JOURNAL OF BACTERIOLOGY, July 1995, p. 3897–3903 Vol. 177, No. 14
0021-9193/95/$04.00+0

MINIREVIEW

Prokaryotic Introns and Inteins: a Panoply of Form and Function

MARLENE BELFORT, MARY E. REABAN, TIMOTHY COETZEE,† AND JACOB Z. DALGAARD‡

Molecular Genetics Program, Wadsworth Center and School of Public Health, State University of New York at Albany, New York State Department of Health, P.O. Box 22002, Albany, New York 12201-2002

INTRODUCTION

Following their initial discovery, introns and RNA splicing were considered, along with the nuclear envelope, as characteristics that distinguish eukaryotes from prokaryotes (the eubacteria, referred to simply as bacteria, and the archaebacteria, referred to as archaea). This dogma became shaky with the identification of putative introns in tRNA genes of archaea (35) and finally crumbled with the discovery of self-splicing group I introns in the phages of purple bacteria (9), which was followed by the discovery of group I (40, 59, 72) and group II introns (24) in bacterial cells. Another breakthrough was the recent discovery of a nuclear pre-mRNA-like intron in a bacterial plasmid (48). The variety of intervening sequences (IVSs) that span the phylogenetic spectrum was further extended by the discovery of inteins, elements that can splice at the protein level, in eukaryotes (36), archaea (56, 71), and bacteria (19, 20). This minireview will illustrate that interrupted genes can no longer be considered the province of eukaryotes, but rather that many forms of IVSs are phylogenetically diverse, occurring also in bacterial and archaeal genomes (Fig. 1).

GROUP I INTRONS

The first intron to be discovered in a prokaryote was a group I intron, located in the thymidylate synthase gene of bacteriophage T4 (9). Subsequently, group I introns were found in other T4 coliphage genes (29) as well as in *B. subtilis* phage genes (1, 28) and in tRNA genes of cyanobacteria and proteobacteria (40, 59, 72). Additionally, group I introns were found in the genomes of both chloroplasts and mitochondria, which are endosymbiotic descendents of the cyanobacteria and proteobacteria, respectively (reviewed in reference 3). As with the canonical group I intron of *Tetrahymena thermophila* (8), many of these introns can self-splice in vitro in the absence of proteins.

Group I introns are spliced from the precursor RNA by a series of transesterification reactions initiated by an exogenous guanosine cofactor, resulting in ligated exons and a free intron bearing the initiating guanosine at its 5′ end (Fig. 2A) (reviewed in references 8 and 60). This latter feature of the reaction has been widely exploited to identify group I introns: bacterial RNA extracts were incubated with [α-^{32}P]GTP and then screened for radiolabeled introns (29). It is therefore not surprising that the majority of phage and eubacterial group I introns discovered thus far have been shown to be self-splicing in vitro. Nevertheless, it appears that certain proteins that bind RNA nonspecifically, such as the *E. coli* ribosomal protein S12, may provide chaperone-like activity to promote proper folding of the RNA (11). This contrasts with the many fungal mitochondrial group I introns that require specific protein splicing factors, some encoded by nuclear genes, others (maturases) encoded by the introns themselves (7, 42).

Despite a lack of primary sequence conservation, group I introns share common short- and long-range pairings, P1 through P9 (Fig. 2A), which form the basis of a conserved, catalytically active tertiary structure (49, 51). In addition, many group I introns have open reading frames (ORFs) looped out of these conserved structures (reviewed in reference 53). Several of these ORFs encode endonucleases which play a role in intron mobility, as is described below.

GROUP II INTRONS

Group II introns have been recently identified in cyanobacteria (*Calothrix* spp.) and proteobacteria (*Azotobacter vinelandii*), the respective progenitors of chloroplasts and mitochondria (24), and in a bacterium close to the hearts of many, *Escherichia coli* (23, 39). These discoveries are particularly intriguing from an evolutionary standpoint, for group II introns had previously been known to exist only in chloroplasts and mitochondria of fungi and plants and because the group II splicing mechanism is similar to that of nuclear pre-mRNA introns. Indeed, it is widely held that group II introns served as the ancestors of nuclear pre-mRNA introns (62).

As with nuclear pre-mRNA splicing, group II splicing initiates when the 2′ OH of a bulged nucleotide attacks the 5′ splice site, resulting in formation of a 2′-5′ bond and a lariat structure (Fig. 2B). Subsequently, the terminal OH of the 5′ exon attacks the 3′-splice site, resulting in ligation of the exons and release of the intron lariat (reviewed in references 8 and 60). Like group I introns, group II introns adopt a conserved secondary structure that is essential for splicing (50). This structure consists of six conserved helices that emanate from a central wheel (Fig. 2B). Sequence conservation is confined to short strings of intron residues and the bulged A in domain VI that forms the lariat branch site.

Although some group II introns can self-splice, they do so only under nonphysiological conditions, and many require protein cofactors, some of which are encoded by the intron (reviewed in reference 42). The products of these internal ORFs can act as maturases to facilitate splicing or as reverse transcriptases (RTs) to promote group II intron mobility (41, 60).

A NUCLEAR PRE-mRNA-LIKE INTRON

An Ri plasmid gene of *Agrobacterium rhizogenes* has been found to contain an intron in its untranslated leader region.

† Present address: Brain and Development Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7250.
‡ Present address: Frederick Cancer Research and Development Center, P.O. Box B, Frederick, MD 21702-1201.

