

# Retrotransposition of a yeast group II intron occurs by reverse splicing directly into ectopic DNA sites

Lorna Dickson<sup>\*\*†</sup>, Hon-Ren Huang<sup>\*†</sup>, Lu Liu<sup>\*</sup>, Manabu Matsuura<sup>§¶</sup>, Alan M. Lambowitz<sup>§</sup>, and Philip S. Perlman<sup>\*||</sup>

<sup>\*</sup>Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390-9148; and <sup>§</sup>Institute for Cellular and Molecular Biology, Department of Chemistry and Biochemistry, and Section of Molecular Genetics and Microbiology, University of Texas, Austin, TX 78712

Communicated by Marlene Belfort, New York State Department of Health, Albany, NY, September 19, 2001 (received for review July 23, 2001)

**Group II introns, the presumed ancestors of nuclear pre-mRNA introns, are site-specific retroelements. In addition to “homing” to unoccupied sites in intronless alleles, group II introns transpose at low frequency to ectopic sites that resemble the normal homing site. Two general mechanisms have been proposed for group II intron transposition, one involving reverse splicing of the intron RNA directly into an ectopic DNA site, and the other involving reverse splicing into a site in RNA followed by reverse transcription and integration of the resulting cDNA by homologous recombination. Here, by using an “inverted-site” strategy, we show that the yeast mtDNA group II intron aI1 retrotransposes by reverse splicing directly into an ectopic DNA site. This same mechanism could account for other previously described ectopic transposition events in fungi and bacteria and may have contributed to the dispersal of group II introns into different genes.**

**G**roup II introns are both catalytic RNAs and mobile genetic elements (1). These introns splice by means of a lariat intermediate and may be the ancestors of nuclear pre-mRNA introns (2). As mobile elements, the introns insert efficiently into unoccupied sites in intronless alleles by a process called homing or retrohoming and also transpose at low frequency to ectopic sites that resemble the normal homing site. The basic features of the retrohoming mechanism were originally elucidated in studies of the yeast mtDNA introns aI1 and aI2, and more recent studies have extended those findings to several other introns (reviewed in refs. 1 and 3). Retrohoming occurs by a target DNA-primed reverse transcription (TPRT) mechanism mediated by a ribonucleoprotein (RNP) particle containing the intron-encoded reverse transcriptase (RT) and the excised intron RNA lariat. The intron RNA in this RNP particle reverse splices directly into the sense strand of a DNA target site, whereas the intron-encoded protein (IEP) cleaves the opposite strand at a specific site and uses the 3' end of the cleaved strand as a primer for reverse transcription of the inserted intron RNA. The resulting cDNA copy of the intron is then integrated by recombination or repair mechanisms, which may differ among organisms (4–6). In yeast mitochondria, retrohoming can occur either by complete reverse splicing and synthesis of a full-length cDNA, which may be integrated by repair enzymes, or by synthesis of a partial cDNA, which invades an intron-containing allele to initiate double-strand break repair (DSBR) recombination (4, 5).

Group II intron homing sites span  $\approx 30$  bp, from about position  $-20$  upstream of the intron insertion site to position  $+10$  downstream (see Fig. 1A for the aI1 site) (7). The IEP first recognizes several nucleotide residues in the distal 5' region of the site (positions  $-13$ ,  $-18$ ,  $-19$ , and  $-21$  for aI1) and promotes local DNA unwinding, enabling the intron RNA to base pair to sense-strand positions  $-12$  through  $+1$  for reverse splicing. This DNA–RNA interaction involves three short sequence elements in the DNA target site (IBS2, IBS1, and  $\delta'$ ), which base pair with complementary intron RNA sequences denoted EBS2, EBS1, and  $\delta$  (Fig. 1A). Antisense-strand cleavage occurs after reverse splicing and requires additional interactions between the protein and the 3' exon (including positions  $+4$ ,  $+6$ , and  $+9$  for aI1) (7–9). The same pairings between the

intron and flanking exons also occur in precursor RNA and are required for RNA splicing (2). The use of similar base-pairing interactions for RNA splicing and DNA target-site recognition ensures that the introns insert only at sites from which they can be spliced.

The yeast intron aI1 can also transpose to ectopic sites that resemble the natural homing site but at very low frequencies (10, 11). Initially, ectopic insertion sites were identified in the downstream introns of the *COXI* (cytochrome *c* oxidase subunit I) gene, including the group I introns aI5 $\alpha$  and aI5 $\beta$  and the group II intron aI5 $\gamma$  (Fig. 1). The intermediate containing the intron inserted into the ectopic site was detected by PCR (Fig. 1Bi). However, the final products of those transpositions were large deletions resulting from recombination between the duplicated copies of the inserted intron (Fig. 1Bii). aI1 also inserts at the E1/aI1 junction in mtDNA to form intron dimers (or excised circles) (Fig. 1Biii). Similar group II intron ectopic integration events resulting in large deletions of mtDNA sequences between the two intron copies have been reported for two other fungi (11, 12).

It was proposed that the ectopic integration events in fungal mtDNAs are initiated by the excised intron RNA reverse splicing into a transcript containing the ectopic site, followed by reverse transcription and integration of the resulting cDNA into mtDNA by homologous recombination (10, 13). Initial evidence suggested that the *Lactococcus lactis* Ll.LtrB intron retrotransposes by this mechanism (14). Those bacterial events require the RT activity of the IEP and were reported to be RecA-dependent, in contrast to retrohoming, which is RecA-independent. Recent experiments, however, have called into question the RecA dependence of Ll.LtrB transposition and have also shown that ectopic transposition may not have a strong strand bias (K. Ichihyanagi, A. Beauregard, and M. Belfort, personal communication). We previously suggested an alternative model for ectopic retrotransposition that involves reverse splicing of the intron RNA directly into DNA target sites (7, 15). The feasibility of this mechanism was established by the finding that the aI1 intron could reverse splice *in vitro* into double-stranded DNA substrates containing the ectopic sites (7). Reverse splicing directly into ectopic DNA sites was reported for transposition events of a *Sinorhizobium meliloti* group II intron, but the generality of the mechanism was not clear because the ectopic sites analyzed were all closely related to the normal intron homing site (16).

