*RNA* (2001), *7*:1522–1530. Cambridge University Press. Printed in the USA.
Copyright © 2001 RNA Society.
DOI: 10.1017.S1355838201014522

# RNA interference in *Trypanosoma brucei*: Cloning of small interfering RNAs provides evidence for retroposon-derived 24–26-nucleotide RNAs

**APPOLINAIRE DJIKENG,[1] HUAFANG SHI,[1] CHRISTIAN TSCHUDI,[1] and ELISABETTA ULLU[1,2]**

[1]Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06520-8022, USA
[2]Department of Cell Biology, Yale University School of Medicine, New Haven, Connecticut 06520-8022, USA

## ABSTRACT

**In animals and protozoa, gene-specific double-stranded RNA (dsRNA) triggers degradation of homologous cellular RNAs, a phenomenon known as RNA interference (RNAi). In vitro and in vivo dsRNA is processed by a nuclease to produce 21–25-nt small interfering RNAs (siRNAs) that guide target RNA degradation. Here we show that activation of RNAi in *Trypanosoma brucei* by expression or electroporation of actin dsRNA results in production of actin siRNAs and that 10% of these RNAs sediment as high-molecular-weight complexes at 100,000 × *g*. To characterize actin siRNAs, we established a cloning and enrichment strategy starting from 20–30 nt RNAs isolated from high-speed pellet and supernatant fractions. Sequence analysis revealed that actin siRNAs are 24–26 nt long and their distribution relative to actin dsRNA was similar in the two fractions. By sequencing over 1,300 fragments derived from the high-speed pellet fraction RNA, we found abundant 24–26-nt-long fragments homologous to the ubiquitous retroposon INGI and the site-specific retroposon SLACS. Northern hybridization with strand-specific probes confirmed that retroposon-derived 24–26-nt RNAs are present in both supernatant and high-speed pellet fractions and that they are constitutively expressed. We speculate that RNAi in trypanosomes serves a housekeeping function and is likely to be involved in silencing retroposon transcripts.**

**Keywords: dsRNA; retroposon; RNAi; siRNA; *Trypanosoma brucei***

## INTRODUCTION

RNA interference or RNAi refers to the posttranscriptional gene silencing mechanism through which gene-specific dsRNA triggers degradation of homologous cellular transcripts (Fire, 1999; Sharp, 1999, 2001; Bosher & Labouesse, 2000). RNAi is widespread through eukaryotic phyla, is mechanistically and genetically related to PTGS in plants (Kooter et al., 1999; Dalmay et al., 2000; Fagard et al., 2000; Mourrain et al., 2000; Plasterk & Ketting, 2000) and "quelling" in *Neurospora* (Cogoni & Macino, 1999), and is considered a defense mechanism against aberrant transcripts such as those produced during viral infection (Ratcliff et al., 1997, 1999; Mourrain et al., 2000) and mobilization of transposons (Ketting et al., 1999; Tabara et al., 1999).

Reprint requests to: Elisabetta Ullu, Department of Internal Medicine, Yale University School of Medicine, PO Box 208022, 333 Cedar Street, New Haven, Connecticut 06520-8022, USA; e-mail: elisabetta.ullu@yale.edu.

Current models for RNAi suggest that upon entry into cells, dsRNA is processed by a nuclease to produce 21–25-nt small RNAs that act as guide sequences for degradation of target RNA and have been termed small interfering RNAs or siRNAs (Zamore et al., 2000; Elbashir et al., 2001a). The most significant pieces of evidence supporting the role of siRNA in silencing by RNAi come from in vitro studies showing that addition of synthetic dsRNA to a *Drosophila* extract results in cleavage of the target mRNA with a periodicity of 21–23 nt (Zamore et al., 2000) and that 21–22-nt synthetic RNAs with short 3′ extensions are inducers of RNAi (Elbashir et al., 2001b). The latter study established that the information of siRNAs is sufficient to guide degradation of target RNA; however it is still unclear whether larger dsRNA fragments also play a role in the RNAi silencing mechanism. Biochemical fractionation of a *Drosophila* cultured cell extract showed that the activity degrading mRNA requires an RNA component and that small 21–23-nt RNAs cofractionate with the

activity (Hammond et al., 2000). A candidate gene coding for the siRNA nuclease has been recently identified in *Drosophila* and has been termed *Dicer* for its ability to digest dsRNA to homogeneously sized pieces that most likely are equivalent to siRNAs (Bernstein et al., 2001). Homologs of *Dicer* have been recently shown to function in the production of small temporal RNAs (stRNAs) in *Caenorhabditis elegans*, *Drosophila*, and human cells, demonstrating that the pathways of RNAi and stRNA production intersect (Grishok et al., 2001; Hutvagner et al., 2001). Further evidence that siRNAs are part of the PTGS pathway(s) comes from experiments in plant systems, where expression of a viral- or cellular-encoded repressor of PTGS severely reduces the accumulation of siRNAs (Llave et al., 2000; Mallory et al., 2001).

Small RNAs with characteristics of siRNAs, namely, similar size and with sense and antisense polarity, were first described in transgenic plants engineered for PTGS (Hamilton & Baulcombe, 1999). Subsequently, they were detected in *Drosophila* cells following transfection of synthetic dsRNA (Hammond et al., 2000), and in *Drosophila* embryos (Yang et al., 2000) and *C. elegans* following microinjection of radiolabeled dsRNA (Parrish et al., 2000). Thus, in vivo expression of dsRNA from transgenes or delivery of synthetic dsRNA into cells is always accompanied by the appearance of siRNA-size molecules, making siRNAs the hallmark of RNAi.