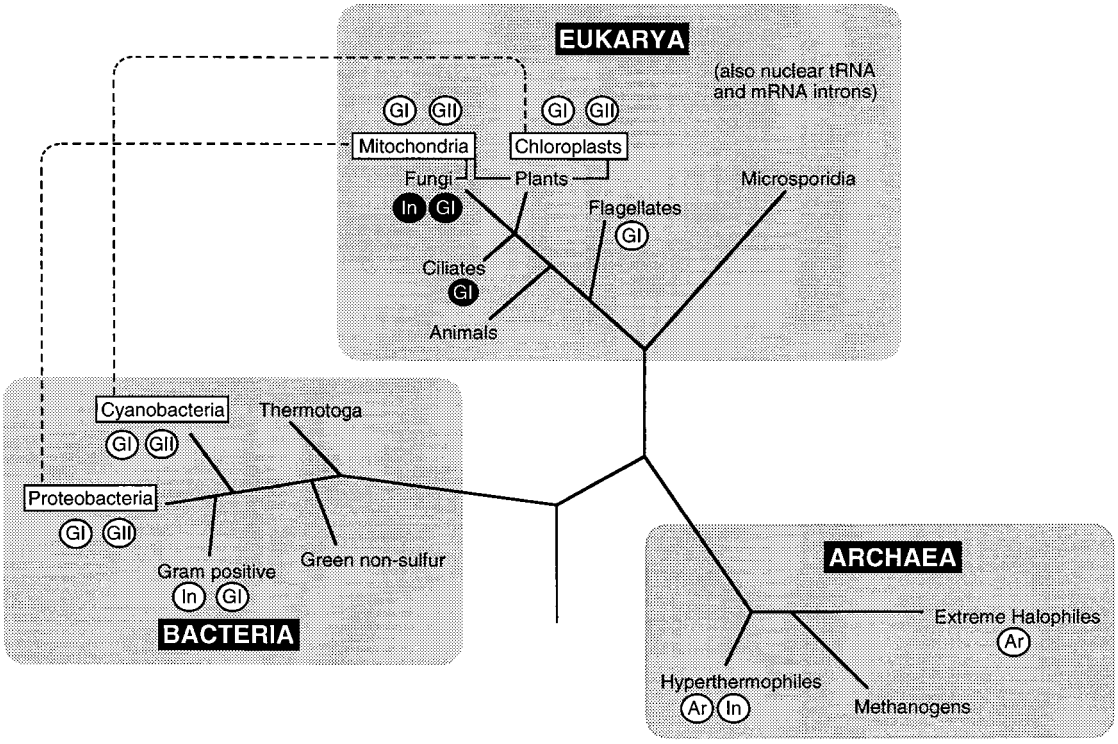

FIG. 1. Phylogenetic distribution of introns and inteins. The distribution of introns and inteins is shown on a phylogenetic tree of the three primary kingdoms (modified from reference 69). IVS types represented in circles are as follows: GI, group I; GII, group II; Ar, archaeal; In, intein. Those IVSs represented by filled circles are nuclear, while the others are prokaryotic or organellar. The group I introns in gram-positive bacteria are in phage genomes. See reference 46 for the flagellate mitochondrial group I intron; see text for other references.

Surprisingly, this intron has the characteristic features of a plant nuclear pre-mRNA intron: the dinucleotides GT at the 5′ end and AG at the 3′ end of the intron, a T-rich region preceding the 3′ splice site, the appropriate AT content, and homology to the plant consensus sequence at the 5′ exon-intron junction (48). Transcription and splicing take place in transformed *Arabidopsis thaliana* plant cells, and it is therefore assumed that the plant spliceosomal machinery effects processing of the bacterial pre-mRNA.

ARCHAEAL INTRONS

A single class of introns, unrelated to any of the aforementioned ones, has been described in the archaeal kingdom. Unique to this phylogenetic domain, they have been found in tRNA genes of halophiles and in tRNA and rRNA genes of hyperthermophiles (18, 25). Archaeal introns vary in size upwards of 15 nucleotides, are catalytically inert, and share a structural motif only at the intron-exon boundaries (38, 64, 65) (Fig. 2C). This motif, a bulge-helix-bulge structure, is recognized by a protein enzyme, an RNA endoribonuclease, that excises the intron from the precursor RNA by cleavage in each of the two bulges. The cleavage event creates 2′-3′ cyclic phosphates and 5′ hydroxyl termini, a characteristic of nuclear tRNA introns of eukaryotes. A poorly defined ligation reaction creates a normal 5′-3′ phosphodiester bond joining the exons (37, 38, 64). Several archaeal introns possess ORFs, some of which encode mobility-type endonucleases (16, 17, 47). Remarkably, these archaeal enzymes contain the LAGLIDADG amino acid motif, a conserved feature of DNA endonucleases encoded by eukaryotic group I introns and by inteins (reviewed in reference 53).

OTHER RNA INTERRUPTIONS

The 23S and 16S rRNA genes of diverse bacterial species contain small (100 to 200 bp) IVSs, which are excised from the RNA without exon ligation, resulting in fragmented rRNAs (6, 44, 66). Analogous bacterial 23S rRNA IVSs that contain ORFs of 121 to 133 amino acids have also been found (58). The ability of these fragmented rRNAs to assemble into an active conformation appears to obviate the need for exon ligation. These IVSs have the potential to fold into similar stem-loop structures; the *Salmonella* IVS has been shown to be excised by RNAse III, which has a preference for such duplex conformations (6). The sporadic distribution of these sequences among closely related organisms, and even among the multiple rRNA gene copies of the same organism, plus the presence of ORFs in some, suggest that these IVSs too are mobile, having invaded rRNA genes.