Abbreviations: RNP, ribonucleoprotein; RT, reverse transcriptase; IEP, intron-encoded protein.

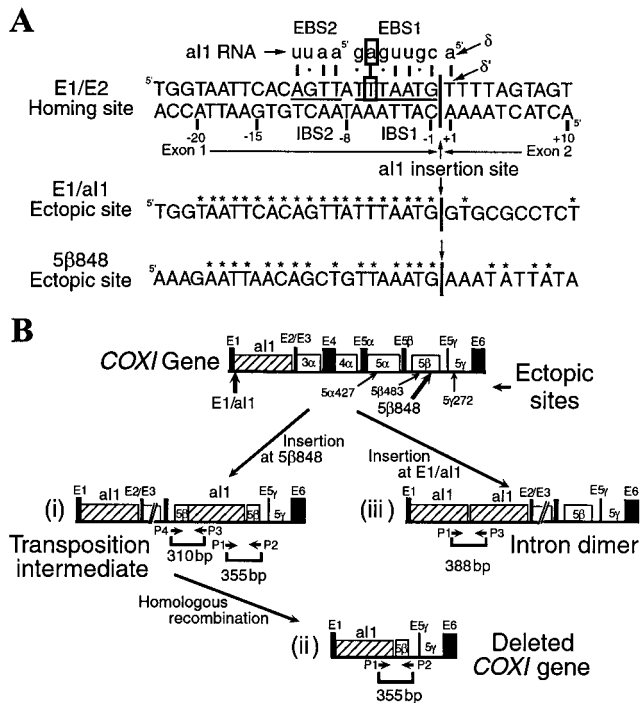
<sup>†</sup>L.D. and H.-R.H. contributed equally to this work.

<sup>\*</sup>Present address: Pacific Northwest Research Institute, 720 Broadway, Seattle, WA 98122.

<sup>§</sup>Present address: National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan, 305-8562.

<sup>||</sup>To whom reprint requests should be addressed. E-mail: philip.perlman@utsouthwestern.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Fig. 1.** DNA target sites and *COXI* (cytochrome c oxidase subunit I) gene rearrangements caused by transposition of a11. (A) DNA target-site sequences. The sequence of the a11 homing site is shown from position -23 in exon 1 (E1) to +10 in exon 2 (E2). The IBS1, IBS2, and  $\delta'$  sequences are indicated, with the complementary intron RNA sequences EBS1, EBS2, and  $\delta$  shown above. The base-pairing interaction at position -6 of EBS1/IBS1, which is disrupted by the A262T mutation in the intron RNA, is boxed. The sense-strand sequences of the E1/a11 and 5 $\beta$ 848 ectopic sites are shown below. Asterisks indicate nucleotide residues that are conserved between the natural site and the ectopic sites. DNA and RNA sequences are indicated in uppercase and lowercase letters, respectively. (B) Diagram of the *COXI* gene and products of a11 transposition in strain 1<sup>+t20</sup>. Exons are filled rectangles and intron ORFs are hatched (a11) or open (a13 $\alpha$ -a15 $\beta$ ) rectangles. Five ectopic sites for a11 insertion are indicated below with bold arrows indicating the sites analyzed here. (i) a11 insertion at the 5 $\beta$ 848 site. (ii) *COXI* deletion resulting from a11 insertion at the 5 $\beta$ 848 ectopic site, followed by recombination between the two copies of a11. (iii) a11 insertion at the E1/a11 site resulting in two tandem copies of a11 ("intron dimer"). It is possible that one intron copy is excised from mtDNA by recombination to yield a circular intron DNA. PCR primers used to detect novel junctions resulting from ectopic insertions and the ensuing recombination events are indicated below the diagrams.

The purpose of this study was to characterize the retrotransposition pathway of a11 and to determine whether the initial target is RNA or DNA. Our findings show that a11 retrotransposition occurs by reverse splicing directly into a DNA target site. This mechanism can also account for other ectopic transposition events in fungi and bacteria and for the dispersal of group II introns into different genes, possibly including eukaryotic nuclear genes in the course of evolution to spliceosomal introns.

## Materials and Methods

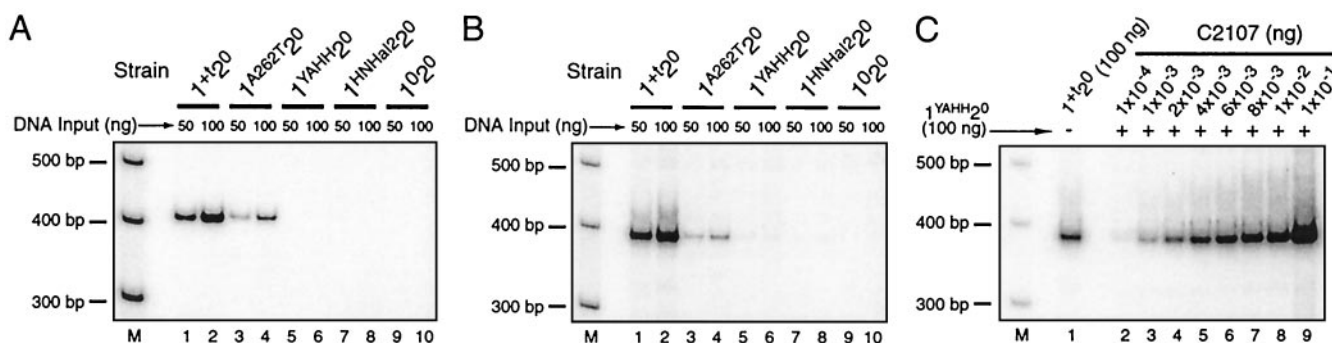
**Yeast Strains and Genetic Manipulations.** Yeast mtDNA genotypes are denoted by a convention in which a superscript + indicates the wild-type intron, a superscript 0 indicates that the intron is absent, and other superscripts refer to specific mutations. The primary yeast strain, 1<sup>+t20</sup>, is a derivative of wild-type strain ID41-6/161 (*MATa ade1 lys1*) and has the six-intron form of the *COXI* gene shown in Fig. 1B (4). Strain C2107 has the *COXI* deletion diagrammed in Fig. 1Bii (4). Strain 1<sup>020</sup> (Fig. 2) is identical to 1<sup>+t20</sup> except that it lacks both a11 and a12 (17). Strain