Among the organisms where genetic interference by dsRNA has been demonstrated, the protozoan parasite *Trypanosoma brucei* represents the most ancient branch of the eukaryotic lineage. The target RNA is cytoplasmic mRNA, as pre-mRNA is not sensitive to RNAi (Ngo et al., 1998). In trypanosomes, mRNA degradation is triggered by either synthetic dsRNA delivered to cells by electroporation (Ngo et al., 1998), dsRNA produced in vivo from transgenes transcribed from opposing constitutive (Shi et al., 2000), or inducible T7 RNA polymerase promoters (LaCount et al., 2000; Wang et al., 2000), or hairpin RNA transcribed from the tetracycline (tet)-inducible procyclic acidic repetitive protein (PARP) promoter (Bastin et al., 2000; Shi et al., 2000). Whereas electroporation of dsRNA produces only a transient RNAi response, stable in vivo expression of dsRNA results in persistent RNAi. In particular, we have engineered a cell line (ACTI) expressing a hairpin RNA targeting nt 1–426 of the actin mRNA coding region (Shi et al., 2000). ACTI cells are viable and thus represent a useful system to study the mechanism of RNAi in trypanosomes.

Here we have analyzed by hybridization, cell fractionation, cloning and sequencing siRNAs derived from actin dsRNA in *T. brucei*, a protozoan parasite representing one of the earliest branch of the eukaryotic lineage. In the course of these studies, we discovered the presence of abundant siRNA-like molecules representing two retroposons that inhabit the trypanosome genome, and propose that in trypanosomes RNAi is a housekeeping mechanism to silence retroposon transcripts.

## RESULTS

### Identification of actin siRNAs in trypanosomes

We first asked whether in *T. brucei* siRNA-like molecules are produced upon activation of RNAi. For this purpose, we used the ACTI cell line expressing actin hairpin RNA (Shi et al., 2000) under the control of a tet-inducible promoter. Upon induction of actin hairpin RNA, 24–26-nt-long RNAs were generated (Fig. 1A), which had the characteristics of siRNAs first described in plants (Hamilton & Baulcombe, 1999): They contained information derived from both the sense (lane 4)

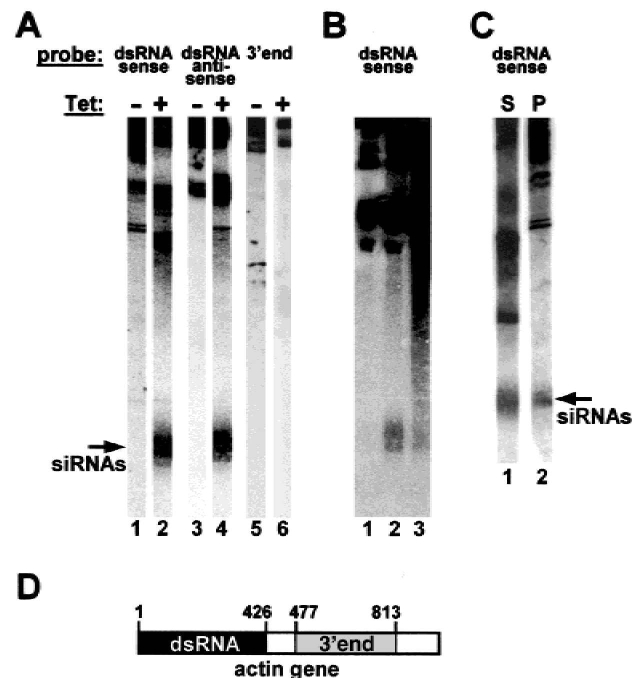


**FIGURE 1.** siRNAs are produced in trypanosomes upon expression or transfection of actin dsRNA. **A:** Low-molecular-weight RNAs from ACTI cells induced (+) or uninduced (−) with tetracycline (Tet) were fractionated through a 15% sequencing gel, blotted, and hybridized with a sense (lanes 1 and 2) or antisense (lanes 3 and 4) riboprobe representing actin dsRNA (solid box) or a 3′ end probe representing nucleotides 477–813 of the actin coding region (gray box; lanes 5 and 6). **B:** Northern blot analysis of low-molecular-weight RNAs from wild-type trypanosomes electroporated with poly(IC) (lane 1) or synthetic actin dsRNA (lanes 2 and 3; Ngo et al., 1998). RNA was isolated 2 h (lanes 1 and 2) or 6 h (lane 3) after electroporation. The probe was the sense strand of actin dsRNA. **C:** A proportion of actin siRNAs pellets at 100,000 × *g*. A cytoplasmic extract from tet-induced ACTI cells was separated into a high-speed pellet (P) and soluble fraction (S) as described in Materials and Methods. RNA from equivalent amounts of extracts was fractionated as described above and hybridized with a sense riboprobe representing actin dsRNA. The arrow indicates the position of a 28-nt DNA marker. **D:** Schematic representation of the *T. brucei* actin coding region not drawn to scale.

and antisense strand (lane 2) of actin dsRNA and the sequence information present in the small actin RNAs was restricted to the expressed dsRNA (lane 6). A similar observation was made when synthetic actin dsRNA was electroporated into trypanosomes (Fig. 1B), but in this case, the abundance of siRNA-like molecules declined over time (compare lanes 2 and 3), consistent with the transient nature of RNAi under these conditions (Ngo et al., 1998). About 10% of actin siRNA were pelleted by centrifugation of a cytoplasmic extract at 100,000 × *g* for 60 min (Fig. 1C), suggesting that, similar to the *Drosophila* system (Bernstein et al., 2001), a proportion of the actin siRNA is present in high-molecular-weight complexes. These centrifugation conditions are known to pellet large ribonucleoprotein complexes including ribosomes, ribosomal subunits, and polysomes. In contrast, soluble material, as well as small ribonucleoprotein complexes, are not pelleted under these conditions.