Yet another type of IVS is found in bacteriophage T4, in gene 60, encoding a topoisomerase subunit (32, 68). This 50-nt, atypical IVS is neither spliced nor excised. Instead, it is bypassed during translation by a gigantic frameshift event. Several features have been identified as being required for ribosomal hopping: a *cis*-acting 16-amino-acid segment of the nascent peptide, matching "takeoff" and "landing" codons optimally separated by 50 nt, a stem-loop structure at the takeoff site, and a stop codon immediately following the takeoff site. Ribosomal protein L9 also appears to play a role in the process (30). The *E. coli trpR* gene also appears to produce a minor product via a frameshift event that bypasses 55 nt (5). However, the mechanism of the *trpR* frameshifting event differs from that of T4 gene 60: there is no requirement for matching

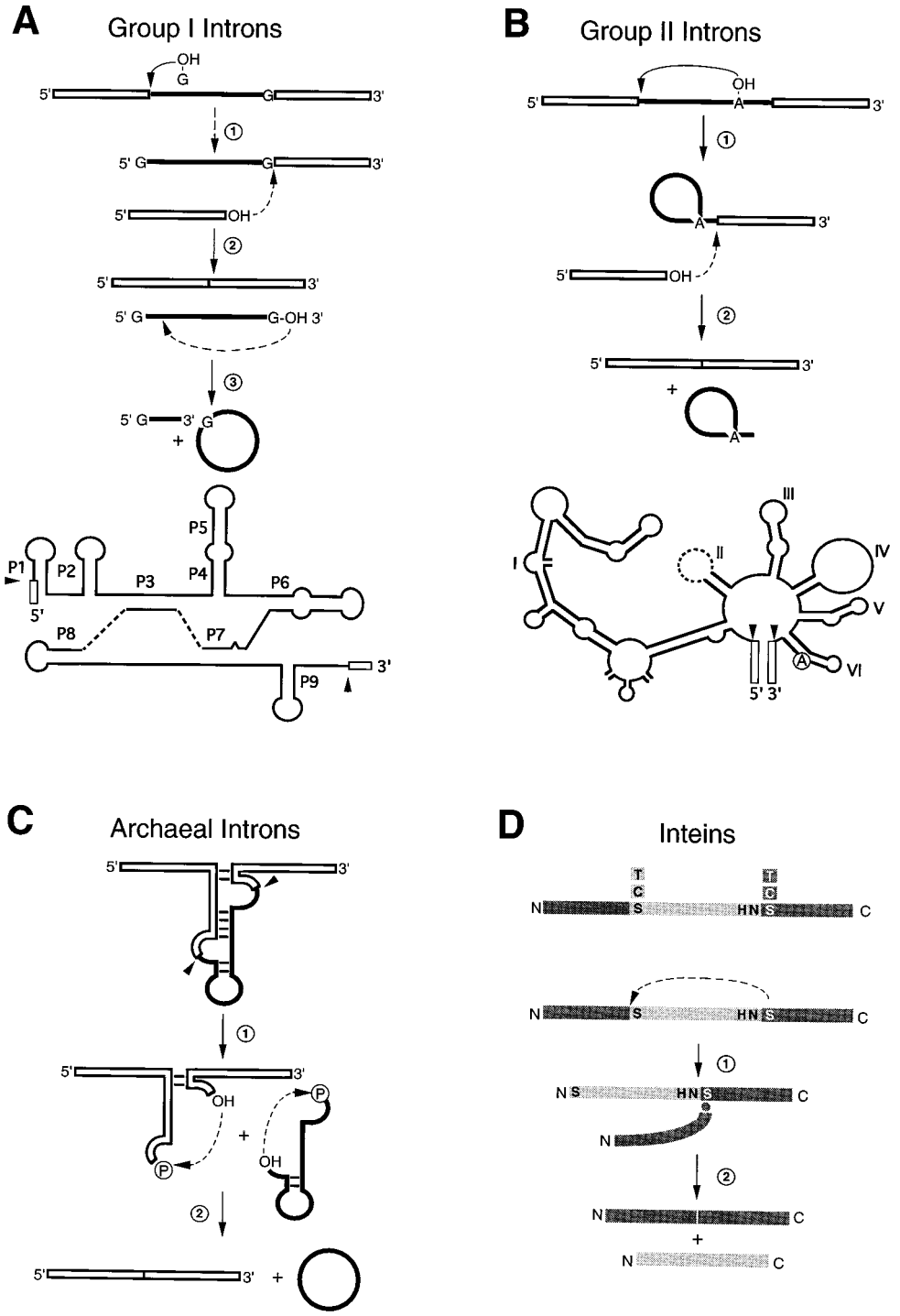

FIG. 2. Splicing pathways and requisite structures. (A) Group I introns. Step 1, nucleophilic attack at the 5′ splice site by the 3′ OH of an exogenous guanosine cofactor. Step 2, attack at the 3′ splice site by the 3′ OH of the 5′ exon, resulting in exon ligation and release of the intron. Step 3, a third, optional, transesterification reaction between the 3′ OH of the intron and an internal site can result in cyclization. Exons are represented by boxes, and introns are represented by lines. The secondary structure is shown below, with exons depicted as boxes and splice sites indicated by arrowheads. (B) Group II introns. Step 1, attack at the 5′ splice site by the 2′ OH of an adenosine near the 3′ end of the intron results in formation of a 2′-5′ bond. Step 2, the 3′ OH of the 5′ exon attacks the 3′ splice site resulting in exon ligation and release of the intron lariat. The secondary structure is shown below, with conserved pairings designated I through VI, and with the bulged A in helix VI. (C) Archaeal introns. Step 1, endonucleolytic cleavage at the bulge-helix-bulge motif, creating 2′,3′ cyclic phosphates (circles). Step 2, exon ligation and optional cyclization. (D) Inteins. Conserved amino acids are indicated at the junction of inteins (light shading) and exteins (dark shading). Step 1, protein splicing is initiated either by nucleophilic attack on the carbonyl group of the upstream splice junction by the serine hydroxyl at the downstream junction or by an N-O acyl shift in the upstream serine, followed by attack and transesterification by the downstream serine. Cysteine and threonine could replace serine in any of these reactions. Both pathways yield a branched intermediate containing an alkali-labile ester bond linking the N-extein to the intein-C-extein. Step 2, the intein is freed by cyclization of the conserved asparagine to form succinamide, and the ester linkage between the two exteins undergoes an O-N acyl shift, producing a standard linkage between the two exteins (70).