5 $\beta$ <sup>0</sup> has a *COXI* gene containing only introns 3 $\alpha$  and 4 $\alpha$  (18). Strain 1<sup>A262T20</sup> carries the mutation A262T in the EBS1 of a11, resulting in a mismatch with nucleotide E1-6T in the EBS1/IBS1 pairing (see Fig. 1A and ref. 4). In strain 1<sup>YAHH20</sup> the conserved YADD motif of the RT domain is mutated to YAHH (4). In strain 1<sup>AAVA20</sup> the conserved HHVK motif of the endonuclease domain is mutated to AAVA; it was constructed similarly to the analogous a12 mutant (19). In strain 1<sup>HNH $\alpha$ 1220</sup>, the HNH endonuclease domain of a11 (the last 66 aa) is replaced by PCR with the homologous sequence from a12 (the last 64 aa). Intron mutations were constructed in plasmid pJVM155, transformed into mtDNA by biolistic transformation, and confirmed by DNA sequencing (4).

Strains 1<sup>+t20</sup>5 $\beta$ <sup>+</sup> $\gamma$ <sup>0</sup>, 1<sup>+t20</sup>5 $\beta$ <sup>848S</sup> $\gamma$ <sup>0</sup>, and 1<sup>+t20</sup>5 $\beta$ <sup>848AS</sup> $\gamma$ <sup>0</sup> contain a wild-type, tagged wild-type, or inverted a15 $\beta$ 848 ectopic site, respectively, and are deleted for a15 $\gamma$ . They were constructed by biolistic transformation of plasmid pTZ5 $\beta$ <sup>+</sup> $\gamma$ <sup>0</sup>-*COXII* or its derivatives containing the specified mutations into a  $\rho$ <sup>0</sup> derivative of strain MCC109 (*MAT $\alpha$  ade2-101 ura3-52 kar1-1*), followed by recombination into the mtDNA of a derivative of strain 1<sup>+t20</sup> with a point mutation that blocks a15 $\gamma$  splicing (see ref. 20). Selection for respiratory sufficient diploids yielded strains in which the recombinant a15 $\beta$  was present in a mtDNA lacking a15 $\gamma$  (5 $\gamma$ <sup>0</sup>). Plasmid pTZ5 $\beta$ <sup>+</sup> $\gamma$ <sup>0</sup>-*COXII* contains a segment of the *COXI* gene from position 348 in intron a15 $\beta$  through the first *EcoRI* site in exon 6, amplified from strain SUV3-1 (21) by RT-PCR and cloned between the *SacI* and *XbaI* sites of pTZ18U (Bio-Rad). The strains were characterized by sequencing a15 $\beta$  after PCR amplification. The inverted ectopic site contains an inadvertent G-to-T mutation at position -8 between IBS1 and IBS2, a position not critical for efficiency of the target site. In our strains, the 5 $\beta$ 848 site is 880 bp from the 5' end of the intron, compared with 848 bp in strains used in another study (10). Growth of yeast cultures and genetic manipulations were as described (22).

**PCR Analysis of a11 Transposition.** Semiquantitative PCR was carried out in 50- $\mu$ l reaction mixtures containing 10-300 ng of mtDNA, 30 pmol of each primer, 0.2 mM each dNTP, PCR buffer with 3 mM MgCl<sub>2</sub>, 2.5 units of *Taq* DNA polymerase, and 0.1  $\mu$ Ci of [<sup>32</sup>P]dATP (3,000 Ci/mmol; 1  $\mu$ Ci = 37 kBq). mtDNA inputs were equalized by Southern hybridization using a random primer-labeled PCR product containing 407 bp of *COXI* exon 6. The annealing temperature for PCR was optimized and the number of cycles was varied to find conditions where the signal intensity is proportional to the amount of template. The reaction mixtures were prepared and denatured at 95°C for 3 min and then cycled at 95°C for 30 sec, 50°C for 30 sec, and 72°C for 45 sec, with the number of cycles and pairs of primers indicated for each experiment. PCR products were analyzed in nondenaturing 5% polyacrylamide gels, which were dried onto filter paper and quantitated with a PhosphorImager. In initial experiments, the relevant bands were sequenced by the dideoxynucleotide method. The positions of the PCR primers are shown in Fig. 1B and Fig. 4. The sequences are P1, 5'-CATCATGTTAAACAATTACATAGAGG; P2, 5'-TATTGGTCATATCTTTTATATG; P3, 5'-TCCCCGTAAGTTAGCCCCTACTGAG; and P4, 5'-CTGGATTATT-TGAAGGAGATGG.