## Cloning of siRNAs

Having established that actin siRNA molecules are generated upon induction of RNAi, we next wished to characterize their structure and their relationship to actin dsRNA. To this end, we established the cloning scheme shown in Figure 2A and described in Materials and Methods. Briefly, 20–30-nt RNAs were gel purified from the soluble and high-speed pellet fraction of tet-induced ACTI cells, ligated to oligonucleotide 1B and reverse transcribed using the 45-nt-long oligonucleotide 1A as a primer. cDNAs were gel purified, ligated to oligonucleotide 2B, and the products were subsequently PCR amplified and digested with *Dpn*II. The resulting fragments (herein referred to as tags) were concatemerized and cloned. Because in preliminary experiments we found that actin tags represented about 1% of the cloned tags, we decided to enrich for actin-specific tags by hybrid selection with a biotinylated actin RNA (Fig. 2B). A synthetic transcript corresponding to the entire actin mRNA coding region was used in order to capture actin fragments derived from both the actin dsRNA and actin mRNA sequences located downstream from the dsRNA region. To reduce potential artifacts introduced by this selection procedure, hybridization was carried out over a wide range of temperatures. Next, we sequenced 174 tags each from the soluble and high-speed pellet fraction. The majority of the tags were derived from actin sequences (147 and 146 for the soluble and pellet fractions, respectively) and specifically from the actin dsRNA sequence (Fig. 3A,B). Only two tags represented sequences outside the dsRNA sequence demonstrating that the starting RNA material was enriched for sequences derived from actin dsRNA and did not represent a random collection of degradation products of actin mRNA. This latter result was also confirmed by southern hybridization of the amplified material before cloning (data not shown). The ends of 59% of the actin tags contained one or a few additional nucleotides not encoded in the actin gene that were most likely added by T4 RNA ligase. By analyzing the terminal nucleotides of the tags we found that C and G residues were underrepresented relative to A and T. This could be due to a preference of T4 RNA ligase to use as substrate for ligation RNAs ending with the latter nucleotides, or be a true

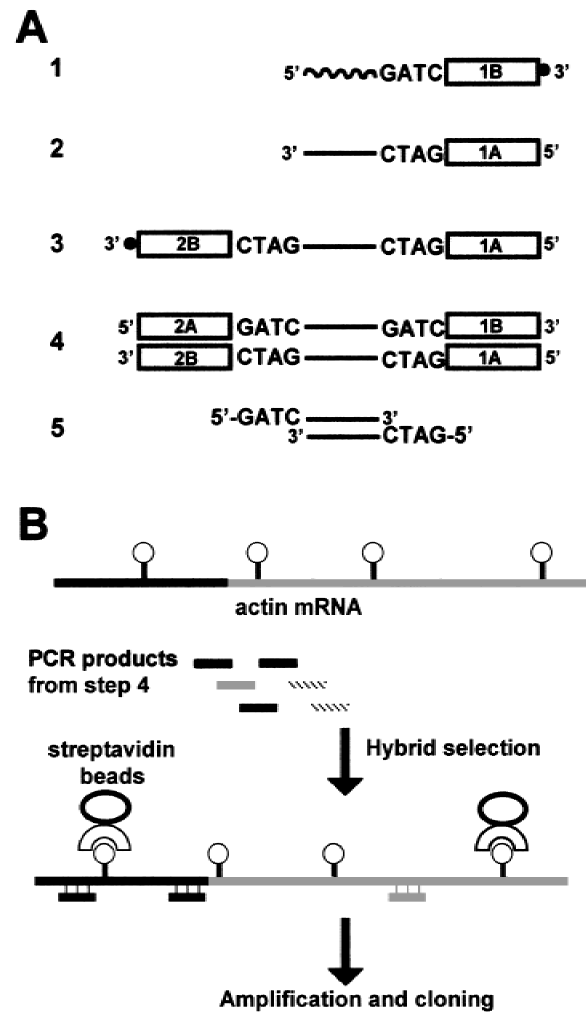


**FIGURE 2.** Schematic representation of the cloning procedure of 20–30-nt RNAs and of the hybridization selection for actin-specific small RNAs. **A:** 1. Ligation of oligonucleotide 1B to the 3′ end of 20–30-nt small RNAs. 2. cDNA synthesis. 3. Ligation of oligonucleotide 2B to the 3′ end of the cDNA. 4. PCR amplification of the cDNA. 5. Products of digestion of the amplified material with *Dpn*II. GATC indicates the recognition sequence of *Dpn*II; filled circles indicate the 3′ blocking group that prevents self-ligation and ligation of the oligonucleotides to the 5′ end of the RNA or cDNA. **B:** The structure of the synthetic transcript representing the actin mRNA coding region is shown at the top. Black lines: actin dsRNA-homologous region and corresponding PCR products; gray lines: sequence of actin mRNA downstream from the dsRNA region and corresponding PCR products; hatched lines: PCR products unrelated to actin mRNA; open circle: biotin moiety. For simplicity, the PCR products are represented without the oligonucleotide sequences used for amplification.

