Additionally, by binding to either the miR390-AGO7 or miR173-AGO1 complexes, SG35 stalls ribosomes on TAS3 or TAS1/2, promoting tasiRNA production (28). In this study, we identified a critical role for SG3S as essential for RDR6 recruitment to TAS precursors (Figs. 2G and 5). Together, the multifunctional roles of SG3S maximize secondary siRNA production from TAS precursors. In addition to SG3S, we also revealed a role for SDE5 in recruitment of RDR6 (Figs. 2G and 5). Communoprecipitation assays suggested that SDE5 transiently interacts with AGO7, SG3S, and RDR6 (SI Appendix, Fig. S3C) and acts as a connector between these factors. Previous reports demonstrated that SG3S, RDR6, and AGO7 are colocalized in siRNA bodies (26, 27). In addition, a recent study showed that SG3S forms droplets with liquid-like properties (52). Based on these results, we propose a model for RDR6 recruitment to the precursor RNAs of phasi/tasiRNAs. First, the precursor RNA bound to the initiator RISC—22-nt small RNA-loaded AGO1 or miR390-AGO7—interacts with SG35s, which form liquid droplets. Second, SDE5 pulls RDR6 into the droplets by transiently interacting with AGO1/7, SG3S, and RDR6. Third, the locally concentrated RDR6s recognize
the 3’ end of the precursor RNAs (Fig. 6). The siRNA bodies formed by SGS3, special RISC, and SDE5 might also activate RDR6 and/or protect the synthesized dsRNAs from unwanted nucleases. Future studies combining our in vitro system and imaging approaches will test this model.

It is known that TAS3 tasiRNAs have specific “hotspots” of increased accumulation between the two miR390 binding sites in Arabidopsis (Fig. 1E) (33). Our in vitro system mimics this pattern of tasiRNA biogenesis (Fig. 1E and SI Appendix, Fig. S1E). Interestingly, the most prominent hotspot overlaps with 5’D7(+) — the most physiologically important TAS3 tasiRNA — both in vivo and in vitro (Fig. 1E). This suggests that the biologically important small RNAs produced from this region are selectively stabilized by an unknown mechanism. The most plausible explanation for this is that endogenous AGO proteins specifically load a subset of siRNAs, protecting them from degradation by nucleases. The hotspot contains four consecutive U nucleotides, whereas 18 to 21 nt downstream are four consecutive Cs. These features are ideal for AGO1-RISC assembly; plant AGO1 has a binding preference for 5’U small RNAs (53, 54). In addition, AGO proteins generally preferentially select the small RNA strand with lower thermodynamic stability at the 5’ end (low-GC content) (55). Thus, the unique sequence surrounding the hotspot may enhance selective loading and stabilization by AGO1.

Our in vitro tasiRNA production system demonstrates that a nonpolyadenylated 3’ end is preferred as an initiation site for dsRNA conversion rather than a polyadenylated one (Fig. 3 and SI Appendix, Fig. S2). This can be explained by the intrinsic properties of RDR6; recombinant RDR6 selects poly(A)-less mRNAs over polyadenylated mRNAs as templates at the initiation step of complementary strand RNA synthesis (17). We envision that the 3’ site cleavage (which effectively removes the poly(A) tail) in the TAS3 pathway creates a non-poly(A) 3’ end, thus enhancing tasiRNA production (Fig. 5).

A previous study has suggested that the 3’ miR390 site is essential for dsRNA synthesis. (A, Left) A schematic illustrating wild-type TAS3 and the TAS3 variants with compensatory mutations in the 5’ or 3’ binding site (SP3W and SW3P). The base-pairing configurations between miR390 (red) and the wild-type miR390 binding sites (orange) or the binding sites with compensatory mutations (blue). (B) In vitro tasiRNA biogenesis with wild-type TAS3 (5SW3W or SP3W in the presence or absence of FHV B2. The compensatory mutations in the 5’ miR390 binding site did not affect tasiRNA production. The single asterisk indicates a degradation intermediate of wild-type TAS3 only detected in the presence of AGO7-RISC. The double asterisk indicates the longer antisense TAS3 RNA specifically detected when the 5’ miR390 binding was cleaved by AGO7-RISC. TAS3 sense RNA, TAS3 antisense RNA, and 5’D7 (+) tasiRNA were detected by northern blotting with specific probes. U6 spliceosomal RNA was used as a loading control. (C) In vitro tasiRNA biogenesis with wild-type TAS3 (5SW3W or SP3W in the presence or absence of FHV B2. While both SW3W and SW3P were efficiently cleaved by the miR390-AGO7-RISC, neither the complementary strand nor 5’D7 (+) tasiRNA were produced from TAS3 5P3W target RNA. TAS3 sense RNA, TAS3 antisense RNA, and 5’D7 (+) tasiRNA were detected by northern blotting with specific probes. U6 spliceosomal RNA was used as a loading control. (D) Coimmunoprecipitation experiments with F-AtAGO7 in the presence of wild-type TAS3 (5SW3W) or its variant with compensatory mutations in the 3’ miR390 binding site (SW3P) shown in A. F-AtAGO7 was immunoprecipitated (IP) using anti-FLAG antibody, and blotted with anti-FLAG, anti-SGS3, anti-HA, anti-NtRDR6, or anti-tubulin antibodies. α-tubulin was used as a loading control as well as a negative control for coimmunoprecipitation experiments.

