

Two for the price of one: a bifunctional intron-encoded DNA endonuclease-RNA maturase

Marlene Belfort

Molecular Genetics Program, Wadsworth Center, New York State Department of Health, and School of Public Health, State University of New York at Albany, Albany, New York 12201, USA

Homing endonucleases are enzymes that are typically encoded by intervening sequences of various kinds—introns that splice at the RNA level, or inteins that splice at the protein level (for review, see Belfort et al. 2002). The endonucleases function primarily to mobilize these elements, which are often self-splicing, by facilitating their integration to new sites in DNA. Because these sites are most often allelic intronless or inteinless sites, this form of mobility is termed homing, and the target sequences are termed homing sites.

Homing endonucleases sometimes do double duty, also functioning as maturases, which promote RNA splicing (for review, see Lambowitz and Perlman 1990; Lambowitz and Belfort 1993). Maturases are usually intron specific, and act as cofactors that bind their intron-containing precursor RNA to facilitate splicing. Nevertheless, catalysis required for splicing remains a property of the RNA.

The endonuclease-gene invasion hypothesis

How might a single protein have evolved to process two substrates, DNA and RNA? It is widely held that DNAs encoding self-splicing introns, particularly group I introns, and self-splicing inteins were colonized by endonuclease genes, which are considered the primordial invasive genetic elements (Fig. 1; for review, see Lambowitz and Perlman 1990; Lambowitz and Belfort 1993). Maturase activity is postulated to have been acquired secondarily, as an adaptation by which to minimize the adverse consequences of endonuclease gene invasion on the splicing function of the intron (Fig. 2). This sequence of events would result in a mutualistic relationship, in which the endonuclease gene is afforded a phenotypically neutral hiding place in a self-splicing element, which in turn acquires the ability to spread to new sites via its newly found endonuclease partner that cleaves DNA.

There is strong circumstantial evidence in favor of the “endonuclease-gene invasion” hypothesis, and in favor of such events having occurred multiple times in the course of evolution of mobile self-splicing elements. A persuasive point is that endonucleases of different families are encoded by highly similar introns. More compelling yet is that inteins, which are molecular entities distinct from introns, house endonucleases of the same family as those associated with some group I introns (Fig. 1; for review, see Belfort and Roberts 1997).

The LAGLIDADG endonuclease-maturase family

Despite this assumed sequence of events, maturase function of these intron-encoded proteins was discovered before their endonuclease activity (Lazowska et al. 1980). Almost a quarter-century ago, maturase activity was ascribed to several *Saccharomyces cerevisiae* mitochondrial intron-encoded proteins, on the basis of splicing defects caused by mutation of the intron-encoded proteins (for review, see Lambowitz and Perlman 1990; Lambowitz and Belfort 1993). Most of these maturases are now classified as members of the LAGLIDADG family of proteins, so named for strings of conserved amino acid residues. Yeast mitochondrial LAGLIDADG proteins were shown subsequently to promote intron mobility by virtue of their endonuclease activity (for review, see Dujon 1989; Perlman and Butow 1989), and for some yeast mitochondrial introns, maturase and endonuclease functions were shown to reside in the same molecule (Delahodde et al. 1989; Wenzlau et al. 1989). The specificity of the maturases suggests that each evolved to exploit and/or accommodate specific features of its host intron. After acquiring a dedicated splicing factor, an intron was free to lose its self-splicing activity, and most maturase-dependent introns are no longer self-splicing under physiological conditions. However, in some cases, endonuclease activity can be reactivated by one or a few amino acid changes (for review, see Lambowitz et al. 1999).

Separate binding sites for DNA and RNA?

A key question that arose in relation to the structural and functional basis of specificity of one protein for both

E-MAIL belfort@wadsworth.org; FAX (518) 474-3181.

Article and publication are at <http://www.genesdev.org/cgi/doi/10.1101/gad.1162503>.