takeoff and landing codons or for a stop codon, and ribosomal hopping is less efficient by an order of magnitude.

INTEINS—PROTEIN SPLICING ELEMENTS

Another dogma-breaking discovery was of genes which contain in-frame inserts that are removed at the protein, rather than at the RNA, level. The intervening polypeptide (intein) is excised precisely from the precursor protein, and the flanking polypeptides (exteins) are ligated to form the mature protein (12, 14, 57). Inteins are widespread (Fig. 1), having been found in the nuclear genes for a vacuolar ATPase subunit in the eukaryotes *Saccharomyces cerevisiae* and *Candida tropicalis*, in the DNA polymerase genes of the archaea *Thermococcus litoralis* and *Pyrococcus* sp. strain GB-D, and in the *recA* genes of the eubacteria *Mycobacterium leprae* and *M. tuberculosis* (compiled in reference 57).

There are several conserved amino acids at the intein splice junctions, including a histidine-asparagine pair at the C terminus of the intein and serine, cysteine, or threonine residues downstream of both splice junctions (compiled in reference 71). Several models for the molecular mechanism of protein splicing have been proposed (13, 67, 70, 71), but the details have not yet been resolved. Figure 2D shows the most recent model, in which an alkali-labile branched intermediate is resolved by cyclization of the C-terminal asparagine of the intein (71). In addition to these remarkable properties, all known inteins contain the LAGLIDADG motifs characteristic of homing endonucleases, and some have been shown to possess analogous nucleolytic activities (53).

INTRON AND INTEIN MOBILITY

Some of the group I intron ORFs have been found to encode double-stranded DNA endonucleases, which recognize and cleave a cognate intronless allele at or near the site of intron insertion. It has been shown for several of the phage and fungal mitochondrial introns that repair of this double-strand break results in the incorporation of the intron, including its ORF (Fig. 3A). Thus, the ORF confers mobility on its host intron. This process is referred to as homing, and the endonucleases are referred to as homing endonucleases (reviewed in reference 41). Homing is a highly efficient gene conversion event and involves coconversion of the adjacent exon markers, which results from exonucleolytic degradation of the cleaved recipient. Interestingly, the *S. cerevisiae* vacuolar ATPase intein ORF has been shown to home into inteinless alleles by the same pathway, initiated by the intein's endonuclease activity (26).

Although endonucleases with LAGLIDADG motifs similar to those in the eukaryotic homing endonucleases and inteins have been found in archaeal introns, homing remains to be demonstrated in this system. Nevertheless, the discovery of an endonuclease ORF related to the group I intron ORFs in a different type of intron (16) fuelled arguments that the endonuclease ORF is the true mobile element (41). By finding refuge in a splicing element, host viability would be maintained, while the intron acquires mobile properties. This scenario has been further supported by the demonstration that endonuclease recognition and cleavage sites lie hidden within some introns. The occurrence of such sites which are split by the ORF lends credence to the hypothesis that the endonuclease-ORF invaded the intron as the primary agent of mobility (45).

Mobility of group II introns is mechanistically different from that of group I introns (Fig. 3B). These introns are capable of homing (43, 52) as well as transposition reactions (54, 61), both of which appear to be mediated by the conserved intron-encoded RTs. Group II intron homing, which thus far has been demonstrated only for yeast mitochondrial introns, is still a poorly defined process. While homing involves coconversion of flanking exon markers, implicating exon sequences in the mobility event, the process has a remarkable requirement for splicing proficiency of the intron. Thus, although intron cDNA for homing is likely to be derived from pre-mRNA, splicing or a competent intron structure is needed. Two proposals advanced for this requirement are as follows. (i) An RNA structure may be required for cDNA synthesis by RT (52). (ii) There may be a role for the intron in endonuclease function required in the cDNA integration step (4, 52).

Recent evidence favors group II intron transposition to nonallelic sites through RNA intermediates (54, 61). The most likely pathway is via reverse splicing into heterologous RNA. Subsequent reverse transcription and homologous recombination into the heterologous site would fix the intron at the nonallelic locus on the genome (Fig. 3C). Theoretically, group I introns could move by a similar mechanism. Although no group I introns are known to contain cognate RTs, and no complete transposition events have yet been reported, the self-splicing group I intron in the *Chlamydomonas reinhardtii* chloroplast 23S rRNA gene has been found to undergo the first step of reverse splicing into the host cytoplasmic 5.8S rRNA (63).

Although the self-propelled mobility pathways of group II introns have not yet been observed in bacteria, the occurrence of RT-ORFs in these introns raises expectations of their mobility. Furthermore, like the RT-containing bacterial retrons, which have been discovered both in freestanding form and in prophages (34), some group II introns are contained within DNAs that have the hallmarks of mobile elements (23, 39). These findings suggest that group II introns have superimposed mobility mechanisms. They have the potential to home or transpose by virtue of their built-in RTs, while mobility may also be conferred by the putative transposition ability of the elements in which they reside.