**Biochemical Methods.** Yeast mtDNA and RNA were isolated from cells grown to early logarithmic phase in YP-2% raffinose medium (22, 23). Southern and Northern hybridizations were as described (22, 24). Biochemical assays of reverse splicing and DNA endonuclease activity were carried out with mitochondrial RNP particle preparations (19, 25). Endogenous primer-dependent RT activity of RNP particle preparations was measured as described (26) except that the primer was complemen-



**Fig. 2.** PCR analysis of retrotransposition into ectopic sites. (A and B) PCR assay for aI1 insertions into the E1/aI1 and 5β848 sites. mtDNAs ( $\approx$ 50 and 100 ng) from the indicated strains were analyzed by PCR to detect products of ectopic transposition into the E1/aI1 site (A; 22 cycles with primers P1 and P3) and 5β848 site (B; 25 cycles with primers P1 and P2). Similar results for the 5β848 site were obtained by using primers P3 and P4 (25 cycles) (not shown). Strain 1<sup>020</sup> (lanes 9 and 10) has mtDNA identical to that of strain 1<sup>+20</sup> except that aI1 is deleted. (C) Estimate of the frequency of ectopic transposition events at the 5β848 site. mtDNA from strain 1<sup>+20</sup> (100 ng) was subject to 25 cycles of PCR using primers P1 and P2 to detect the downstream junction between aI1 and aI5β (see Fig. 1B*ii*), and that signal (lane 1) was compared with a calibration curve obtained with mtDNA from strain C2107, which contains the *COXI* deletion shown in Fig. 1B*ii* (lanes 2–9). In the experiment shown, decreasing amounts of mtDNA from strain C2107 were mixed with mtDNA from the inactive strain 1<sup>YAHH20</sup> to maintain a constant 100 ng of DNA, and the same result was obtained in reaction mixtures containing only the C2107 mtDNA dilution series (not shown). A 5'-labeled 100-bp ladder was used for molecular weight markers (M) here and in the subsequent figures.

tary to aI1 positions 86–110 and the reactions were chased with 2 mM unlabeled dTTP for 10 min.

## Results

**Assays for Ectopic Transposition.** Fig. 1B shows the *COXI* gene of yeast strain 1<sup>+20</sup>, with five known ectopic transposition sites indicated by arrows. The sequences of the two sites analyzed in detail here, E1/aI1 and 5β848, are shown in Fig. 1A (10, 11). Because all of the ectopic sites are in the sense strand of the gene, each is present in both *COXI* transcripts and mtDNA, so that retrotransposition could, in principle, be initiated by reverse splicing into the RNA or DNA.

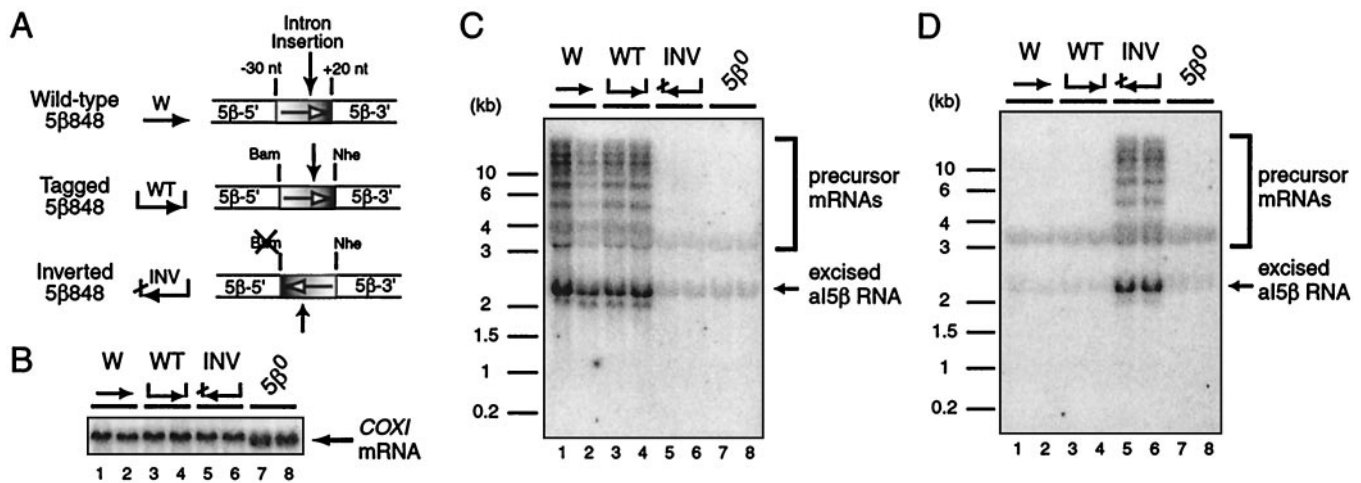
First, we confirmed the previous reports of ectopic transposition of aI1 into the E1/aI1 and 5β848 sites in the *COXI* gene by semiquantitative PCR, using primer pairs that detect novel junctions resulting from intron insertion (Fig. 1B *i* and *iii*). Spontaneous transposition of aI1 into the E1/aI1 site of the 1<sup>+20</sup> strain yields a strong radioactive band after 22 PCR cycles with 50 and 100 ng of template (Fig. 2A, lanes 1 and 2), whereas insertion into the 5β848 site yields similarly intense bands after 25 PCR cycles with the same amounts of template (Fig. 2B, lanes 1 and 2). We confirmed by sequencing that the PCR products contain the expected junction sequences and that intron insertion does not alter the target-site sequence (not shown). The strength of each signal varies over a wide range of DNA template amounts (Fig. 2A and B, lanes 1 and 2 and not shown), and no signal was detected by using the same or higher amounts of DNA from the negative control strain 1<sup>020</sup>, which lacks aI1 (Fig. 2A and B, lanes 9 and 10).

**Effects of Intron Mutations on Ectopic Transposition.** The EBS/IBS pairings are required for reverse splicing into both RNA and DNA target sites. We showed previously that the aI1 mutation A262T, which results in a U·U mismatch in EBS1/IBS1 (see Fig. 1A), inhibits reverse splicing and homing into the E1/E2 site >100-fold, and that both are restored by a compensatory change in the DNA target site (4). The mutant aI1 still splices efficiently *in vivo* even though the EBS1/IBS1 pairing in the pre-mRNA is weakened. The A262T mutation results in the same mismatch at position –6 of EBS1/IBS1 in the ectopic E1/aI1 and 5β848 sites and results in reduced, but detectable, ectopic transposition at both sites (3- to 4-fold at E1/aI1 and 16- to 18-fold at 5β848) (Fig. 2A and B; lanes 3 and 4). We conclude that the EBS1/IBS1

pairing influences retrotransposition, but that the mismatch has less effect on transposition than on homing.

