Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 1692–1697, February 1998
Genetics

Introns and intein coding sequence in the ribonucleotide reductase genes of *Bacillus subtilis* temperate bacteriophage SPβ

VLADIMIR LAZAREVIC*, BLAZENKA SOLDO*, ANDREAS DÜSTERHÖFT†, HELMUT HILBERT†, CATHERINE MAUËL*, AND DIMITRI KARAMATA*‡

*Institut de Génétique et de Biologie Microbiennes, Rue César-Roux 19, CH-1005 Lausanne, Switzerland; and †Qiagen GmbH, Max-Volmer-Strasse 4, D-40724 Hilden, Germany

Communicated by Werner Arber, University of Basel, Basel, Switzerland, November 18, 1997 (received for review September 8, 1997)

ABSTRACT **The two putative ribonucleotide reductase subunits of the *Bacillus subtilis* bacteriophage SPβ are encoded by the *bnrdE* and *bnrdF* genes that are highly similar to corresponding host paralogs, located on the opposite replication arm. In contrast to their bacterial counterparts, *bnrdE* and *bnrdF* each are interrupted by a group I intron, efficiently removed *in vivo* by mRNA processing. The *bnrdF* intron contains an ORF encoding a polypeptide similar to homing endonucleases responsible for intron mobility, whereas the *bnrdE* intron has no obvious trace of coding sequence. The downstream *bnrdE* exon harbors an intervening sequence not excised at the level of the primary transcript, which encodes an in-frame polypeptide displaying all the features of an intein. Presently, this is the only intein identified in bacteriophages. In addition, *bnrdE* provides an example of a group I intron and an intein coding sequence within the same gene.**

Synthesis of the four deoxyribonucleotides, the DNA building blocks, requires the reduction of the four corresponding ribonucleotides by the ribonucleotide reductase (RR). Three distinct classes of RR are defined on the basis of their primary structure, oxygen tolerance/requirement and radical generator (1). A similar allosteric control and some sequence identity suggest that they might have evolved from a common, most likely anaerobic, ancestor (1, 2).

Eubacterial species encode one to three different RR. *Escherichia coli* possess three such enzymes, aerobic essential NrdA/NrdB, aerobic nonessential NrdE/NrdF and anaerobic NrdD/NrdG (3–8). In addition to host DNA degradation, certain bacteriophages supply nucleotides by a *de novo* pathway, in which phage-encoded RR play(s) a central role. The *E. coli* specific virulent bacteriophage T4 encodes two different RR. Interestingly, *nrdB*, the structural gene of the small subunit of aerobic RR, and *nrdD*, encoding the anaerobic enzyme, both are interrupted by a self-splicing group IA2 intron (9–13). Like many other group I introns, that of *nrdD* corresponds to a gene encoding the site-specific endonuclease involved in intron mobility, i.e., intron homing (14).

Inteins (15) are self-splicing elements that catalyze their own excision from the protein precursor and the concomitant ligation of C- and N-terminal segments called exteins (16, 17). Insertion of an intein DNA sequence into an intein-less allele, the process known as intein homing, is mediated by the intein endonuclease activity that is not required for splicing (18, 19). Based mainly on a computer search of conserved motifs, a total of 36 inteins have been detected in various organisms including yeasts, mycobacteria, cyanobacteria, archaea, and algal chloroplast (compiled in ref. 20). The archaeon *Pyrococcus furiosus* gene for anaerobic RR harbors two intein coding sequences (2). In contrast to introns, which are more common in phages than in eubacteria (21), no inteins have been identified in phage or viral genomes.

In *B. subtilis* the essential four-gene operon required for ribonucleotide reduction is located at 164° (22). Its two central genes, *nrdE* and *nrdF*, encode, respectively, the large and the small subunit of RR (22). Hereafter, we show that homologs of these two genes, named *bnrdEF* (from SP*beta nrdEF*) located at 185°, i.e., within the chromosomal segment corresponding to SPβ prophage, harbor each an intron. In addition, *bnrdE* contains an intein coding sequence.