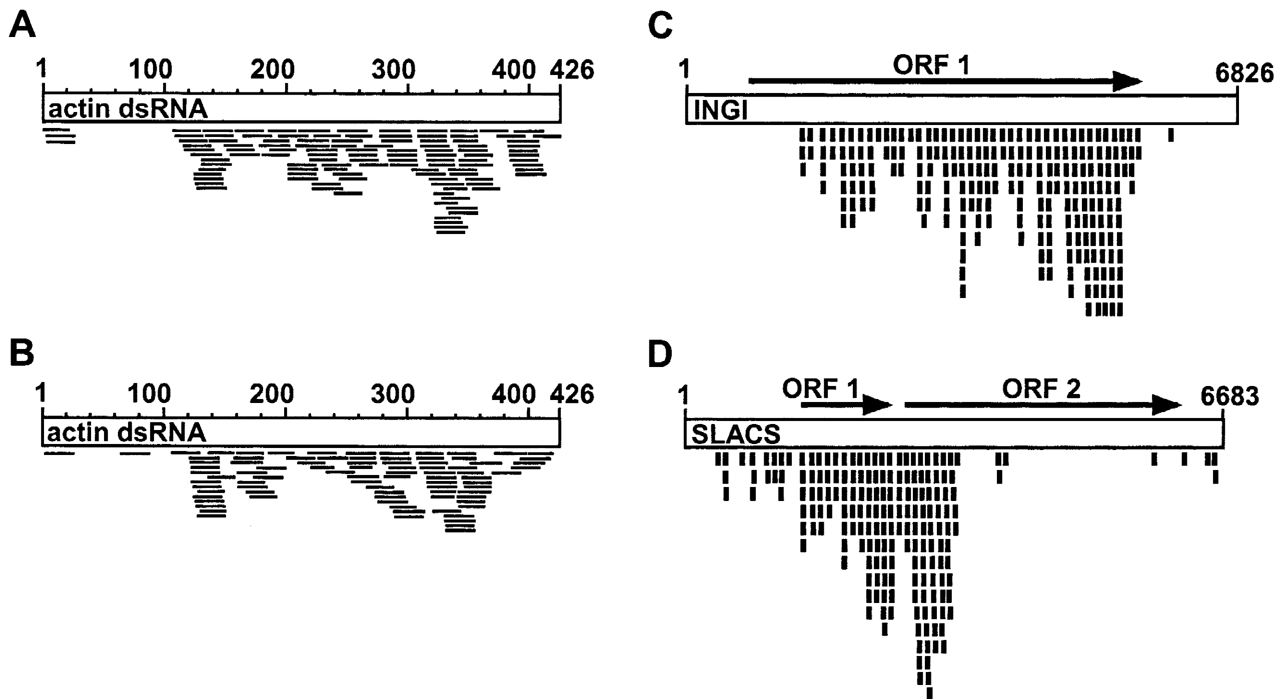


**FIGURE 3.** Distribution of actin, INGI, and SLACS siRNA tags. Actin siRNA tags (horizontal bars) were derived from the soluble fraction (**A**) or the high-speed pellet (**B**) of tet-induced ACTI cells. Before cloning,the PCR products were enriched for actin-specific tags as described in Materials and Methods. Only nonredundant actin tags are shown. INGI (**C**) and SLACS (**D**) tags (vertical bars) originated from the high-speed pellet-derived PCR products prior to the enrichment step. Solid arrows indicate the position of the predicted ORFs of INGI (accession number X05710) and SLACS (accession number X17078). Each drawing is to scale.

feature of the RNA fragments in the starting RNA material. Similar findings have been recently described for siRNAs produced in a *Drosophila* in vitro extract (Elbashir et al., 2001a).

The size of the actin tags varied from 20 to 30 nt, with 72% being between 24 and 26 nt long (10% 20–23 nt, 20% 24 nt, 29% 25 nt, 23% 26 nt, 11% 27 nt, and 5% 28–30 nt). This size distribution was confirmed by comparing the mobility of actin siRNAs relative to synthetic RNA oligonucleotides of 20, 23, and 27 nt in length (data not shown).

The distribution of unique actin tags from the soluble and pellet fraction was comparable along the dsRNA region, suggesting that the two siRNA populations are very similar (Fig. 3A,B). In both populations the tags were less represented towards the 5′ end relative to the remainder of the dsRNA sequence. In particular, only four tags originated from the very 5′ end and so far we have recovered one tag from the region between nt 27 and 107.

## Constitutive production of retroposon-specific siRNA-like molecules

Having established that actin siRNAs molecules were specifically represented in our starting RNA samples, we chose to analyze the population of tags derived from the high-speed pellet fraction RNA prior to selection for actin sequences. This choice was based on preliminary sequence results that indicated that the corresponding clones showed a higher prevalence of mRNA-derived tags as compared to those derived from the high-speed supernatant fraction. So far, we have analyzed 1,383 tags and a summary is shown in Table 1. Thirty-eight percent of the tags were derived from structural RNAs (ribosomal RNAs, 5S RNA, and tRNAs among others). Eighty-five tags (6.1%) could be assigned to various mRNAs and of these, 18 or 1.3% were actin tags. In contrast, we found only 3 tags derived from $\alpha$-tubulin mRNA, a mRNA that is at least 10 times more abundant than actin mRNA in trypano-

**TABLE 1.** Distribution of siRNA tags.

| | |
|---|---|
| Structural RNAs (rRNAs, tRNAs, etc.) | 519 (38%) |
| mRNAs/ESTs | 67 (4.8%) |
| Actin mRNA | 18 (1.3%) |
| Retroposons | |
| INGI | 245 (18%) |
| SLACS | 182 (13%) |
| HTGS/GSS | 242 (17%) |
| No hits | 110 (8%) |
| Total number of tags analyzed | 1,383 |

somes. Seventeen percent of the tags generated hits in the current HTGS/GSS *T. brucei* database, whereas 8% of the tags did not reveal matches in any of the databases.