INTRONS AND INTEINS—THE WHYS AND WHEREFORES

Why are introns and inteins present at all in prokaryotes? The streamlined nature of prokaryotic genomes, with their pressure to shed excess DNA to effect rapid replication, is a major force at work to eliminate them. Furthermore, the physical constraints of the prokaryotic environment may be restrictive: uninfected, wild-type *E. coli* cells have been shown to lack the requisite accessory functions for homing (10), and coupling of transcription and translation may interfere with splicing, thus decreasing the tolerance of introns. Finally, opportunities for dissemination are limited in prokaryotes, compared with eukaryotes, which enjoy efficient sexual reproduction (31, 33).

Countering these eradicative pressures is the sheer invasiveness of these elements. Introns and inteins can be considered selfish DNAs, and, as has been proposed for transposable elements and nuclear mRNA introns (21, 55), may survive by their ability to invade and multiply without significantly compromising their host. In fact, inteins would represent the ultimate parasites, having the self-contained ability to invade a gene, independent of the splicing function of a host intron. The fact that homing has evolved numerous times independently for different kinds of IVSs suggest that mobility is a strong evolutionary driving force and that their parasitic nature wins out over forces to eliminate them. As an additional incentive to

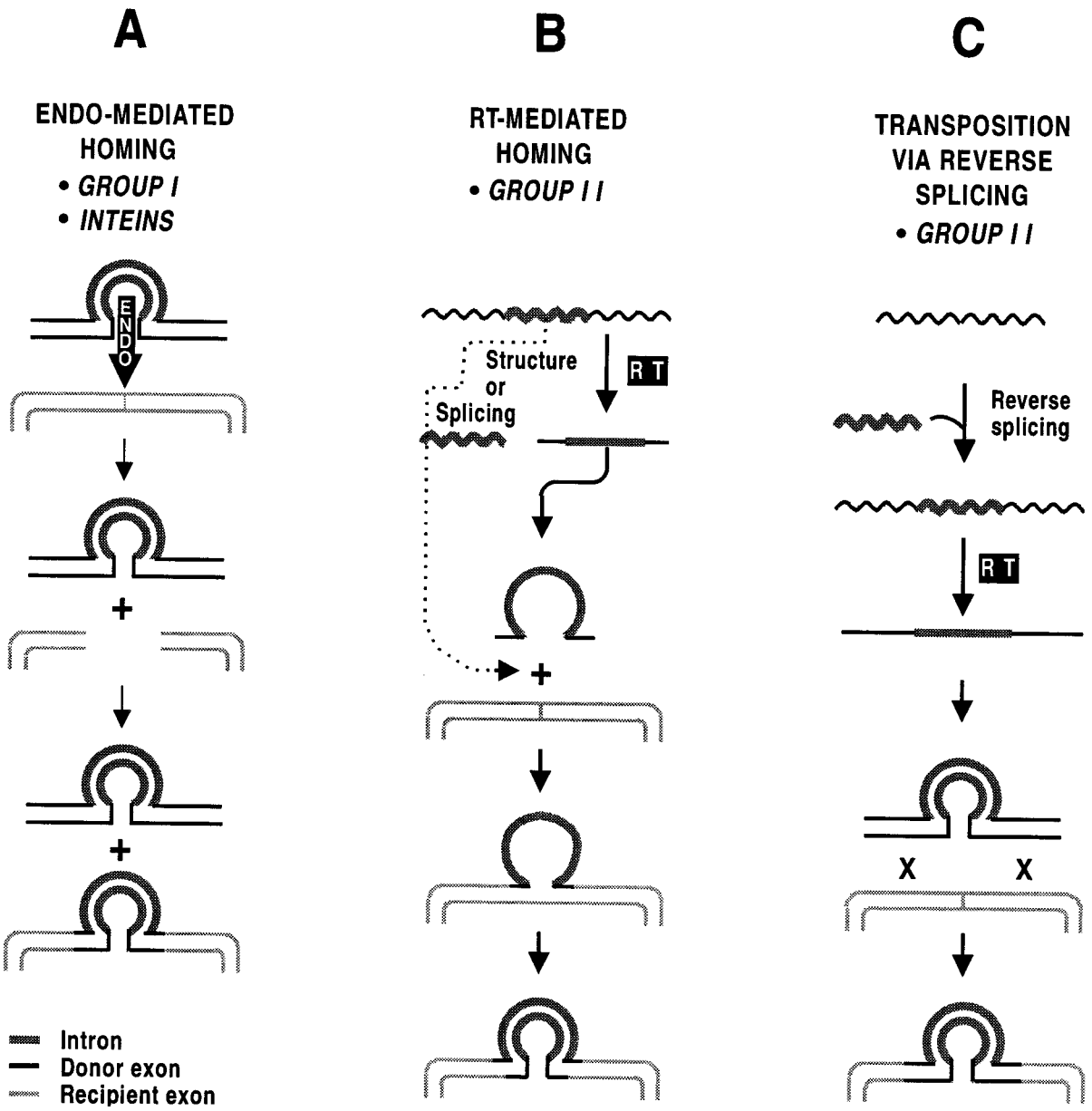

FIG. 3. Intron mobility pathways. Wavy lines, RNA; straight lines, DNA; thick lines, introns; thin lines, exons. (A) Group I intron and intein homing through repair of double-stranded DNA breaks. Endonuclease-mediated cleavage of the recipient alleles stimulates a gene conversion event that results in intron or intein inheritance in a DNA-based pathway. (B) Group II intron homing through an RT-mediated pathway. Here, pre-mRNA is postulated to act as a template for cDNA synthesis by RT. The role of intron structure or splicing (dashed line) and the mechanism of site-specific integration of the cDNA remain conjectural. (C) Intron transposition by reverse splicing. Although reverse splicing is depicted here with RNA as the recipient, it is also mechanistically possible for DNA to be the recipient. The cDNA is then proposed to act as a recombination substrate with genomic DNA in double-stranded (as shown) or single-stranded form. (Adapted from reference 4, with permission from the publisher.)

retain them, the host organism may find some aspect of the intron/intein beneficial, or the IVS may evolve properties beneficial to their host and/or themselves (2, 15).