MATERIALS AND METHODS

Bacteria and Plasmids. *E. coli* strain DH5α (23) was used as the host for plasmid constructs. Plasmids pPS344 and pPS394 contain, respectively, a 1,851-bp (sequence position 228-2078) and a 4,558-bp (sequence position 2,073–>5,700) SPβ *Eco*RI insert in vector pMTL20EC (24). PCR-amplified segments of cDNA obtained with oligonucleotide pairs VL264/VL265 and VL262/VL263 (see below) were cloned into pUC18 (25), yielding plasmids pPS609 and pPS610, respectively. *B. subtilis* strains CU1147 (26) and CU1050 (=su^{+3}) (27) were used for SPβ(*c2*) induction and amplification, respectively.

DNA Preparation. For manual sequencing and cloning, plasmid DNA was prepared by the alkaline lysis method (28). For automated sequencing, plasmids were purified using QIAGEN-tip 100 columns and QIAwell 8 Plasmid Kit (Qiagen, Hilden, Germany). PCR products were purified using QIAquick PCR Purification Kit (Qiagen).

RNA Isolation and Reverse Transcription (RT)–PCR. RNA was isolated with a RNeasy Total RNA Kit (Qiagen) from cells of strain CU1050, 15 min after infection with SPβ. After removal of traces of DNA with RNase-free DNase I (Pharmacia), RNA was repurified with the same kit. The reverse-transcriptase reaction was carried out with a First-strand cDNA Synthesis kit (Pharmacia) using 5 μg of total RNA and 30 pmol of downstream primer in a 15 μl volume. After addition of 2.5 units of *Taq* DNA polymerase (Pharmacia) and 30 pmol of each downstream and upstream primer, a segment of cDNA was amplified in total volume of 50 μl by 30 cycles of PCR. Each cycle included 1 min of melting at 95°C, 1 min of annealing at 45°C, and 1 min of extension at 72°C. The upstream primers VL264 5′-AGCATTAAATCTAAACAAACTAAGAGC-3′, VL262 5′-AGCAACATCTTTCTAACATTGGCTC-3′, and VL268 5′-GGGAAATGCTCGAAAGTTGTTGGAGC-3′ were used in reactions initiated with downstream oligonucleotides VL265 5′-GCTCCAACAACTTTC-

Abbreviations: RR, ribonucleotide reductase(s); RT-PCR, reverse transcription–PCR.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF020713).

A commentary on this article begins on page 1356.

‡To whom reprint request should be addressed. e-mail: dimitri.karamata@igbm.unil.ch.

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.