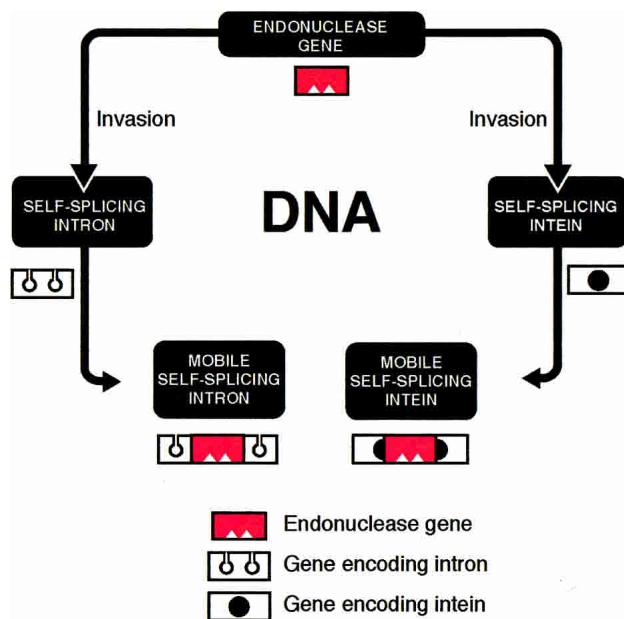


Figure 1. Hypothetical DNA invasion by an endonuclease gene. The schematic depicts invasion of self-splicing intron or intein DNA by an endonuclease gene, to generate mobile intervening sequences.

DNA and RNA was—one binding site, or two? Although endonuclease and maturase functions have been separated by mutation (Henke et al. 1995; Szczepanek et al. 2000), that question can only be answered through application of biochemical and structural tools applied to purified endonuclease-maturases. The yeast proteins have been a challenge in this respect, because of the mitochondrial genetic code and the proteins' toxicity, and because many of these proteins are expressed as fusions with the upstream exon (for review, see Lambowitz and Belfort 1993). However, the breakthrough came with expression in *Escherichia coli* of the LAGLIDADG endonuclease-maturase I-AniI, encoded by an *Aspergillus nidulans* group I intron (Ho et al. 1997). The stage was set by the demonstration that this protein has readily assayable maturase and endonuclease activities in vitro. Earlier this year, two groups, headed by Mark Caprara and by Richard Waring, showed independently, by nucleic acid competition experiments, that the I-AniI-

binding sites for RNA and DNA appear to be functionally distinct (Chatterjee et al. 2003; Geese et al. 2003). These laboratories joined forces with pioneers in the homing-endonuclease structure-function field, led by Barry Stoddard, to report the definitive experiments in this issue of *Genes and Development* (Bolduc et al. 2003).

The crystal structure of the 250-amino acid I-AniI monomer in complex with a 31-bp DNA substrate, which spans the homing site, was determined to 2.6 Å resolution. The pseudosymmetric structure has a core α - β fold typical of a handful of previously determined LAGLIDADG endonucleases (Fig. 3A; for review, see Chevalier and Stoddard 2001). Thus, two pseudosymmetric domains are held in close proximity by the repeated LAGLIDADG motifs, which form parts of interfacial α -helices. This signature sequence also contains the acidic endonuclease active site residues Asp 16 and Glu 148. These residues, which are represented by D in LAGLIDADG (underscored), each bind a Mg^{2+} ion, and protrude into the DNA-binding cleft. The DNA-binding surface is formed by two four-stranded antiparallel β -sheets, one contributed by each domain. Specific residues in these β -sheets engage in base-specific interactions, which dictate the DNA-binding specificity. The target DNA is curved at $\sim 45^\circ$, with the scissile phosphates positioned about 5 Å apart, near the conserved acidic residues. Cleavage at these phosphates generates a double-strand break with the 4-nucleotide 3' overhangs characteristic of the LAGLIDADG endonucleases.

What, then, of the maturase function of I-AniI? By alignment and modeling of homologous maturases, a battery of conserved basic residues, mostly remote from the DNA-binding cleft, were identified as forming a possible RNA-binding surface (Fig. 3A,B). Strikingly, an arginine-to-glutamate substitution at residue 239 implicated one of these residues in maturase function. Whereas this mutant maintained DNA-binding and cleavage activities, both its RNA-binding and maturase functions were reduced ~ 10 -fold (Bolduc et al. 2003). This experiment fits well with a reciprocal mutational study, in which the two acidic active-site residues of the endonuclease were changed to alanines. Although endonuclease function was obliterated, maturase function remained intact (Chatterjee et al. 2003). Clearly, the business centers for DNA and RNA reside in two distinct regions of I-AniI.