Fig. 4. The mismatches in the 3’ miR390 binding site are essential for dsRNA synthesis. (A, Left) A schematic illustrating wild-type TAS3 and the TAS3 variants with compensatory mutations in the 5’ or 3’ binding site (SP3W and SW3P). The base-pairing configurations between miR390 (red) and the wild-type miR390 binding sites (orange) or the binding sites with compensatory mutations (blue). (B) In vitro tasiRNA biogenesis with wild-type TAS3 (5SW3W or SP3W in the presence or absence of FHV B2. The compensatory mutations in the 5’ miR390 binding site did not affect tasiRNA production. The single asterisk indicates a degradation intermediate of wild-type TAS3 only detected in the presence of AGO7-RISC. The double asterisk indicates the longer antisense TAS3 RNA specifically detected when the 5’ miR390 binding was cleaved by AGO7-RISC. TAS3 sense RNA, TAS3 antisense RNA, and 5’D7 (+) tasiRNA were detected by northern blotting with specific probes. U6 spliceosomal RNA was used as a loading control. (C) In vitro tasiRNA biogenesis with wild-type TAS3 (5SW3W or SP3W in the presence or absence of FHV B2. While both SW3W and SW3P were efficiently cleaved by the miR390-AGO7-RISC, neither the complementary strand nor 5’D7 (+) tasiRNA were produced from TAS3 5P3W target RNA. TAS3 sense RNA, TAS3 antisense RNA, and 5’D7 (+) tasiRNA were detected by northern blotting with specific probes. U6 spliceosomal RNA was used as a loading control. (D) Coimmunoprecipitation experiments with F-AtAGO7 in the presence of wild-type TAS3 (5SW3W) or its variant with compensatory mutations in the 3’ miR390 binding site (SW3P) shown in A. F-AtAGO7 was immunoprecipitated (IP) using anti-FLAG antibody, and blotted with anti-FLAG, anti-SGS3, anti-HA, anti-NtRDR6, and abundant NtSGS3 were coimmunoprecipitated with F-AtAGO7 in the presence of SW3P in a manner dependent on the miR390 duplex. The asterisk indicates nonspecific bands. α-tubulin was used as a loading control as well as a negative control for coimmunoprecipitation experiments.
for setting the proper phasing register (34). Thus, the cleavable 3′ miR390 site has a dual role in determination of the initiation site for phase of tasiRNAs and promotion of dsRNA synthesis, both of which maximize the production of functional tasiRNAs. Although poly(A) removal also enhances dsRNA synthesis in the one-hit TAS1 pathway (SI Appendix, Fig. S2), detectable levels of dsRNAs and tasiRNAs were produced when polyadenylated TAS1 was added into the cell-free system (SI Appendix, Fig. S2). This suggests a two-step production mechanism of tasiRNAs from TAS1; low levels of tasiRNAs are first produced from polyadenylated TAS1 RNA, then the 21-nt tasiRNAs cleave TAS1 in cis to generate non-poly(A) template that efficiently produce tasiRNAs (Fig. 5). Indeed, it is known that some tasiRNAs, including 3′ miR390* duplex was added and further incubated at 25 °C for 1.5 h. The re- action mixture was added to the target cleavage site after target cleavage. (D) RNA coimmunoprecipitation experiments with F-AtAGO7 in the NtSGS3-depleted lysate. The supplementation of AtSGS3 protein did not promote the interaction between AGO7-RISC and the 5′ cleavage fragment of 5m3P.

Materials and Methods

General Methods. Preparation of lysate from tobacco BY-2 cells, substrate mixture (containing adenosine triphosphate [ATP], ATP-regeneration system, amino acid mixtures), 1× lysis buffer (30 mM 2-(4-(2-hydroxyethyl)piperazin-1-yl)ethane-sulfonic acid [Hepes] at pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate), and small RNA duplexes have been previously described in detail (43). Anti-N-RRDR6 was raised in rabbits using synthetic peptides (NH2-CLGPEN-PYRLNQRRRTTM-COOH) as antigens (Medical and Biological Laboratories).