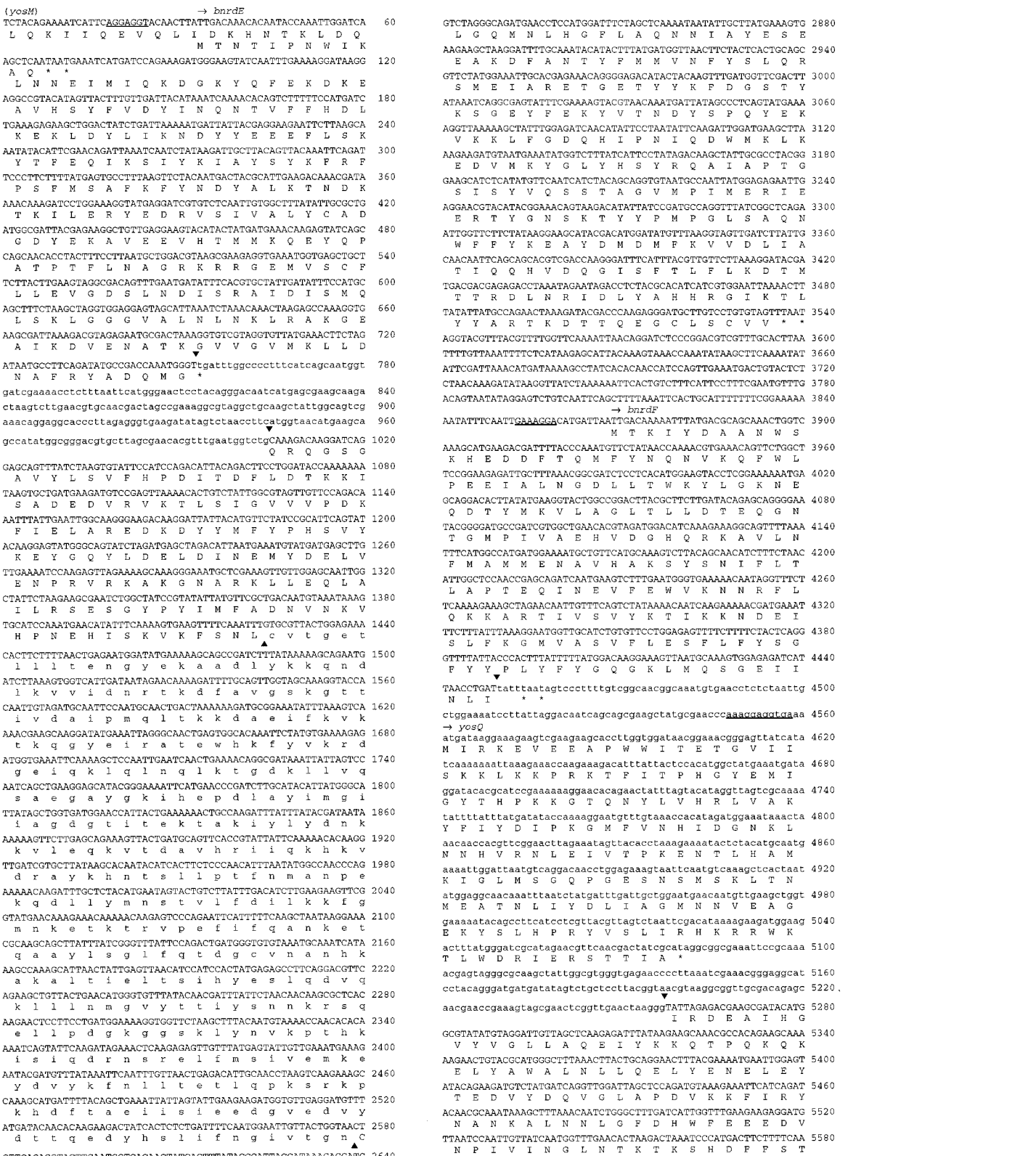

FIG. 1. The nucleotide and the deduced amino acid sequence of SPβ *bnrdE* and *bnrdF* genes. The two introns and an intein are indicated by lowercase type in the corresponding sequences. Residues complementary to *B. subtilis* 16S rRNA are underlined. Asterisks denote stop codons. Arrows indicate translational start codons. ▲ and ▼, extein-intein and exon-intron junctions.

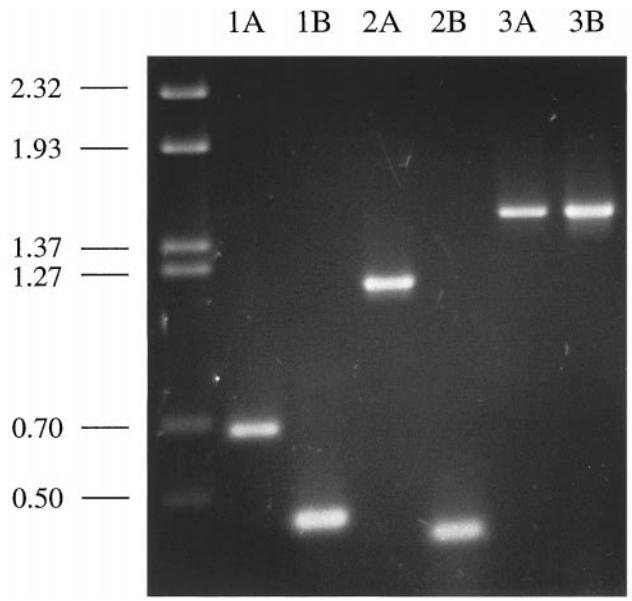

FIG. 2. *In vivo* splicing of intron RNA. PCR amplification of phage genomic DNA segments containing *bnrdE* intron (1A), *bnrdF* intron (2A) and *bnrdE* intein coding sequences (3A). Lanes B, RT-PCR products obtained using RNA isolated from phage-infected cells and pairs of oligonucleotides identical to those in corresponding lanes A.

GAGCATTTCCC-3′, VL263 5′-CAGTAAGTTTAAAGCC-CATGCGTAC-3′, and VL269 5′-TGCTCTCGCAACTGCT-GGAGCATTTAC-3′, respectively.