Strain 1<sup>YAHH20</sup> lacks aI1 RT activity because of a mutation in the conserved YADD motif in the RT domain. The mutant shows no detectable retrohoming in crosses, although it still supports  $\approx$ 50% residual homing by an RT-independent double-strand break repair pathway, initiated by reverse splicing and antisense-strand cleavage by the aI1 endonuclease (4). This RT domain mutation inhibits aI1 ectopic transposition >100-fold at both sites, indicating an RT-dependent mechanism (Fig. 2A and B, lanes 5 and 6).

Finally, we tested whether aI1 transposition requires the function of the HNH endonuclease domain. Mutating the conserved HHVK motif of the domain to AAVA (strain 1<sup>AAVA20</sup>) inhibited homing into the E1/E2 site and transposition into both ectopic sites by >95% (not shown). A complication, however, is that mutations of the HNH motif not only block endonuclease activity but also inhibit the RT activity of the IEP (ref. 26; J. San Filippo and A.M.L., unpublished data). To address this problem, we constructed strain 1<sup>HNHΔI20</sup>, in which the portion of the endonuclease domain of aI1 containing the HNH motif was replaced with the homologous sequence from aI2. Biochemical assays showed that this hybrid mutant intron has wild-type levels of primer-dependent RT activity and reverse splicing using the E1/E2 substrate, but lacks detectable antisense-strand cleavage activity, presumably because the aI2-derived portion of the hybrid protein cannot recognize the E2 sequence (not shown). Nevertheless, the mutant splices aI1 efficiently *in vivo* and has  $\approx$ 70% of control homing activity in crosses, 44% of which occurs by retrohoming as judged by coconversion of upstream but not downstream markers (data not shown; see refs. 4 and 5). These findings indicate that this mutant intron retains RT activity *in vivo* and imply that inhibiting antisense-strand cleavage in this case only moderately inhibits homing. This phenotype resembles that of a previously described aI2 endonuclease mutant, P714T, which is surprisingly active for retrohoming at some target sites but inactive at others (5). Either these mutants retain some antisense-strand cleavage activity *in vivo* or they rely on an alternative mechanism for priming or antisense-strand cleavage at some sites. Regardless, as shown in Fig. 2A and B, lanes 7 and 8, the 1<sup>HNHΔI20</sup> mutation blocks insertion of aI1 into both the E1/aI1 and 5β848 sites, suggesting that the endonuclease domain of the protein contributes to retrotransposition.



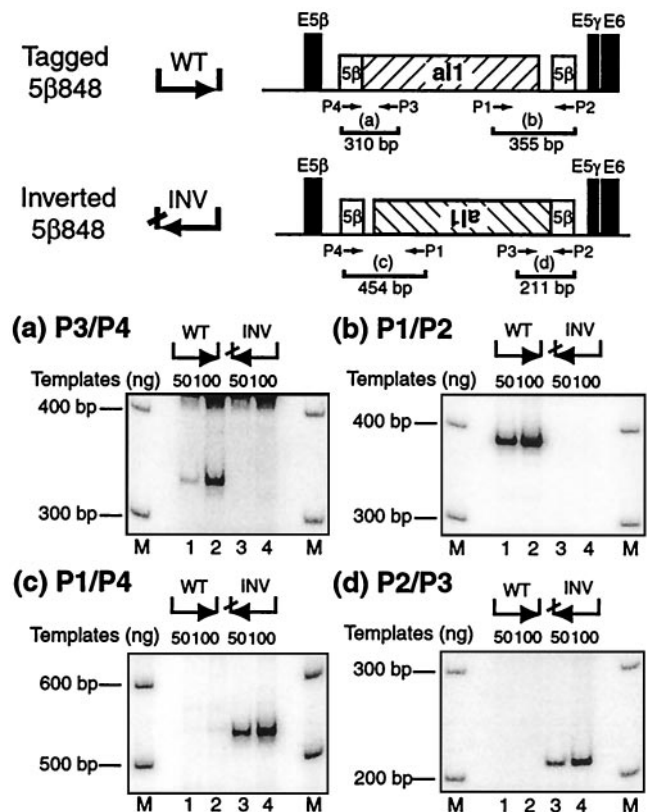
**Fig. 3.** Diagram of wild-type and modified 5 $\beta$ 848 sites and Northern hybridization analysis of sense- and antisense-strand transcripts of aI5 $\beta$ . (A) Diagrams of wild-type and constructed 5 $\beta$ 848 alleles. The wild-type (W) strain contains the normal 5 $\beta$ 848 site and lacks aI5 $\gamma$  (5 $\gamma^0$ ). The tagged wild-type 5 $\beta$ 848 site (WT), containing flanking *Bam*HI and *Nhe*I sites, was constructed by site-directed mutagenesis. Then the 50 bp between the *Bam*HI and *Nhe*I sites were inverted (with disruption of the *Bam*HI site) to make the inverted 5 $\beta$ 848 site (INV). Both mutated sites were transformed into mtDNA to yield strains with the tagged WT or INV 5 $\beta$ 848 sites. (B–D) Northern hybridizations. Northern blots of two independent preparations of cellular RNAs from the indicated strains were hybridized with a *COXI* exon 6-specific probe to detect *COXI* mRNA (loading control) (B) or oligonucleotide probes complementary to the sense (C) and antisense (D) strands of the 5 $\beta$ 848 ectopic site. Strain 5 $\beta^0$  contains a *COXI* gene lacking aI5 $\beta$ . Faint signals at  $\approx$ 2.3 and 3.4 kb in all strains are due to cross-hybridization with rRNA.