A popular theory is that the self-splicing introns are molecular fossils from the primitive RNA world. This hypothesis is fueled by demonstrations that group I introns can function as replicases (22) and is consistent with the presence of group I and group II introns in bacteria. These introns, specifically the group II introns, are postulated to have given rise to the nuclear pre-mRNA introns prevalent in modern metazoans (62). The discovery of a nuclear pre-mRNA-like intron in an agrobacterium (48) provides a provocative conceptual bridge between the bacterial self-splicing group II introns and the spliceosome-dependent pre-mRNA introns in eukaryotes.

Although neither the catalytic group I and group II introns nor the pre-mRNA introns have yet been found in archaea, archaeal tRNA introns are similar in both position and splicing pathway to tRNA introns of eukaryotes (reviewed in reference 41). It is conceivable that the relatively close phylogenetic relationship with the archaea may explain the source of tRNA introns in eukarya. These speculations on the origins of eukaryotic nuclear pre-mRNA and tRNA introns are consistent with the postulated chimeric nature of eukarya, as organisms which arose from a fusion between members of the bacterial

and archaeal kingdoms (27). Thus, introns are not only of functional interest to the molecular biologist, but their occurrence and properties continue to feed into theories of molecular evolution and the origins of biological diversity.

ACKNOWLEDGMENTS

We are thankful to Maryellen Carl for manuscript preparation, Maureen Belisle for illustrations, and Mary Bryk, Vicky Derbyshire, Joyce Huang, Monica Parker, and Elisabeth Tillier for comments on the manuscript.

Work in the laboratory of M.B. is funded by NIH grants GM39422 and GM44844. M.E.R. is supported by NIH postdoctoral grant GM14915.