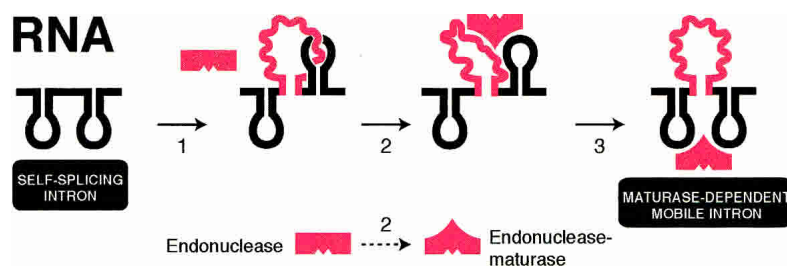


Figure 2. Evolution of a maturase-dependent mobile intron from a self-splicing intron. (Step 1) Intron invasion by an endonuclease gene is followed by synthesis of the endonuclease, in this case I-AniI. (Step 2) The resulting impairment of intron self-splicing drives acquisition of maturase function by the endonuclease, likely in several stages (dashed arrow). (Step 3) Maturase binds elements of a misfolded intron, which is chaperoned into an active conformation that is in turn stabilized by the maturase, to generate a maturase-dependent mobile intron. (Black lines) Ribozyme RNA; (red lines) endonuclease-maturase-encoding RNA.

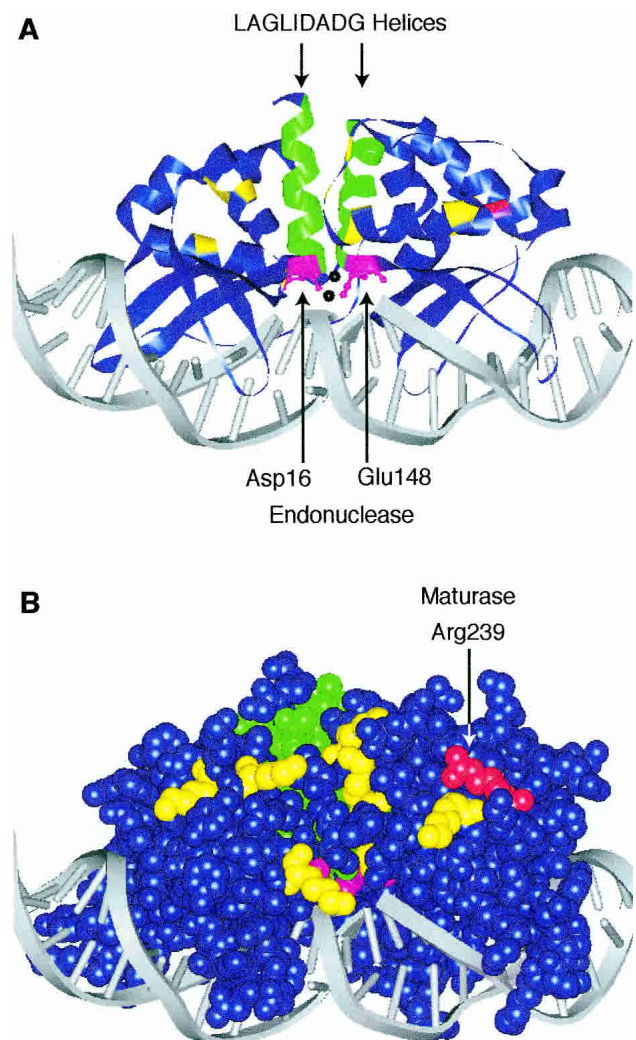


Figure 3. Pseudosymmetric structure of I-AniI on DNA homing site. (A) Ribbon diagram. (B) Space-filling model. View in B is similar to that in A. In both A and B, the α -helices containing LAGLIDADG motifs are colored green. The catalytic endonuclease residues Asp 16 and Glu 148 are shown in magenta, and Arg 239, a key amino acid in maturase function, is shown in red. Two bound Mg^{2+} ions are depicted as black spheres. Positively-charged surface residues conserved in closely related proteins are highlighted in yellow.

The action of I-AniI maturase

Although we already have detailed mechanistic insights into LAGLIDADG endonuclease function (for review, see Chevalier and Stoddard 2001), our understanding of maturase activity is in its infancy. However, we do know that the RNA-I-AniI ribonucleoprotein (RNP) complex forms with a 1:1 stoichiometry (Solem et al. 2002; Geese et al. 2003), and that I-AniI seems to recognize broad features of RNA structure. Also, it appears that I-AniI binds loosely to an intron with limited tertiary structure, and then facilitates formation of the active fold, to yield a stable RNP complex (Fig. 2; Ho and Waring 1999; Solem et al. 2002).