**Transpositions at the 5 $\beta$ 848 Site Are Quite Rare.** To estimate the frequency of aI1 ectopic transposition, mtDNA from strain 1<sup>+t20</sup> was subjected to PCR with primers P1 and P2 to detect the downstream junction between aI1 and aI5 $\beta$  (Fig. 1B), and that signal was compared with a calibration series using mtDNA from strain C2107, which has the *COXI* deletion resulting from retrotransposition at the 5 $\beta$ 848 site (Fig. 1Bii). As shown in Fig. 2C, 100 ng of wild-type 1<sup>+t20</sup> mtDNA (lane 1) yields the same signal as does  $2\text{--}4 \times 10^{-3}$  ng of C2107 mtDNA (lanes 4 and 5). From these data, we calculate that there are about three aI1–5 $\beta$ 848 junctions per 100,000 mtDNAs. Because a haploid yeast cell contains about 20 copies of mtDNA, one product of ectopic transposition is present in  $\approx$ 1,700 cells.

**A Critical Test of the RNA-Target Hypothesis.** The findings that the EBS1/IBS1 pairing and aI1 RT activities are important for ectopic transposition indicate that the mechanism involves reverse splicing and reverse transcription, but do not establish whether the target for reverse splicing is RNA or DNA. To address this issue, we devised a test of the RNA-target hypothesis by analyzing the effect of inverting the aI5 $\beta$ 848 site in mtDNA. Because the promoters for all of the protein-coding genes of yeast mtDNA are on the same DNA strand (27), the correct target sequence should no longer be present for reverse splicing into mitochondrial transcripts. The aI5 $\beta$ 848 site is a very favorable one to test this hypothesis because it is in the aI5 $\beta$  ORF, which is not needed for splicing aI5 $\beta$  (28).

We first added *Bam*HI and *Nhe*I sites to the flanks of a 50-bp sequence that includes the wild-type 5 $\beta$ 848 site (W), forming the tagged wild-type 5 $\beta$ 848 site (WT) (see Fig. 3A). The segment between the *Bam*HI and *Nhe*I sites was then replaced with a double-stranded oligonucleotide that inverts the entire sequence (inactivating the *Bam*HI site) to make the inverted-5 $\beta$ 848 site (INV). Both the tagged and inverted sequences were introduced into the *COXI* gene of strain 1<sup>+t20</sup> in place of the wild-type 5 $\beta$ 848 site by mitochondrial transformation. The resulting strains have *COXI* genes that lack aI5 $\gamma$ , but are otherwise identical to the *COXI* gene of 1<sup>+t20</sup> (see Fig. 1B, Fig. 4, and *Materials and Methods*).

The validity of this strategy depends critically on there being few if any transcripts of the antisense strand of aI5 $\beta$ . The strains



**Fig. 4.** The inverted 5 $\beta$ 848 site is a target for aI1 retrotransposition. The diagrams at the top show the expected products of aI1 retrotransposition into the tagged wild-type (WT) and inverted (INV) orientations of the 5 $\beta$ 848 site. Insertion of aI1 in the sense orientation is detected by PCR using primers P3 and P4 (a) and P1 and P2 (b), whereas insertion of aI1 in the antisense orientation is detected by using primers P1 and P4 (c) and P2 and P3 (d). The gels show the products obtained in these reactions when mtDNA from strains with the WT and INV 5 $\beta$ 848 sites was used. Reaction a used 27 cycles; reactions b and c used 25 cycles; and reaction d used 32 cycles.

all respire well, and the Northern blots of Fig. 3B show that all contain similar levels of *COXI* mRNA, indicating that aI5 $\beta$  is spliced efficiently. The Northern blots of Fig. 3C and D measure the levels of transcripts of each strand of aI5 $\beta$ . With a probe complementary to the sense strand of the 5 $\beta$ 848 target site, the two control strains (W and tagged WT) show detectable excised aI5 $\beta$  intron RNA and larger pre-mRNAs (Fig. 3C, lanes 1–4). In the strain with the inverted site, the only signals with this probe result from a low level of cross-reaction with rRNA (lanes 5 and 6), also seen in the null control strain lacking aI5 $\beta$  (lanes 7 and 8). Conversely, a probe complementary to the antisense strand of the 5 $\beta$ 848 site detects excised aI5 $\beta$  RNA and *COXI* pre-mRNAs in the strain with the inverted site (Fig. 3D, lanes 5 and 6), but only cross-hybridization with rRNAs is evident in the other strains (lanes 1–4 and 7 and 8). Thus, the data show that there are no detectable transcripts of the antisense strand of aI5 $\beta$  in any of these strains.

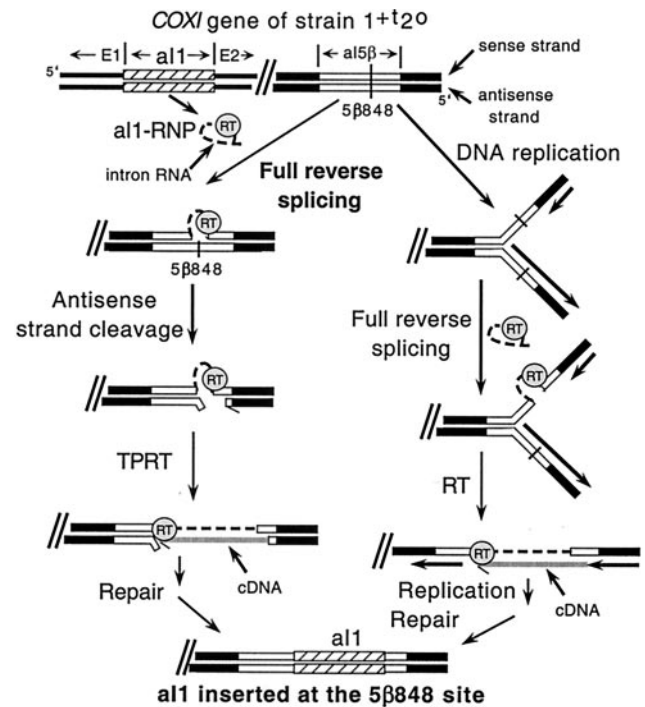
The RNA-target hypothesis predicts that the inverted site will block retrotransposition, because the correct site is not present in transcripts, whereas the DNA-target hypothesis predicts that the inverted site will permit retrotransposition, but that the intron will insert in the opposite orientation. mtDNAs from the strains containing the wild-type tagged (WT) and inverted (INV) sites were used as templates in four PCRs that assay for the 5' and 3' junctions between aI1 and aI5 $\beta$  sequences in both possible orientations (Fig. 4; reactions a–d). The results show that the 5 $\beta$ -tagged site in the normal orientation supports retrotransposition in the sense orientation but not in the antisense orientation (lanes 1 and 2), as expected. By contrast, inversion of the site permits insertion in the antisense orientation (reactions c and d, lanes 3 and 4) but blocks it in the sense orientation (reactions a and b, lanes 3 and 4). These data show that DNA is the major target for aI1 retrotransposition.