REFERENCES

1. **Bechhofer, D. H., K. K. Hue, and D. A. Shub.** 1994. An intron in the thymidylate synthase gene of *Bacillus* bacteriophage β22: evidence for independent evolution of a gene, its group I intron, and the intron open reading frame. Proc. Natl. Acad. Sci. USA **91:**11669–11673.
2. **Belfort, M.** 1989. Bacteriophage introns: parasites within parasites? Trends Genet. **5:**209–213.
3. **Belfort, M.** 1991. Self-splicing introns in prokaryotes: migrant fossils? Cell **64:**9–11.
4. **Belfort, M.** 1993. An expanding universe of introns. Science **262:**1009–1010.
5. **Benhar, I., and H. Engelberg-Kulka.** 1993. Frameshifting in the expression of the E. coli *trpR* gene occurs by the bypassing of a segment of its coding sequence. Cell **72:**121–130.
6. **Burgin, A. B., K. Parodos, D. J. Lane, and N. R. Pace.** 1990. The excision of intervening sequences from Salmonella 23S ribosomal RNA. Cell **60:**405–414.
7. **Burke, J. M.** 1988. Molecular genetics of group I introns: RNA structures and protein factors required for splicing—a review. Gene **73:**273–294.
8. **Cech, T. R.** 1993. Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P, p. 239–269. *In* R. F. Gesteland and J. F. Atkins (ed.), The RNA world. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
9. **Chu, F. K., G. F. Maley, F. Maley, and M. Belfort.** 1984. Intervening sequence in the thymidylate synthase gene of bacteriophage T4. Proc. Natl. Acad. Sci. USA **81:**3049–3053.
10. **Clyman, J., and M. Belfort.** 1992. *trans* and *cis* requirements for intron mobility in a prokaryotic system. Genes Dev. **6:**1269–1279.
11. **Coetzee, T., D. Herschlag, and M. Belfort.** 1994. *E. coli* proteins, including ribosomal protein S12, facilitate in vitro splicing of phage T4 introns by acting as RNA chaperones. Genes Dev. **8:**1575–1588.
12. **Colston, M. J., and E. O. Davis.** 1994. The ins and outs of protein splicing elements. Mol. Microbiol. **12:**359–363.
13. **Cooper, A. A., Y. Chen, M. A. Lindorfer, and T. H. Stevens.** 1993. Protein splicing of the yeast *TFP1* intervening protein sequence: a model for self-excision. EMBO J. **12:**2575–2583.
14. **Cooper, A. A., and T. H. Stevens.** 1993. Protein splicing: excision of intervening sequences at the protein level. Bioessays **15:**667–674.
15. **Dalgaard, J. Z.** 1994. Mobile introns and inteins: friend or foe? Trends Genet. **10:**306–307.
16. **Dalgaard, J. Z., R. A. Garrett, and M. Belfort.** 1993. A site-specific endonuclease encoded by a typical archaeal intron. Proc. Natl. Acad. Sci. USA **90:**5414–5417.
17. **Dalgaard, J. Z., R. A. Garrett, and M. Belfort.** 1994. Purification and characterization of two forms of I-*Dmo*I, a thermophilic site-specific endonuclease encoded by an archaeal intron. J. Biol. Chem. **269:**28885–28892.
18. **Daniels, C. J., R. Gupta, and W. F. Doolittle.** 1985. Transcription and excision of a large intron in the tRNA gene of an archaebacterium, *Halobacterium volcanii*. J. Biol. Chem. **260:**3132–3134.
19. **Davis, E. O., P. J. Jenner, P. C. Brooks, M. J. Colston, and S. G. Sedgwick.** 1992. Protein splicing in the maturation of the *M. tuberculosis* RecA protein: a mechanism for tolerating a novel class of intervening sequence. Cell **71:**201–210.
20. **Davis, E. O., H. S. Thangaraj, P. C. Brooks, and M. J. Colston.** 1994. Evidence of selection for protein introns in the RecAs of pathogenic mycobacteria. EMBO J. **13:**699–703.
21. **Doolittle, W. F., and C. Sapienza.** 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature (London) **284:**601–603.
22. **Doudna, J. A., and J. W. Szostak.** 1989. RNA-catalysed synthesis of complementary-strand RNA. Nature (London) **339:**519–522.
23. **Ferat, J., M. L. Gouar, and F. Michel.** 1994. Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. Life Sci. **317:**141–148.
24. **Ferat, J., and F. Michel.** 1993. Group II self-splicing introns in bacteria. Nature (London) **364:**358–361.
25. **Garrett, R. A., J. Dalgaard, N. Larsen, J. Kjems, and A. S. Mankins.** 1991. Archaeal rRNA operons. Trends Biochem. Sci. **16:**22–26.
26. **Gimble, F. S., and J. Thorner.** 1992. Homing of a DNA endonuclease gene by meiotic gene conversion in *Saccharomyces cerevisiae*. Nature (London) **357:**301–306.
27. **Golding, G. B., and R. S. Gupta.** 1995. Protein based phylogenies support a chimeric origin for the eukaryotic genome. Mol. Biol. Evol. **12:**1–6.
28. **Goodrich-Blair, H., V. Scarlato, J. M. Gott, M. Xu, and D. A. Shub.** 1990. A self-splicing group I intron in the DNA polymerase gene of Bacillus subtilis bacteriophage SPO1. Cell **63:**417–424.
29. **Gott, J. M., D. A. Shub, and M. Belfort.** 1986. Multiple self-splicing introns in bacteriophage T4: evidence from autocatalytic GTP labeling of RNA in vitro. Cell **47:**81–87.
30. **Herbst, K. L., L. M. Nichols, R. F. Gesteland, and R. B. Weiss.** 1994. A mutation in ribosomal protein L9 affects ribosomal hopping during translation of gene 60 from bacteriophage T4. Proc. Natl. Acad. Sci. USA **91:**12525–12529.
31. **Hickey, D. A.** 1982. Selfish DNA: a sexually transmitted nuclear parasite. Genetics **101:**519–531.
32. **Huang, W. M., S. Ao, S. Casjens, R. Orlandi, R. Zeikus, R. Weiss, D. Winge, and M. Fang.** 1988. A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. Science **239:**1005–1012.
33. **Hurst, L. D.** 1994. The uncertain origin of introns. Nature (London) **371:**381–382.
34. **Inouye, S., M. G. Sunshine, E. W. Six, and M. Inouye.** 1991. Retronphage R73: An *E. coli* phage that contains a retroelement and integrates into a tRNA gene. Science **252:**969–971.
35. **Kaine, B. P., R. Gupta, and C. R. Woese.** 1983. Putative introns in tRNA genes of prokaryotes. Proc. Natl. Acad. Sci. USA **80:**3309–3312.
36. **Kane, P. M., C. T. Yamashiro, D. F. Wolczyk, N. Neff, M. Goebl, and T. H. Stevens.** 1990. Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H^+-adenosine triphosphatase. Science **250:**651–657.
37. **Kjems, J., and R. A. Garrett.** 1988. Novel splicing mechanism for the ribosomal RNA intron in the archaebacterium Desulfurococcus mobilis. Cell **54:**693–703.
38. **Kjems, J., and R. A. Garrett.** 1991. Ribosomal RNA introns in archaea and evidence for RNA conformational changes associated with splicing. Proc. Natl. Acad. Sci. USA **88:**439–443.
39. **Knoop, V., and A. Brennicke.** 1994. Evidence for a group II intron in *Escherichia coli* inserted into a highly conserved reading frame associated with mobile DNA sequences. Nucleic Acids Res. **22:**1167–1171.
40. **Kuhsel, M. G., R. Strickland, and J. D. Palmer.** 1990. An ancient group I intron shared by eubacteria and chloroplasts. Science **250:**1570–1573.
41. **Lambowitz, A. M., and M. Belfort.** 1993. Introns as mobile genetic elements. Annu. Rev. Biochem. **62:**587–622.
42. **Lambowitz, A. M., and P. S. Perlman.** 1990. Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. Trends Biochem. Sci. **15:**440–444.
43. **Lazowska, J., B. Meunier, and C. Macadre.** 1994. Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstream-located markers. EMBO J. **13:**4963–4972.
44. **Linton, D., J. P. Clewley, A. Burnens, R. J. Owen, and J. Stanley.** 1994. An intervening sequence (IVS) in the 16S rRNA gene of the eubacterium *Helicobacter canis*. Nucleic Acids Res. **22:**1954–1958.
45. **Loizos, N., E. R. M. Tillier, and M. Belfort.** 1994. Evolution of mobile group I introns: Recognition of intron sequences by an intron-encoded endonuclease. Proc. Natl. Acad. Sci. USA **91:**11983–11987.
46. **Lonergan, K. M., and M. W. Gray.** 1994. The ribosomal RNA gene region in Acanthamoeba castellanii mitochondrial DNA. A case of evolutionary transfer of intron between mitochondria and plastids? J. Mol. Biol. **239:**476–499.
47. **Lykke-Andersen, J., H. P. Thi-Ngoc, and R. A. Garrett.** 1994. DNA substrate specificity and cleavage kinetics of an archaeal homing-type endonuclease from *Pyrobaculum organotrophum*. Nucl. Acids Res. **22:**4583–4590.
48. **Magrelli, A., K. Langenkemper, C. Dehio, J. Schell, and A. Spena.** 1994. Splicing of the *rolA* transcript of *Agrobacterium rhizogenes* in *Arabidopsis*. Science **266:**1986–1988.
49. **Michel, F., L. Jaeger, E. Westhof, R. Kuras, F. Tihy, M. Xu, and D. A. Shub.** 1992. Activation of the catalytic core of a group I intron by a remote 3′ splice junction. Genes Dev. **6:**1373–1385.
50. **Michel, F., K. Umesono, and H. Ozeki.** 1989. Comparative and functional anatomy of group II catalytic introns—a review. Gene **82:**5–30.
51. **Michel, F., and E. Westhof.** 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. **216:**585–610.
52. **Moran, J. V., S. Zimmerly, R. Eskes, J. C. Kennell, A. M. Lambowitz, R. A. Butow, and P. S. Perlman.** 1995. Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. Mol. Cell. Biol. **15:**2828–2838.
53. **Mueller, J. E., M. Bryk, N. Loizos, and M. Belfort.** 1993. Homing endonucleases, p. 111–143. *In* S. M. Linn, R. S. Lloyd, and R. J. Roberts (ed.), Nucleases. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
54. **Mueller, M. W., M. Allmaier, R. Eskes, and R. J. Schweyen.** 1993. Transposition of group II intron al1 in yeast and invasion of mitochondrial genes at new locations. Nature (London) **366:**174–176.