Surprisingly, 31% of the tags represented sequences from two retroposon elements, namely the ubiquitous abundant element INGI (Kimmel et al., 1987; Murphy et al., 1987) and the site-specific retroposon SLACS (Aksoy et al., 1990) that interrupts the spliced leader RNA genes and is present in about 10 copies per genome. Both elements have the hallmarks of non-LTR retroposons. The full-length INGI element (6,826 bp) contains a single long ORF with homology to reverse transcriptase, whereas SLACS (6,683 bp) contains two ORFs. ORF1 has no known function, whereas ORF2 has homology to reverse transcriptase. As shown in Figure 3C,D, INGI and SLACS tags originated from selective portions of the retroposon sequences. For INGI, all but one of the tags (245) came from the single ORF, whereas in the case of SLACS, 163 tags were derived from ORF1 and the beginning of ORF2, and very few tags (19) came from outside the two ORFs. Interestingly, within the tag-rich regions, the representation of the tags was considerably less prominent towards the 5′ end. Dot blot analysis of total RNA with gene-specific probes (Table 2) revealed that, compared to $\alpha$-tubulin mRNA, INGI and SLACS transcripts were about 15- and 150-fold less abundant, respectively. Thus, relative to the representation of INGI and SLACS transcripts in total RNA, the 20–30 nt RNAs from the high-speed pellet fraction were highly enriched for retroposon sequences. Northern blot hybridization of RNA derived from uninduced and tet-induced ACTI cells (Fig. 4A) confirmed the presence of INGI (lanes 3 and 4) and SLACS (lanes 5 and 6) siRNA-size molecules in total trypanosome small RNAs. Importantly, INGI and SLACS siRNA-size molecules were present at the same level in RNA derived from tet-induced or uninduced cells. Similarly to actin siRNAs, INGI and SLACS small RNAs partitioned between the high-speed pellet and soluble fractions and had both sense and antisense polarity (Fig. 4B).

**TABLE 2**. Relative abundance of retroposon transcripts.[a]

| | 4.0 $\mu$g | 2.0 $\mu$g | 1.0 $\mu$g |
|---|---|---|---|
| Tubulin | 100 (160,259) | 100 (79,631) | 100 (45,803) |
| INGI | 5.9 (9,507) | 5.9 (4,734) | 6.2 (2,838) |
| SLACS | 0.7 (1,183) | 0.9 (677) | 0.7 (300) |

[a]Different amounts of total RNA from insect form trypanosome cells were loaded onto a nitrocellulose membrane and hybridized to $^{32}$P-labeled DNA probes corresponding to the coding region of $\alpha$-tubulin (nt 1–1030), INGI (nt 3421–4820), and SLACS (nt 1512–2726). The relative abundance of retroposon transcripts was calculated from the $\alpha$-tubulin hybridization set at 100%. Numbers in parenthesis are the original values in arbitrary units obtained after phosphorimager analysis of the membrane.

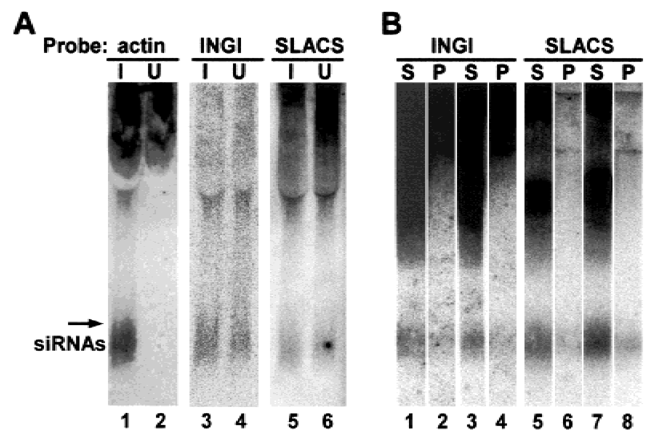


**FIGURE 4.** INGI and SLACS siRNAs. **A:** Production of INGI and SLACS siRNAs is constitutive. RNA from induced (I) or uninduced (U) ACTI cells was fractionated as described in the legend of Figure 1. Hybridizations were carried out with DNA probes representing nt 3421 to 4620 of INGI (lanes 3 and 4) or nt 1512 to 2726 of SLACS (lanes 5 and 6), or with an actin sense riboprobe corresponding to actin dsRNA (lanes 1 and 2). **B:** Sense and antisense INGI and SLACS siRNAs are present in both soluble and high-speed pellet fractions. A cytoplasmic extract from ACTI cells was fractionated into soluble and high-speed pellet as described in Materials and Methods. RNA representing equivalent portions of the two fractions were separated by gel electrophoresis as described in Figure 1. S: RNA from the soluble fraction; P: RNA from the high-speed pellet. Individual strips were hybridized to a sense (lanes 1 and 2) or antisense (lanes 3 and 4) INGI DNA probe or to a sense (lanes 5 and 6) or antisense (lanes 7 and 8) SLACS DNA probe. The arrow indicates the position of a 28-nt DNA marker.

## DISCUSSION

Our results show that, in trypanosomes, siRNAs are produced in response to activation of RNAi either by regulated expression or transfection of actin dsRNA. Furthermore, we have established a cloning strategy to survey 20–30-nt RNAs produced in vivo and report the discovery of abundant siRNA-like molecules derived from retroposon transcripts.

The actin-specific small RNAs we have characterized have all the features of siRNAs that have been described in other systems, namely, they are of similar size, they contain information from both strands of the dsRNA, and, as described in *Drosophila* cells (Bernstein et al., 2001), partition between a high-speed soluble and pellet fraction. Thus, it is very likely that these small RNAs perform the function of guiding degradation of the target mRNA as has been shown in other systems. Nevertheless, at present, it is difficult to discount the possibility that these short RNAs are produced by cleavage of the actin dsRNA by an RNAse III-like activity unrelated to the RNAi pathway.

To clone 20–30 nt RNAs, we employed a multistep protocol and PCR amplification using oligonucleotides and procedures that have been successfully established for serial analysis of gene expression (Virlon et al., 1999). By sequencing hybrid-selected actin tags,

we determined that they were derived almost exclusively from the actin dsRNA-homologous region and that their average size of 24–26 nt was indistinguishable from that of actin siRNAs present in the starting RNA material. Thus, in trypanosomes, siRNAs are a few nucleotides longer than siRNAs produced in vitro in a *Drosophila* extract, which are 20–22 nt long (Elbashir et al., 2001a), but interestingly very similar in size to the small RNAs described in plants, which have been reported to be about 25 nt long (Hamilton & Baulcombe, 1999). The difference in size between the trypanosome and *Drosophila* siRNAs might reflect differences in the cleavage specificity of the nuclease that degrades dsRNA (Bernstein et al., 2001) or differences between the in vitro and in vivo systems.