Sequencing. The nucleotide sequence of plasmids pPS609 and pPS610 was determined by dideoxy chain termination with the Sequenase version 2.0 kit (United States Biochemical) and [α-^{35}S]dATP (Amersham), using M13 forward and reverse primers. Plasmids pPS344 and pPS394 were sequenced using a primer walking strategy. Applied Biosystems dye-terminators and Ampli*Taq* DNA polymerase FS were used for cycle sequencing reactions. Automated set-up of sequencing reactions was carried out on the BioRobot 9600 laboratory workstation (Qiagen). Automated sequencing was done on an Applied Biosystems 377 DNA sequencer (Perkin–Elmer). Assembly, editing and finishing of data were carried out using the SeqMan II module of the Lasergene software package (DNAstar, Madison, WI). The final sequence was analyzed by the University of Wisconsin Computer Group software (29).

SPβ Preparation. A stock of SPβ was obtained by heat shock of strain CU1147. The DNase-treated lysate was used to infect the SPβ-cured strain CU1050 grown in Luria–Bertani medium supplemented with 0.1% glucose, 10 mM $CaCl_2$, and 10 mM $MgCl_2$, at an OD_{595} of 0.3. The incubation at 37°C was continued until lysis. Phage DNA was isolated from the lysate using the Qiagen Lambda Midi kit.

PCR. PCRs were set up with 0.1 ng of SPβ genomic DNA, 100 pmol of each primer, 20 nmol of four dNTPs (Pharmacia) in 100 μl of reaction buffer (Pharmacia) containing 2.5 units of *Taq* DNA polymerase (Pharmacia). The reactions were run with denaturation for 2 min at 95°, followed by 30 cycles of amplification (95°C for 30 s, 45°C for 1 min, 72°C for 1 min/kb of the segment to be amplified), and hold extension for 10 min at 72°C. A PCR product of 2,483 bases (sequence position <1–349), generated on SPβ genomic DNA using oligonucleotides BS304 5′-TTCAATGCGTAGTCATTGTAG-3′ and BS319 5′-ATGGGTTGAACTAGGCGGTGT-3′ was sequenced directly.

RESULTS AND DISCUSSION

Sequencing of the temperate phage SPβ allowed the identification of a putative 22-gene operon (unpublished data) specifying among others dUTPase, thioredoxin, and RR, all involved in the synthesis of DNA precursors. The putative large and small subunits of the phage RR are encoded by ORFs *bnrdE* and *bnrdF*, respectively, (Fig. 1) which exhibit over 70% identity (not presented) to their host counterparts *nrdE* and *nrdF* (22), located at 164°. *bnrdE* has a 38-nt overlap with the upstream ORF *bnrdI* whose homolog *nrdI* (*ymaA*) precedes the host *nrdE* gene (22). Comparison of corresponding *B. subtilis* strain 168 (22) and phage sequences revealed two intervening segments of 252 and 1,155 bp in *bnrdE*, and one such 808-bp segment in *bnrdF*. Absence of these segments at analogous positions in the *nrdE* and *nrdF* genes of the lysogenic host strain CU1147 and the SPβ-cured strain CU1050 was confirmed by PCR (not presented). The proximal nonhomologous insert in *bnrdE* starts with a UGA stop codon, whereas in that found in *bnrdF*, the termination codon UAA is separated from the upstream coding sequence by four nucleotides (Fig. 1). These two intervening sequences are similar to group IA2 phage introns (30, 31). Evidence for their *in vivo* excision was obtained by RT-PCR. RNA isolated 15 min upon phage infection of the SPβ-cured strain CU1050 was used as template for cDNA synthesis initiated with the oligonucleotides VL263 and VL265 specific for downstream exons. To amplify the entire introns and parts of the flanking exons, the oligonucleotides VL262 and VL264 corresponding to the upstream exons were included in the subsequent second strand synthesis and PCR reactions. For both pairs of primers, the reaction products were smaller than those generated on phage genomic DNA as template (Fig. 2). Differences in size correspond to introns predicted from the nucleotide sequence, revealing intron excision from the primary transcript. Exact intron boundaries were determined by sequencing of cloned PCR-amplified cDNA segments (Fig. 3).