In Vitro tasiRNA Biogenesis Assay. Basically, 10 μL of BY-2 lysate, 5 μL of substrate mixture, and 0.5 μL each of 1 μM F-AtAGO7 mRNA and 1 μM F-AtDSDE5 mRNA were mixed and incubated at 25 °C. After 30 min, 1 μL of 1 μM miR390/ miR390* duplex was added and further incubated at 25 °C for 1.5 h. The re- action mixture and incubated at 25 °C for 10 min with or without 1 μL of 1 μM viral silencing suppressor FHV-B2 recombinant protein, purified as previously described (56). Next, 1 μL of 40 nM TAS3 RNAs was added to the reaction mixture and incubated at 25 °C for 30 min. Aliquots of the reaction mixtures were used for Western blotting by adding sodium dodecyl sulfate (SDS) protein sample buffer (40% glycerol, 240 mM Tris·HCl pH 6.8, 8% SDS, 0.04% bromophenol blue, and dithiothreitol), and for northern blotting sequences are required for the secondary cleavage site to trigger tasi/phasiRNA biogenesis; otherwise, the strong base pairing between small RNAs and the end of the cleavage fragment retains RISCs at the 3′ end for extended periods of time, resulting in inhibition of dsRNA synthesis by RDR6.

Taken together, we have revealed key mechanisms underlying phasi/tasiRNA biogenesis. Our in vitro system will pave the way for understanding the molecular bases of RNA silencing amplification, including not only phasi/tasiRNA biogenesis but also transgene-triggered and virus-triggered gene silencing in plants.
after deproteinization with proteinase K (Nacalai Tesque, Inc.) and ethanol precipitation followed by dissolving the pellet in formamide (40% formamide, 5 mM ethylenediamine tetraacetic acid [EDTA], 0.01% xylene cyanol, and 0.01% bromophenol blue). For Fig. 1 C and D and SI Appendix, Fig. S1B, 12 μL of BY-2 lysate, 6 μL of substrate mixture, 0.6 μL of 1 μM F-AgTAS6 mRNA, 0.6 μL of 1 μM F-AgTAS5 mRNA, 0.5 μL of 0.25 μM F-AgTAS3 mRNA, and 1 μL of 0.5 μM F-AgrDR6 mRNA were used.

Commmunoprecipitation Assay. The anti-FLAG antibody (Sigma) was immobilized using Dynabeads protein G (Invitrogen) in 1X lysis buffer and rotated at 4 °C for 1 h. After immobilization, the beads were washed 3 times with lysis wash buffer containing 200 mM NaCl and 1% Triton X-100 (polyoxyethylene (10) octylphenyl ether, Wako). Afterward, the tasiRNA biogenesis assay reaction mixture was added to the beads and rotated at 4 °C for 1 h. The beads were washed with lysis wash buffer three times, and finally, proteins were resuspended in SDS protein sample buffer. For SG33 immunoprecipitation (Fig. 2F), the anti-NSG33 antibody (29) was added directly to the tasiRNA biogenesis reaction mixture and incubated for 30 min on ice. After incubation, Dynabeads protein G (Invitrogen), washed with lysis wash buffer, was added to the reaction mixture and rotated at 4 °C for 30 min. The wash step was performed as outlined. For RNA immunoprecipitation, cap-labeled TAS3 target RNAs were used (Fig. 4D). After wash step, the beads were incubated with 1X lysis buffer containing proteinase K (200 ng/μL) at 55 °C for 10 min. Finally, the samples were resuspended in the same volume of formamide dye (10 mM EDTA pH 8.0, 98% [v/vol] deionized formamide, 0.025% [v/vol] xylene cyanol, and 0.025% bromophenol blue). After incubation at 95 °C for 2 min, the communoprecipitated RNAs were separated on an 8% polyacrylamide denaturing gel.

Immunodepletion of Endogenous SG33. Immunodepletion of NSG33 from BY-2 lysate was performed as previously described (28).