Proteins that facilitate ribozyme function by binding to the RNA have been classified into two broad categories. First, the RNA chaperones bind RNA nonspecifically and guide RNA folding by destabilizing misfolded RNA structures that constitute kinetic traps (Herschlag 1995). RNA chaperones are not required to remain bound to RNA to facilitate splicing. Second, non-intron-encoded splicing accessory factors stabilize catalytically active structures by tertiary-structure capture and RNA scaffolding (for review, see Lambowitz et al. 1999; discussed below). These proteins, which tend not to be specific to a particular intron, must stay bound to the RNA in an RNP complex to promote splicing. As a result of the above observations from the Caprara and Waring laboratories, it would appear that I-AniI has features of both an RNA chaperone, facilitating the correct folding of the intron, and a splicing accessory, remaining complexed to the intron to stabilize the active conformation (Fig. 2; Ho and Waring 1999; Solem et al. 2002). Thus, a naive surface, distinct from the DNA-binding cleft of I-AniI, seems to have evolved to bind intron RNA, thereby limiting the conformational options of the intron and locking it into its appropriate configuration to self-splice.

Evolution of function

There is a growing list of bifunctional moonlighting proteins that act on DNA and RNA (Cassiday and Maher 2002; Jeffrey 2003). Some, such as thymidylate synthase and T4 DNA polymerase, perform specific chemistry on DNA while acting as translational repressors by blocking access to other molecules when they bind RNA. Like these DNA processing enzymes, I-AniI performs chemical catalysis on DNA and has evolved recognition specificity for RNA, with its maturase function apparently deriving from its novel binding potential, rather than from *de novo* chemistry. Therefore, it will be of interest to determine how different intron-specific maturases have tailored their binding properties and adapted to the functional needs of individual introns.

What selection might be acting on these intron-encoded proteins? On one hand, intron dissemination demands maintenance of endonuclease activity. On the other hand, there is selective pressure on the maturase to promote efficient splicing. In some cases, such pressure must be extreme, as for a maturase-dependent yeast mitochondrial intron that has recruited an additional cellular protein to collaborate with the intron-encoded LAGLIDADG protein to form a multicomponent RNP (Bassi et al. 2002). This nuclear-encoded protein, Mrs1, is a member of the RNase H-fold superfamily, and bears resemblance to DNA junction-resolving enzymes. There are yet other examples of cellular functions, like tRNA synthetases, that serve as splicing factors (for review, see Lambowitz et al. 1999). The best-understood of these is CYT18, tyrosyl tRNA synthetase, which doubles as a splicing accessory for several mitochondrial introns in *Neurospora crassa*, to stabilize the catalytically active RNA structure. CYT18 also has a collaborator, CYT19,

which is a DEAD-box protein that can bind RNA non-specifically and act as an ATP-dependent RNA chaperone (Mohr et al. 2002). CYT19 is thought to disrupt non-productive intron folding, thereby cooperating with CYT18 to effect splicing. This disparate array of intron-specific maturases and more general splicing accessory proteins speaks to the evolutionary pressure to maintain efficient splicing, for preservation of gene function, and to the varied solutions to promoting RNA catalysis.

Interestingly, homing endonucleases fall into highly conserved structural families characterized by different sequence motifs (for review, see Belfort and Roberts 1997; Belfort et al. 2002), whereas no conserved maturase features have yet been reported. This observation is consistent with the mechanisms of RNA maturase function being more diverse than the restricted catalytic function of homing endonucleases. Nevertheless, the plasticity of these structurally and functionally constrained DNA endonucleases to expand DNA target-site recognition (Chevalier et al. 2003), and to act on RNA, is not to be underestimated. The ability of enzymes like I-AniI to acquire RNA maturase activity allows the endonuclease gene and its host intron to be cemented in one genome, whereas development of an extended repertoire of DNA cleavage sites ensures their spread to others.

Acknowledgments

Thanks to Phil Perlman of Southwestern Medical Center, and Wadsworth colleagues Lori Conlan, Vicky Derbyshire, David Edgell, and Janice Pata for their thoughtful comments on the manuscript. The expert assistance of Maryellen Carl and John Dansereau in manuscript and figure preparation is greatly appreciated. Work in the author's laboratory is supported by grants from NIH (GM39422 and GM44844) and NSF (CTS-0304055).