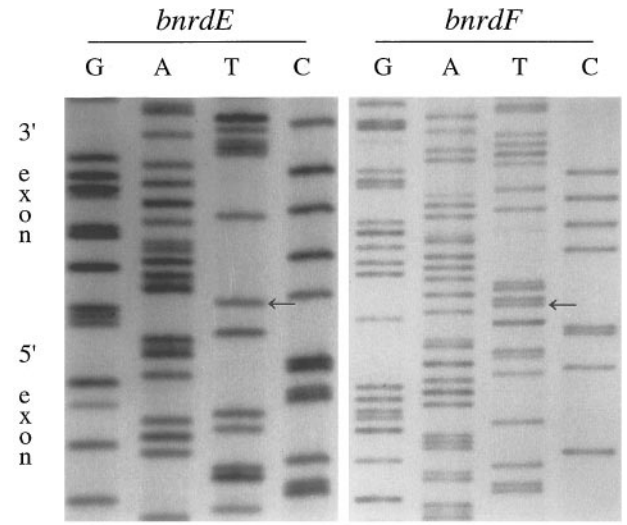

FIG. 3. *bnrdE* and *bnrdF* splice junctions. The nucleotide sequence corresponding to RT-PCR-amplified spliced mRNA. Last residue of the upstream exon to which the downstream exon is ligated is marked by an arrow.

The two SPβ and other group I introns share short conserved sequences (P, Q, R and S) as well as similar secondary structure elements (Fig. 4), represented by both local and long-range complementary base pairing regions (P1–P9), necessary for proper folding and excision (32–35). Like in most other group I introns, the 5′ splice site of the *bnrdE* and *bnrdF* introns is located in P1 (Fig. 4), after a uridine paired to guanosine, whereas the 3′ splice site follows a guanosine (32). The first four 5′-terminal nucleotides of the downstream *bnrdE* exon and the seven residues occupying the analogous position in *bnrdF* are complementary to residues immediately preceding the guanosine paired with uridine at the 5′ splice site. These