When we aligned the actin tags derived from either the soluble or the high-speed pellet fraction along the dsRNA region, we found that their distribution was indistinguishable, suggesting that the two populations of siRNAs are very similar. In both cases, tags derived from the 5′ end were underrepresented as compared to those from the 3′ end of the dsRNA. In contrast, within the regions of the actin sequence that were tag rich, the ends of the tags were closely spaced and there were no major discontinuities. This asymmetric distribution of siRNA sequences has also been observed by hybridization analysis of RNAs derived from transgenes in plant systems (Hutvagner et al., 2000) and might suggest that degradation of dsRNA in vivo proceeds preferentially from one end of the duplex, as has been recently determined to occur in a *Drosophila* in vitro extract competent for RNAi (Elbashir et al., 2001a), or that sequence preference dictates cleavage of the dsRNA. However, because the actin tags were enriched by hybridization, we cannot discount the possibility that the enrichment step led to a skewed representation of the tags.

To search for siRNA-size molecules that are endogenously produced in trypanosomes, we carried out extensive sequencing of unselected tags from the high-speed pellet fraction. The 1,383 tags analyzed so far could be divided into three classes of roughly similar size. The first class represented structural RNAs, with ribosomal fragments being the most abundant species.

Because these RNAs make up about 99% of the cellular RNA, the tags most likely represent unavoidable contaminants from RNA breakage that occurred in the cell or during the purification procedure. Indeed, northern blot analysis of DEAE-sepharose-enriched small RNAs from the high-speed pellet fraction with a 28S ribosomal RNA probe revealed a hybridization pattern consistent with random degradation products and there was no specific accumulation of ribosomal RNA fragments in the size range of 20–30 nt (data not shown). This is in contrast to actin- and retroposon-derived small RNAs, which were detected as discrete 24–26-nt fragments (Figs. 1 and 4).

The second class of tags, comprising 31% of the tags, consisted of a heterogeneous collection of sequences derived from known mRNAs, from unannotated sequences in the *T. brucei* genome database and from unsequenced regions of genome. The interpretation of the mRNA-derived tags will require additional experimentation, as for each mRNA, the number of specific tags is only one or few, and for most of them, the gene identity/function is unknown. At the present time, it is not possible to distinguish whether these tags represent true siRNAs present in low abundance or degradation products.

The third class of tags (31%) originated from the INGI and SLACS retroposon-like elements that inhabit the trypanosome genome. We believe that their abundance in the cloned material is highly significant, considering that transcripts derived from these two elements are much less abundant than $\alpha$-tubulin mRNA (Table 2), which represents about 1% of mRNA, or about 0.01% of total cellular RNA. Furthermore, northern hybridization with INGI- or SLACS-specific probes revealed the presence of siRNA-size molecules with all the characteristics of "true" siRNAs, including their partial fractionation with high-molecular-weight ribonucleoprotein particles. The most surprising result was that these siRNA-like molecules are constitutively expressed. One possible explanation for this observation is that these retroposon-derived small RNAs are involved in silencing retroposon transcripts via the RNAi pathway, and thus represent true siRNAs. As siRNAs are derived from processing of dsRNA, parts or all of the INGI and SLACS elements need to be transcribed from both strands. In the case of INGI transcripts, it is known from previous work that transcripts originate from both strands (Murphy et al., 1987), and we have confirmed these results by northern blotting with strand-specific probes (data not shown). Thus, there is the potential to form INGI dsRNA. In the case of SLACS we know less, because the abundance of SLACS transcripts in steady-state RNA is extremely low. Furthermore, because the transcription polarity of SLACS elements in the genome is not known and secondary structure predictions of the SLACS sequence do not show the potential to form extensive stem structures, it is at present not clear how SLACS dsRNA is made.

The observation that siRNA-like molecules from retroposon elements are constitutively expressed makes us speculate that, in trypanosomes, RNAi is part of a housekeeping quality control mechanism for mRNA surveillance that functions to down-regulate expression of deleterious gene products. An alternative scenario is that retroposon-derived small RNAs function as inhibitors of the RNAi pathway by preventing other siRNAs from getting access to the complex that degrades mRNA or by down-regulating the activity of a potential trypanosome *Dicer* homolog, which remains to be identified. However, we find this possibility unlikely in view of the

fact that the trypanosome RNAi machinery appears to have a high capacity and so far, we have not been able to saturate the system (unpubl. observation).

In the case of retroposon elements that require the synthesis of retroposon-encoded proteins for transposition and transpose via an RNA intermediate, degradation of retroposon transcripts by RNAi would effectively prevent transposon mobilization. Indeed, both in plants and animals, RNAi is considered a defense mechanism against harmful transcripts, such as those produced by transgenes, transposons, and viruses (Ratcliff et al., 1997, 1999; Ketting et al., 1999; Tabara et al., 1999; Mourrain et al., 2000). In particular, our results are consistent with the genetic connection between RNAi and silencing of transposon that has been established in *C. elegans* (Ketting et al., 1999; Tabara et al., 1999). However, in this case, these silenced transposons move via a DNA cut-and-paste mechanism rather than retroposition. Moreover, although the results were not described in the article, sequencing of 20–22-nt small RNAs from a *Drosophila* extract also revealed sequences derived from endogenous retroposons (Elbashir et al., 2001a). Given the fact that trypanosomes are descendants of ancient eukaryotes, it is tempting to speculate that RNAi is active in all organisms that have been invaded by transposons/retroposons during their evolutionary history. It will be interesting to determine whether siRNAs corresponding to transposon sequences are present in all organisms that harbor them.