## Discussion

The finding that the inverted 5 $\beta$ 848 site supports aI1 insertion into mtDNA provides strong evidence that a DNA target is used for group II intron retrotransposition in yeast mitochondria. We showed previously that aI1 can reverse splice *in vitro* into five known ectopic sites, including 5 $\beta$ 848 (7). Here, we show that transposition into ectopic sites *in vivo* depends on the RT activity of the IEP, as well as the IBS1/EBS1 pairing. Together, these findings indicate that ectopic transposition of aI1 to the 5 $\beta$ 848 site occurs by a mechanism in which the intron reverse splices directly into the DNA target site and is then reverse transcribed by the IEP (Fig. 5). Our results do not distinguish whether the initial target site is in double- or single-stranded DNA (e.g., at a replication fork as indicated in Fig. 5). In either case, completion of the transposition events requires the synthesis of a full-length DNA copy of the intron, which is then integrated by DNA replication or repair enzymes. Reverse splicing into DNA can also account for retrotransposition into the other ectopic sites in yeast mtDNA, as well as for other previously described ectopic transposition events in fungi and bacteria.

The previous studies of group II intron retrotransposition in yeast were undertaken to explain mtDNA deletions starting at the last nucleotide of aI1. Consequently, the PCR experiments that were reported analyzed only sense-strand insertions (10, 11). Our inverted-site experiment shows that switching the 5 $\beta$ 848 site to the antisense strand does not strongly inhibit retrotransposition and suggests that there is little if any strand bias, at least for this site.

As discussed previously, all of the yeast ectopic sites can form a reasonable EBS1/IBS1 pairing, albeit with some mismatches, but they show less conservation in IBS2 and in the 5' and 3' flanking regions recognized by the IEP (7). This is essentially the same pattern found for the *L. lactis* LI.LtrB intron transposition sites (14). We show here that a mutation in EBS1 that results in



**Fig. 5.** Retrotransposition mechanisms using DNA targets. The *COXI* gene of strain 1+20 (top) contains both the donor aI1 intron (hatched) and the 5 $\beta$ 848 ectopic site in intron 5 $\beta$  (open rectangle). The mechanism on the left begins with reverse splicing into the ectopic site in double-stranded DNA. Inefficient nicking of the antisense strand forms the primer for full-length cDNA synthesis by the RT with completion of intron insertion by DNA repair. The mechanism on the right begins with reverse splicing into the ectopic site at a replication fork. cDNA synthesis is initiated either *de novo* or by using the 3' end of the newly made leading strand with further replication and repair needed to complete intron insertion. Potential single-stranded DNA target sites may also exist on the lagging strand of the replication fork or in actively transcribed regions of mtDNA.

a mispairing with IBS1 at both the E1/aI1 and 5 $\beta$ 848 sites decreases the level of retrotransposition, but the inhibition is only partial, whereas the same mispairing decreases retrohoming by >100-fold (4). These results are consistent with the finding that ectopic transposition sites can tolerate some mispairings between EBS1 and IBS1 (7) and suggest that the frequency of ectopic transposition is ordinarily limited by events that occur after reverse splicing.

The finding that mutations in the endonuclease domain of the IEP strongly inhibit retrotransposition suggests that the endonuclease activity may be needed for these events. One possibility is that the initial reverse splicing event occurs into double-stranded DNA and the endonuclease is required to generate the primer for target DNA-primed reverse transcription (TPRT), as in retrohoming (Fig. 5, left pathway). Although the ectopic sites lack key nucleotide residues required for efficient antisense-strand cleavage, it is possible that the endonuclease cleaves inefficiently at or near position +10 to generate a primer at levels that cannot be detected biochemically. A caveat is that we cannot exclude the possibility that these endonuclease domain mutations inhibit retrotransposition indirectly by affecting the RT or another activity of the IEP (see ref. 26). With the *L. lactis* intron, a mutation that inhibited DNA endonuclease activity but not RT activity had little effect on the retrotransposition frequency, suggesting that the endonuclease activity is not required in that system (14). As discussed previously, it is possible that some other endonuclease activity nicks the antisense strand to generate the primer for cDNA synthesis (7). Alternatively, if reverse

splicing occurs into single-stranded DNA, priming could occur by an alternative mechanism (e.g., *de novo* or using the 3' end of a newly synthesized strand) without need for the endonuclease activity (see Fig. 5, right pathway) (7). The finding that the RT activity of the IEP is required for transposition argues against mechanisms in which the reverse spliced intron is copied directly by mtDNA polymerase.

PCR experiments estimate that there are three products of aI1 retrotransposition into the 5 $\beta$ 848 site per 100,000 mtDNA molecules. That value is about 10,000-fold lower than the level of retrohoming into the normal aI1 target site. Factors that could contribute to the low level of retrotransposition include deviations from the normal target-site sequence, inefficient antisense-strand cleavage, an inefficient mechanism for generating the DNA primer, and the absolute requirement for both complete reverse splicing and synthesis of a full-length cDNA copy of the intron.

Given the efficiency of homologous recombination in yeast mitochondria, it is surprising that the major retrotransposition pathway does not involve reverse splicing into RNA followed by cDNA synthesis and recombination. An analogous mechanism involving recombination between an intron-containing mtDNA and a cDNA copy of a spliced mRNA appears to be responsible for spontaneous intron deletion from mtDNA (29). A possible explanation is our previous observation that the intron RNA

associated with the IEP preferentially reverse splices into DNA rather than RNA, whereas the free RNA reverse splices efficiently into RNA but not DNA (19). A similar change in target preference as a result of association with the IEP may likewise favor reverse splicing into DNA in other systems.