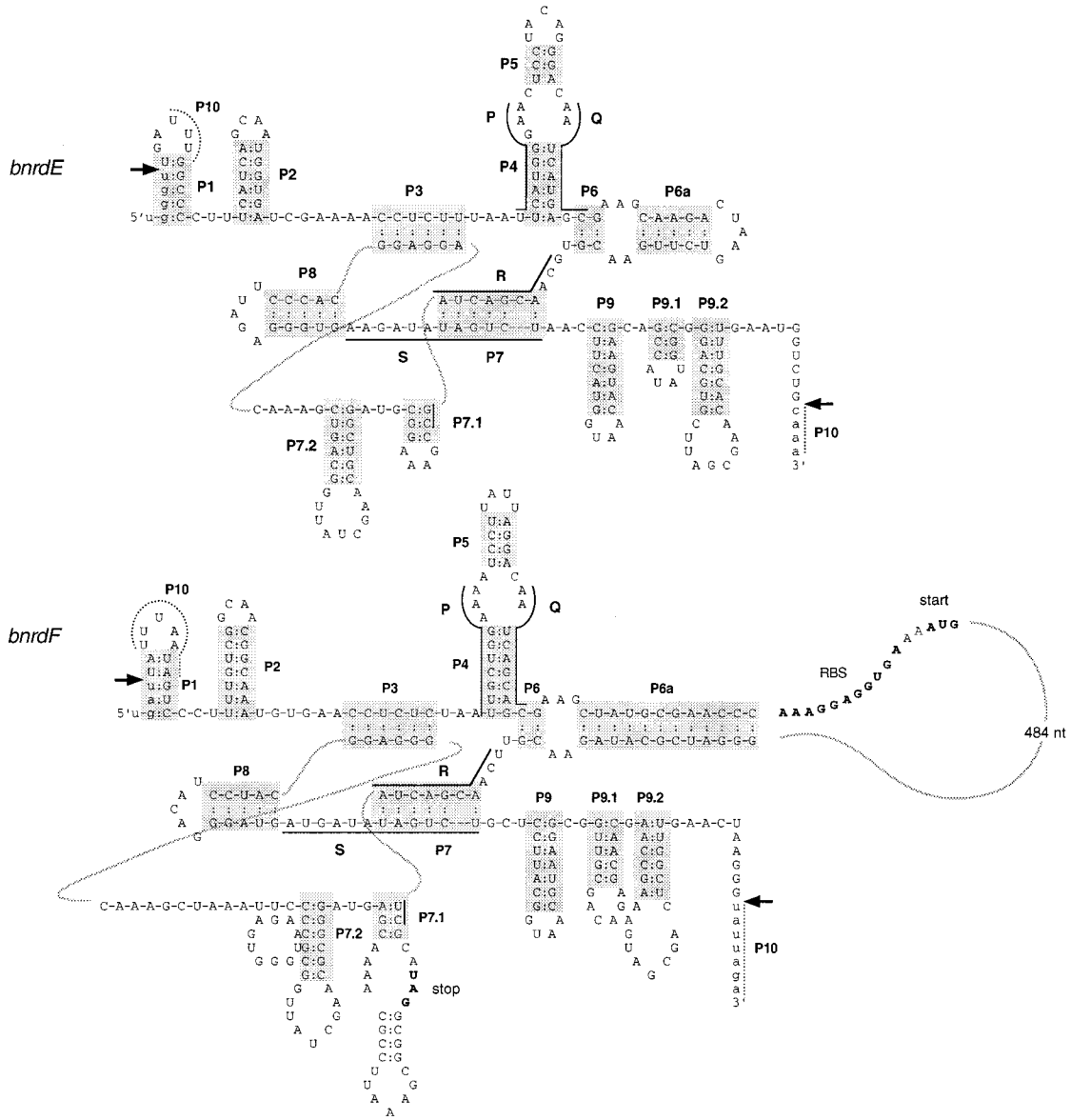

FIG. 4. Proposed secondary structures for *bnrdE* and *bnrdF* introns. Arrows indicate splice boundaries between exons (lowercase) and introns (uppercase). Conserved base-paired regions (P1–P9) are shaded, and conserved primary structure elements (P, Q, R, and S) are marked by solid lines. Complementary intron and downstream exon segments that could pair and form P10 are marked by dotted line. Ribosome-binding site (RBS) and start and stop codons of *yosQ* are indicated and represented in bold.

paired exon sequences (P10) contribute to alignment of 3′ and 5′ splice sites for ligation (34, 36).

The *bnrdF* intron contains a 522-nt ORF, named *yosQ*, which begins in the large peripheral loop of stem P6 and ends in the unpaired region of stem P7.1 (Fig. 4). At appropriate distance, *yosQ* is preceded by a strong ribosome-binding site whose 11-base stretch is complementary to the *B. subtilis* 16S rRNA (Figs. 1 and 4). Inspection of protein databases revealed similarities (not presented) between the N-terminal moiety of YosQ and those of Gram-positive phage intron-encoded and free-standing endonucleases (37). The conserved domain contains a H-N-H motif defining a larger family of phage and bacterial endonucleases, while variations in the C-terminal part might indicate involvement in recognition of different target sequences (37).

After removal of the *bnrdF* intron, translation of the messenger RNA yields a polypeptide 93% identical to *B. subtilis* NrdF (Fig. 5*A*). However, removal of the *bnrdE* intron generates an ORF encoding a 1084 residues protein and exhibiting 87% identity to *B. subtilis* NrdE, but containing an extra domain of 385 amino acids (Fig. 5*B*). The possibility that this intervening sequence is spliced out at the RNA level was ruled out by using RT-PCR. The length of the reaction product, generated under conditions allowing efficient removal of *bnrdE* and *bnrdF* introns, corresponded to unspliced messenger RNA (Fig. 2). Inspection of the amino acid sequence of this nonhomologous insert revealed all known intein features, suggesting splicing at the protein level. Residues identified as critical for splicing (19, 38–40), namely cysteine as one of three possible residues at the C-terminal side of each of the two splice junctions, and asparagine at the last position of the intein, were found (Figs. 1 and 5*b*). Histidine, conserved at the penultimate position in most known inteins (20) but not absolutely required for splicing (38, 41), is here replaced by glycine. It provides the third example of such a substitution and the fourth case of a nonhistidine residue in this position. The putatively excised protein has all the conserved intein motifs termed blocks A–H (20, 42). Blocks C and E correspond to the two copies of the LAGLIDADG motif (43) initially identified in yeast mitochondrial maturases, and later found in endo-