References

- Bassi, G.S., de Oliveira, D.M., White, M.F., and Weeks, K.M. 2002. Recruitment of intron-encoded and co-opted proteins in splicing of the bI3 group I intron RNA. *Proc. Natl. Acad. Sci.* **99**: 128–133.
- Belfort, M. and Roberts, R.J. 1997. Homing endonucleases: Keeping the house in order. *Nucleic Acids Res.* **25**: 3379–3388.
- Belfort, M., Derbyshire, V., Cousineau, B., and Lambowitz, A. 2002. Mobile introns: Pathways and proteins. In *Mobile DNA II* (eds. N. Craig, R. Craigie, M. Gellert, and A. Lambowitz), pp. 761–783. ASM Press, Washington, DC.
- Bolduc, J.M., Spiegel, P.C., Chatterjee, P., Brady, K.L., Downing, M.E., Caprara, M.G., Waring, R.B., and Stoddard, B.L. 2003. Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor. *Genes & Dev.* (this issue).
- Cassiday, L.A. and Maher III, J. 2002. Survey and summary. Having it both ways: Transcription factors that bind DNA and RNA. *Nucleic Acids Res.* **30**: 4118–4126.
- Chatterjee, P., Brady, K.L., Solem, A., Ho, Y., and Caprara, M.G. 2003. Functionally distinct nucleic acid binding sites for a group I intron encoded RNA maturase/DNA homing endonuclease. *J. Mol. Biol.* **329**: 239–251.
- Chevalier, B.S. and Stoddard, B.L. 2001. Homing endonucleases: Structural and functional insight into the catalysis of intron/intron mobility. *Nucleic Acid Res.* **29**: 3757–3774.
- Chevalier, B., Turmel, M., Lemieux, C., Monnat, R.J.J., and Stoddard, B.L. 2003. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J. Mol. Biol.* **329**: 253–269.
- Delahodde, A., Goguel, V., Becam, A.M., Creusot, F., Perea, J., Banroques, J., and Jacq, C. 1989. Site-specific DNA endonuclease and RNA maturase activities of two homologous intron-encoded proteins from yeast mitochondria. *Cell* **56**: 431–441.
- Dujon, B. 1989. Group I introns as mobile genetic elements: Facts and mechanistic speculations—a review. *Gene* **82**: 91–114.
- Geese, W.J., Kwon, Y.K., Wen, X., and Waring, R.B. 2003. *In vitro* analysis of the relationship between endonuclease and maturase activities in the bi-functional group I intron-encoded protein, I-AniI. *Eur. J. Biochem.* **270**: 1543–1554.
- Henke, R.M., Butow, R.A., and Perlman, P.S. 1995. Maturase and endonuclease functions depend on separate conserved domains of the bifunctional protein encoded by the group I intron al4 α of yeast mitochondrial DNA. *EMBO J.* **14**: 5094–5099.
- Herschlag, D. 1995. RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **270**: 20871–20874.
- Ho, Y. and Waring, R.B. 1999. The maturase encoded by a group I intron from *Aspergillus nidulans* stabilizes RNA tertiary structure and promotes rapid splicing. *J. Mol. Biol.* **292**: 987–1001.
- Ho, Y., Kim, S.J., and Waring, R.B. 1997. A protein encoded by a group I intron in *Aspergillus nidulans* directly assists RNA splicing and is a DNA endonuclease. *Proc. Natl. Acad. Sci.* **94**: 8994–8999.
- Jeffrey, C.J. 2003. Multifunctional proteins: Examples of gene sharing. *Ann. Med.* **35**: 28–35.
- Lambowitz, A.M. and Perlman, P.S. 1990. Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. *Trends Biochem. Sci.* **15**: 440–444.
- Lambowitz, A.M. and Belfort, M. 1993. Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62**: 587–622.
- Lambowitz, A.M., Caprara, M.G., Zimmerly, S., and Perlman, P.S. 1999. Group I and group II ribozymes as RNPs: Clues to the past and guides to the future. In *The RNA world*, 2nd ed., pp. 451–485. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Lazowska, J., Jacq, C., and Slonimski, P.P. 1980. Sequence of introns and flanking exons in wild-type and *box3* mutants of cytochrome *b* reveals an interlaced splicing protein coded by an intron. *Cell* **22**: 333–348.
- Mohr, S., Stryker, J.M., and Lambowitz, A.M. 2002. A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. *Cell* **109**: 769–779.
- Perlman, P.S. and Butow, R.A. 1989. Mobile introns and intron-encoded proteins. *Science* **246**: 1106–1109.
- Solem, A., Chatterjee, P., and Caprara, M.G. 2002. A novel mechanism for protein-assisted group I intron splicing. *RNA* **8**: 412–425.
- Szczepanek, T., Jamoussi, K., and Lazowska, J. 2000. Critical base substitutions that affect the splicing and/or homing activities of the group I intron bi2 of yeast mitochondria. *Mol. Gen. Genet.* **264**: 137–144.
- Wenzlau, J.M., Saldanha, R.J., Butow, R.A., and Perlman, P.S. 1989. A latent intron-encoded maturase is also an endonuclease needed for intron mobility. *Cell* **56**: 421–430.