## MATERIALS AND METHODS

### RNA analyses

Expression of actin hairpin RNA was induced by diluting ACTI cells (Shi et al., 2000) 10-fold in fresh medium containing 10 $\mu$g/mL tet and the cells were allowed to grow for 2 days. Total RNA was prepared with the TRIZOL reagent (Gibco). To enrich for low-molecular-weight RNAs, we used either binding and elution to DEAE-53 sepharose beads as described (Walter & Blobel, 1983), or filtration through Centricon-100 filtration units according to the manufacturer's instructions (Millipore Corp.). Both methodologies select for RNAs smaller than a few hundred nucleotides. Low-molecular-weight RNAs were electrophoresed through 15% sequencing gels, electroblotted to an uncharged nylon membrane, and cross-linked. Hybridizations to riboprobes were carried out in 50% formamide buffer at 40 °C for 16–20 h, which we determined was the optimal hybridization temperature. For hybridization with DNA probes, the temperature was lowered to 37 °C.

### Extract preparation and cloning of siRNAs

A cytoplasmic extract from $5 \times 10^9$ ACTI cells was prepared by detergent lysis as described (Sobel & Wolin, 1999). The postnuclear supernatant was further fractionated into soluble and high-speed pellet fractions by centrifugation at $100{,}000 \times g$ for 1 h. RNA was prepared by phenol-chloroform extraction and enriched for low-molecular-weight RNAs by DEAE-53 sepharose chromatography as described above. Enriched material was fractionated by electrophoresis through a 7 M urea-15% polyacrylamide gel, 20–30-nt RNAs were excised and recovered from the gel slice by soaking in DEAE-53 sepharose binding buffer followed by chromatography using 0.5-mL DEAE-53 sepharose beads. The sequence of the oligonucleotides used for cloning can be found in a modified SAGE protocol (Virlon et al., 1999). In particular, oligonucleotides 1B and 2B (45 nt long) carried at the 3′ end a blocking group to prevent self ligation and ligation to the RNA 5′ end, and the 5′ end of the oligonucleotides carried the recognition sequence 5′-GATC-3′ for the *Dpn*II restriction enzyme. Gel-purified small RNAs were ethanol precipitated, dephosphorylated, resuspended in 5 $\mu$L of water, and 2.5 $\mu$L were ligated with T4 RNA ligase to 20 pmol of oligonucleotide 1B in 10 $\mu$L (O'Brien & Wolin, 1994). Ligated material was phenol extracted, precipitated with ethanol, and used for cDNA synthesis with Superscript II reverse transcriptase (Gibco) in the presence of $\alpha$-$^{32}$P-dCTP and using 40 pmol of oligonucleotide 1A as a primer. cDNA was fractionated on a 15% sequencing gel to remove the oligonucleotides, excised from the gel, eluted in DEAE-53 binding buffer, and recovered by DEAE-53 sepharose chromatography. The cDNA was then ligated with RNA ligase to 20 pmol of oligonucleotide 2B as described above. Aliquots of ligated cDNA were amplified by PCR using Platinum Taq DNA polymerase (Gibco) and oligonucleotides 1A and 2A. For large-scale PCR amplification, conditions were chosen in the linear range of the reaction. A 2-mL PCR mix containing 10 $\mu$Ci of $\alpha$-$^{32}$P-dCTP was distributed in 50-$\mu$L aliquots and incubated using the following conditions: 94 °C for 30 s, 65 °C for 1 min, and 72 °C for 1 min for up to 30 cycles. PCR products were purified with the Qiagen PCR pure kit and separated by electrophoresis through a 1.5% agarose gel. The DNA was purified from the agarose gel slices, digested with 100 U of *Dpn*II for 2 h, and the digested material was fractionated by electrophoresis through a native 15% polyacrylamide gel. The released fragments were recovered by DEAE-53 sepharose chromatography, resuspended in 4 $\mu$L of water, and ligated to each other using the rapid ligation kit from Roche. The ligated material was phenol extracted, ethanol precipitated, and fractionated by gel electrophoresis through a 2-mm-thick 1.2% agarose gel. Fragments larger than 200 nt were then recovered from the gel and ligated to *Bam*HI-cleaved and dephosphorylated plasmid vector pUC18. The ligation mixture was electroporated into electrocompetent DH10B cells according to the manufacturer's instruction (Gibco).

Colonies were screened by hybridization and by PCR. Plasmid DNAs were sequenced using either the forward or reverse primer. To enrich for actin tags, about 1 $\mu$g of PCR product from the large-scale amplification reaction was resuspended in 50 $\mu$L of water, denatured at 95 °C for 5 min, and hybridized to 1.5 $\mu$g of sense biotinylated actin coding region RNA in 5 × SET. Hybridization was carried out for 30 min at each of the following temperatures: 65 °C, 60 °C, 55 °C, 50 °C, 45 °C, 40 °C, and 37 °C. Hybridized material was recovered by binding to 1 mg of streptavidin–magnetic beads according to the manufacturer's instructions (Dynal), washed, and eluted at 70 °C in water. Eluted material was PCR amplified using the conditions described above and the

hybridization–selection procedure was repeated once more. PCR products were then processed as described above. To verify the efficiency of the selection procedure, aliquots of PCR products before selection and after each round of selection were analyzed by southern hybridization with two actin-specific probes, one representing the dsRNA region and the other one representing nt 477–813 of actin mRNA. The results of this analysis showed that the twice-selected PCR fragments had been enriched for actin sequences about 300-fold relative to the starting material and that they did not contain detectable sequences downstream of the dsRNA region.