55. **Orgel, L. E., and F. H. Crick.** 1980. Selfish DNA: the ultimate parasite. Nature (London) **284:**604–607.
56. **Perler, F. B., D. G. Comb, W. E. Jack, L. S. Moran, B. Quiang, R. B. Kucera, J. Benner, B. E. Slatko, D. O. Nwankwo, S. K. Hempstead, C. K. S. Carlow, and H. Jannasch.** 1992. Intervening sequences in an archaea DNA polymerase gene. Proc. Natl. Acad. Sci. USA **89:**5577–5581.
57. **Perler, F. B., E. O. Davis, G. E. Dean, F. S. Gimble, W. E. Jack, N. Neff, C. J. Noren, J. Thorner, and M. Belfort.** 1994. Protein splicing elements: inteins and exteins—a definition of terms and recommended nomenclature. Nucleic Acids Res. **22:**1125–1127.
58. **Ralph, D., and M. McClelland.** 1993. Intervening sequence with conserved open reading frame in eubacterial 23S rRNA genes. Proc. Natl. Acad. Sci. USA **90:**6864–6868.
59. **Reinhold-Hurek, B., and D. Shub.** 1992. Self-splicing introns in tRNA genes of widely divergent bacteria. Nature (London) **357:**173–176.
60. **Saldanha, R., G. Mohr, M. Belfort, and A. M. Lambowitz.** 1993. Group I and group II introns. FASEB J. **7:**15–24.
61. **Sellem, C. H., G. Lecellier, and L. Belcour.** 1993. Transposition of a group II intron. Nature (London) **366:**176–178.
62. **Sharp, P. A.** 1991. Five easy pieces. Science **254:**663.
63. **Thompson, A. J., and D. L. Herrin.** 1994. A chloroplast group I intron undergoes the first step of reverse splicing into host cytoplasmic 5.8 S rRNA. J. Mol. Biol. **236:**455–468.
64. **Thompson, L. D., and C. J. Daniels.** 1988. A $tRNA^{Trp}$ intron endonuclease from *Halobacterium volcanii*. Unique substrate recognition properties. J. Biol. Chem. **263:**17951–17959.
65. **Thompson, L. D., and C. J. Daniels.** 1990. Recognition of exon-intron boundaries by the *Halobacterium volcanii* tRNA intron endonuclease. J. Biol. Chem. **265:**18104–18111.
66. **Trust, T. J., S. M. Logan, C. E. Gustafson, P. J. Romaniuk, N. W. Kim, V. L. Chan, M. A. Ragan, P. Guerry, and R. R. Gutell.** 1994. Phylogenetic and molecular characterization of a 23S rRNA gene positions the genus *Campylobacter* in the epsilon subdivision of the *Proteobacteria* and shows that the presence of transcribed spacers is common in *Campylobacter* spp. J. Bacteriol. **176:**4597–4609.
67. **Wallace, C. J. A.** 1993. The curious case of protein splicing: mechanistic insights suggested by protein semisynthesis. Protein Sci. **2:**697–705.
68. **Weiss, R. B., W. M. Huang, and D. M. Dunn.** 1990. A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60. Cell **62:**117–126.
69. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA **87:**4576–4579.
70. **Xu, M., D. G. Comb, H. Paulus, C. J. Noren, Y. Shao, and F. B. Perler.** 1994. Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. EMBO J. **13:**5517–5522.
71. **Xu, M., M. W. Southworth, F. B. Mersha, L. J. Hornstra, and F. B. Perler.** 1993. In vitro protein splicing of purified precursor and the identification of a branched intermediate. Cell **75:**1371–1377.
72. **Xu, M.-Q., S. D. Kathe, H. Goodrich-Blair, S. A. Nierzwicki-Bauer, and D. A. Shub.** 1990. Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in *Cyanobacteria*. Science **250:**1566–1570.