Western Blot Assay. The reactions, which were resuspended in 1X SDS protein sample buffer, were heated at 95 °C for 5 min. After electrophoresis, the proteins were transferred onto polyvinylidene difluoride membrane (Merck) via a semidry transfer method. The membrane was incubated with blocking buffer (1% skim milk and 1X Tris buffered saline with Tween 20 [TBS-T]) at room temperature for 1 h and supplemented with anti-FLAG antibodies (Sigma, #F1804; 1:10,000 dilution) or anti-HA antibodies (abcam, #ab130275; 1:10,000 dilution), anti-tubulin antibodies (Sigma, #T6074; 1:10,000 dilution), anti-NSG33 antibodies (29) (1:10,000 dilution), and anti-NtRDR6 antibodies (1:10,000 dilution) at 4 °C overnight. The membrane was washed with 1X TBS-T and then incubated with blocking buffer (1% skim milk and 1X TBS-T) with anti-mouse immunoglobulin G (IgG)-horseradish peroxidase (HRP) antibodies (MBL, #330; 1:25,000 dilution) or anti-rabbit IgG-HRP antibodies (Jackson, #111-035-003; 1:20,000 dilution) at room temperature for 1 h. After washing with 1X TBS-T, the membrane was treated with Immobilon Western Chemiluminescent HRP Substrate (Millipore). Images were acquired using the Imager 600 (Amersham) or FUSION FX (VILBER).

Preparation of Probes for Northern Blot Assay. DNA oligo probes (SI Appendix, Table S4) were phosphorylated using T4 polynucleotide kinase (TaKaRa) with [γ-32P]-ATP. Multiple DNA oligo (SI Appendix, Table S4) were mixed and radiolabeled as described for the detection of sense, antisense TAS3a, and antisense TAS1a.

Northern Blot Assay. To detect U6 splicedosomal RNA and tasiRNAs, RNAs were separated on an 18% polyacrylamide denaturing gel and transferred onto a Hybond-N membrane (GE Healthcare) using semidry transfer method. The heteroduplex of the 3′ half of tasiRNAs and the 5′ half of these binding sites are as weak as that of the 3′ half of mir390 and the 5′ half of the 3′ binding site in wild-type TAS3, tasiRNA-AgO1 complex can rapidly dissociate from the cleavage site.

and antisense TAS1a.

| siRNA library prep | Number of reads | Data analysis |
|-------------------|----------------|--------------|
| NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB) and analyzed by Illumina HiSEq 4000. | 2 lysate was performed as previously described (28). | Small RNA-seq. Total RNA was extracted from 3-d-old Col-0 and sgs3-17 (21) Arabidopsis seedlings and by 2X-lyase with TRizol reagent. Followed by gel purification, 19 to 28 nt length small RNAs were gel extracted from total RNAs. Small RNA libraries were prepared by NEBNext Multiplex Small RNA Library Prep Set for Illumina (NEB) and analyzed by Illumina HiSEq 4000. | Sequence Analysis of tasiRNAs. After removal of adaptor sequences by cutadapt (58), small RNA sequence reads of 20 to 29 nt length were mapped to the TAS3a sequence used in the tasiRNA biogenesis assay using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx-toolkit/) and Bowtie (59) allowing for up to one mismatch. Sequence Alignment Map (SAM) files were converted to BAM files using SAMtools (60) and then to BED files with BEDTools (61). Processed data were transferred to R, and length distribution and the 5′-end position for each siRNA were mapped onto the TAS3 mRNA using the ggaplot2 package. |
Prediction of the Free Energies of Heterodimers. The free energies of heterodimers of the 3′ half of small RNAAs and the 5′ half of binding sites were predicted at 25 °C via the Vienna RNA Webserver (2B).

Data Availability. The sequencing data reported in this paper are privately available in DNA Data Bank of Japan (DDBJ; http://www.ddbj.nig.ac.jp/) under the accession numbers DRA009601 for in vitro data and DRA009602 for in vivo data.

Acknowledgments. We thank Manabu Yoshikawa for sharing unpublished data and providing anti-NTGS3 and anti-β-NTRD6, Mariko Watanabe for production of FHV B2 protein, Kaori Kiyokawa for experimental assistance, and Life Science Editors for editorial assistance. We also thank all the members of the I.T. laboratory for helpful discussions and critical comments. This work was supported in part by JST, PRESTO (Grant JPMJPR1822 to H.-o.I.), Grant-in-Aid for Young Scientists (A) (Grant 16H06159 to H.-o.I.), Grant-in-Aid for Scientific Research on Innovative Areas (“Nascent-chain Biology”) (Grant 18H01603 to H.-o.I.), Grant-in-Aid for Challenging Exploratory Research (Grant 15K14444 to H.-o.I.), Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) Fellow 16J07290 (to K.B.), and Grant-in-Aid for JSPS Fellows 20J11529 (to Y.S.). DNA libraries were sequenced by the Vincent J. Coates Genomics Sequencing Laboratory at University of California, Berkeley, supported by an NIH 510 S01B174 Instrumentation Grant.