## ACKNOWLEDGMENTS

Our thanks go to Najib El-Sayed for invaluable support; to Shulamit Michaeli, Joan Steitz, and Sandra Wolin for helpful discussions and critical reading of the manuscript; and to Sarah Renzi for invaluable technical help. This work was supported by National Institutes of Health Grant AI28798 to E.U. E.U. is the recipient of a Burroughs Wellcome Fund Scholar Award in Molecular Parasitology and C.T. is the recipient of a Burroughs Wellcome Fund New Investigator Award in Molecular Parasitology.

*Received August 14, 2001; accepted without revision August 30, 2001*

## REFERENCES

Aksoy S, Williams S, Chang S, Richards FF. 1990. SLACS retrotransposon from *Trypanosoma brucei gambiense* is similar to mammalian LINEs. *Nucleic Acids Res 18*:785–792.

Bastin P, Ellis K, Kohl L, Gull K. 2000. Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. *J Cell Sci 113*:3321–3328.

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature 409*:363–366.

Bosher JM, Labouesse M. 2000. RNA interference: Genetic wand and genetic watchdog. *Nat Cell Biol 2*:E31–E36.

Cogoni C, Macino G. 1999. Homology-dependent gene silencing in plants and fungi: A number of variations on the same theme. *Curr Opin Microbiol 2*:657–662.

Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC. 2000. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell 101*:543–553.

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature 411*:494–498.

Elbashir SM, Lendeckel W, Tuschl T. 2001b. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Dev 15*:188–200.

Fagard M, Boutet S, Morel JB, Bellini C, Vaucheret H. 2000. AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. *Proc Natl Acad Sci USA 97*:11650–11654.

Fire A. 1999. RNA-triggered gene silencing. *Trends Genet 15*:358–363.

Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell 106*:23–34.

Hamilton AJ, Baulcombe DC. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science 286*:950–952.

Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature 404*:293–296.

Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD. 2001. A cellular function for the RNA-interference enzyme *Dicer* in the maturation of the let-7 small temporal RNA. *Science 293*:834–838.

Hutvagner G, Mlynarova L, Nap JP. 2000. Detailed characterization of the posttranscriptional gene-silencing-related small RNA in a GUS gene-silenced tobacco. *RNA 6*:1445–1454.

Ketting RF, Haverkamp TH, van Luenen HG, Plasterk RH. 1999. Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell 99*:133–141.

Kimmel BE, ole-MoiYoi OK, Young JR. 1987. Ingi, a 5.2-kb dispersed sequence element from *Trypanosoma brucei* that carries half of a smaller mobile element at either end and has homology with mammalian LINEs. *Mol Cell Biol 7*:1465–1475.

Kooter JM, Matzke MA, Meyer P. 1999. Listening to the silent genes: Transgene silencing, gene regulation and pathogen control. *Trends Plant Sci 4*:340–347.

LaCount DJ, Bruse S, Hill KL, Donelson JE. 2000. Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. *Mol Biochem Parasitol 111*:67–76.

Llave C, Kasschau KD, Carrington JC. 2000. Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. *Proc Natl Acad Sci USA 97*:13401–13406.

Mallory AC, Ely L, Smith TH, Marathe R, Anandalakshmi R, Fagard M, Vaucheret H, Pruss G, Bowman L, Vance VB. 2001. HC-Pro suppression of transgene silencing eliminates the small RNAs but not transgene methylation or the mobile signal. *Plant Cell 13*:571–583.

Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, Remoue K, Sanial M, Vo TA, Vaucheret H. 2000. Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell 101*:533–542.

Murphy NB, Pays A, Tebabi P, Coquelet H, Guyaux M, Steinert M, Pays E. 1987. *Trypanosoma brucei* repeated element with unusual structural and transcriptional properties. *J Mol Biol 195*:855–871.

Ngo H, Tschudi C, Gull K, Ullu E. 1998. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci USA 95*:14687–14692.

O'Brien CA, Wolin SL. 1994. A possible role for the 60-kD Ro autoantigen in a discard pathway for defective 5S rRNA precursors. *Genes & Dev 8*:2891–2903.

Parrish S, Fleenor J, Xu S, Mello C, Fire A. 2000. Functional anatomy of a dsRNA trigger. Differential requirement for the two trigger strands in RNA interference. *Mol Cell 6*:1077–1087.

Plasterk RH, Ketting RF. 2000. The silence of the genes. *Curr Opin Genet Dev 10*:562–567.

Ratcliff F, Harrison BD, Baulcombe DC. 1997. A similarity between viral defense and gene silencing in plants. *Science 276*:1558–1568.

Ratcliff FG, MacFarlane SA, Baulcombe DC. 1999. Gene silencing without DNA. RNA-mediated cross-protection between viruses. *Plant Cell 11*:1207–1216.

Sharp PA. 1999. RNAi and double-strand RNA. *Genes & Dev 13*:139–141.

Sharp PA. 2001. RNA interference–2001. *Genes & Dev 15*:485–490.

Shi H, Djikeng A, Mark T, Wirtz E, Tschudi C, Ullu E. 2000. Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA 6*:1069–1076.

Sobel SG, Wolin SL. 1999. Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes. *Mol Biol Cell 10*:3849–3862.

Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC. 1999. The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell 99*:123–132.

Virlon B, Cheval L, Buhler JM, Billon E, Doucet A, Elalouf JM. 1999. Serial microanalysis of renal transcriptomes. *Proc Natl Acad Sci USA 96*:15286–15291.

Walter P, Blobel G. 1983. Disassembly and reconstitution of signal recognition particle. *Cell 34*:525–533.

Wang Z, Morris JC, Drew ME, Englund PT. 2000. Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *J Biol Chem 275*:40174–40179.

Yang D, Lu H, Erickson JW. 2000. Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in *Drosophila* embryos. *Curr Biol 10*:1191–1200.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell 101*:25–33.