On the basis of the data now available, it seems likely that reverse splicing into DNA is the predominant pathway for group II intron retrotransposition in both fungal mtDNA and bacteria (this work and K. Ichyanagi, A. Beauregard, and M. Belfort, personal communication). Although it is possible that group II intron retrotransposition can also occur by the RNA-based mechanism, this remains to be demonstrated. In the bacterial systems studied thus far, retrotransposition now appears to be largely RecA-independent, arguing against the RNA-based mechanism as a main pathway. Further, the sequence specificity of reverse splicing into DNA is sufficiently flexible to allow low levels of insertion into sites that deviate substantially from the normal target-site sequence, permitting the dispersal of introns to new sites on an evolutionary time scale. Finally, if group II introns were, in fact, the ancestors of spliceosomal introns, reverse splicing into DNA may also have contributed to the initial dispersal of introns in eukaryotic nuclear genomes.

This work was supported by National Institutes of Health Grants GM31480 to P.S.P. and GM37949 to A.M.L. and by Robert A. Welch Foundation Grant I-1211 to P.S.P.

- Lambowitz, A. M., Caprara, M. G., Zimmerly, S. & Perlman, P. S. (1999) in *The RNA World*, eds. Gesteland, R., Cech, T. R. & Atkins, J. (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed., pp. 451–485.
- Michel, F. & Ferat, J. L. (1995) *Annu. Rev. Biochem.* **64**, 435–461.
- Belfort, M., Derbyshire, V., Parker, M. M., Cousineau, B. & Lambowitz, A. M. (2001) in *Mobile DNA*, eds. Craig, N. L., Craigie, R., Gellert, M. & Lambowitz, A. M. (Am. Soc. Microbiol., Washington, DC), 2nd Ed., in press.
- Eskes, R., Yang, J., Lambowitz, A. & Perlman, P. (1997) *Cell* **88**, 865–874.
- Eskes, R., Liu, L., Ma, H., Chao, M. Y., Dickson, L., Lambowitz, A. M. & Perlman, P. S. (2000) *Mol. Cell. Biol.* **20**, 8432–8446.
- Cousineau, B., Smith, D., Lawrence-Cavanagh, S., Mueller, J. E., Yang, J., Mills, D., Manias, D., Dunny, G., Lambowitz, A. M. & Belfort, M. (1998) *Cell* **94**, 451–462.
- Yang, J., Mohr, G., Perlman, P. S. & Lambowitz, A. M. (1998) *J. Mol. Biol.* **282**, 505–523.
- Guo, H., Zimmerly, S., Perlman, P. S. & Lambowitz, A. M. (1997) *EMBO J.* **16**, 6835–6848.
- Singh, N. N. & Lambowitz, A. M. (2001) *J. Mol. Biol.* **309**, 361–386.
- Mueller, M. W., Allmaier, M., Eskes, R. & Schweyen, R. J. (1993) *Nature (London)* **366**, 174–176.
- Schmidt, W. M., Schweyen, R. J., Wolf, K. & Mueller, M. W. (1994) *J. Mol. Biol.* **243**, 157–166.
- Sellem, C. H., Lecellier, G. & Belcour, L. (1993) *Nature (London)* **366**, 176–178.
- Bonen, L. & Vogel, J. (2001) *Trends Genet.* **17**, 322–331.
- Cousineau, B., Lawrence, S., Smith, D. & Belfort, M. (2000) *Nature (London)* **404**, 1018–1021.
- Yang, J., Zimmerly, S., Perlman, P. S. & Lambowitz, A. M. (1996) *Nature (London)* **381**, 332–335.
- Martinez-Abarca, F. & Toro, N. (2000) *Nucleic Acids Res.* **28**, 4397–4402.
- Kennell, J. C., Moran, J. V., Perlman, P. S., Butow, R. A. & Lambowitz, A. M. (1993) *Cell* **73**, 133–146.
- Wenzlau, J. M., Saldanha, R. J., Butow, R. A. & Perlman, P. S. (1989) *Cell* **56**, 421–430.
- Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P. S. & Lambowitz, A. M. (1995) *Cell* **83**, 529–538.
- Butow, R. A., Henke, R. M., Moran, J. V., Belcher, S. C. & Perlman, P. S. (1996) *Methods Enzymol.* **264**, 265–278.
- Conrad-Webb, H., Perlman, P. S., Zhu, H. & Butow, R. A. (1990) *Nucleic Acids Res.* **18**, 1369–1376.
- Moran, J. M., Zimmerly, S., Eskes, R., Kennell, J. C., Lambowitz, A. M., Butow, R. A. & Perlman, P. S. (1995) *Mol. Cell. Biol.* **15**, 2828–2838.
- Farrell, R. E., Jr. (1998) *RNA Methodologies: A Laboratory Guide for Isolation and Characterization* (Academic, San Diego), 2nd Ed., pp. 54–103.
- Podar, M., Chu, V. T., Pyle, A. M. & Perlman, P. S. (1998) *Nature (London)* **391**, 915–918.
- Zimmerly, S., Guo, H., Perlman, P. S. & Lambowitz, A. M. (1995) *Cell* **82**, 545–554.
- Zimmerly, S., Moran, J. V., Perlman, P. S. & Lambowitz, A. M. (1999) *J. Mol. Biol.* **289**, 473–490.
- Foury, F., Roganti, T., Lecrenier, N. & Purnelle, B. (1998) *FEBS Lett.* **440**, 325–331.
- Johnson, C. H. & McEwen, J. E. (1997) *Mol. Gen. Genet.* **256**, 88–91.
- Levra-Juillet, E., Boulet, A., Seraphin, B., Simon, M. & Faye, G. (1989) *Mol. Gen. Genet.* **217**, 168–171.