FIG. 5. Alignment of amino acid sequences of the *bnrdF* (*A*) and *bnrdE* (*B*) products, obtained after intron excision, with their host paralogs. Conserved intein blocks A–H are indicated.

nucleases encoded by group I introns, archeal introns, inteins, and yeast free-standing endonucleases (21, 42). Therefore, it is likely that the predicted intein corresponds to a homing site-specific endonuclease capable of inserting a copy of its DNA site into an intein-less allele. Finally, the intein length falls within the 150–548 residue range, reported for other inteins (20).

Aligned sequences of *bnrdE* and *bnrdF* introns revealed a 74% identity, as well as five gaps within the *bnrdE* intron (not presented). The longest gap corresponds to *yosQ*. Both introns exhibit somewhat lower degrees of homology with bacteriophage introns (not presented) found in genes *thy* (44) of β22 (Gram-positive) and *nrdB* (45) of T4 (Gram-negative). In contrast, YosQ resembles (not presented) intron-encoded and free-standing endonucleases found in phages of Gram-positive organisms only (37) supporting the hypothesis that introns and intron-encoded ORFs evolve independently (44). The G+C content of the core of *bnrdE* and *bnrdF* introns (48%), is over 10% higher than that of the *bnrdEF* coding sequences and that of *yosQ*. This high G+C content is probably essential for intron folding (Fig. 4). Mutations tending to reduce it to the lower host G+C content would most likely be deleterious to phage due to deficient splicing of RR mRNA.

The generation of the functional SPβ BnrdE protein is an example of self-splicing at both RNA and protein levels. This unique intron-intein configuration might have arisen from the mobility of both elements and the presence in the *bnrdE* gene of target sequences for their respective endonucleases. Although the *bnrdE* intron does not encode a protein, it is possible that an ORF specifying the homing enzyme was lost. The *B. subtilis* phage β22 *thy* gene (44) and the coliphage T4 *nrdB* gene (9, 45) offer two examples of phage introns occupied by nonfunctional remnants of an endonuclease gene. Conversely, the *bnrdE* intron may be considered as the archetype of an intron that has not been invaded by an endonuclease. Absence of intervening sequences in *B. subtilis nrdE* and *nrdF* genes (not presented) renders unlikely the possibility that the prophage *bnrdE* intein coding sequence and the *bnrdF* intron were acquired from the actual host DNA. Transmission from phage to host could have occurred only had the expression of the phage homing endonucleases been independent of prophage induction, a phenomenon leading to host cell lysis.

Should introns and inteins be part of a prokaryotic regulatory mechanism(s), the latter, in view of its most seldom occurrence, would seem not to efficiently compete with other prokaryotic regulatory mechanisms. There is no satisfactory explanation for the preferential occurrence of intervening sequences within genes involved in DNA metabolism, in particular in the genes encoding RR, the enzyme that allowed transition from RNA to DNA world. The RR diversity (1) is rendered even more complex by the presence of intron(s) and/or intein(s) (2, 9–13). Presently, the significance of intron or intein splicing mechanisms remains unknown, leaving a possible advantage(s) of a two-level splicing open to speculation. It is likely that understanding the phylogenetic aspects of the intein and intron interactions with host genomes will benefit from the knowledge of the ever increasing number of complete genome sequences.

This work was supported by Grant 96.0245 from the Office Fédéral de l'Education et de la Science (D.K. and C.M.) and Grant BIO4-CT96–0655 from the European Commission (A.D.).

1. Reichard, P. (1993) *Science* **260,** 1773–1777.
2. Riera, J., Robb, F. T., Weiss, R. & Fontecave, M. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 475–478.

3. Carlson, J., Fuchs, J. A. & Messing, J. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4294–4297.
4. Fontecave, M., Nordlund, P., Eklund, H. & Reichard, P. (1992) *Adv. Enzymol.* **65,** 147–183.
5. Jordan, A., Gibert, I. & Barbe, J. (1994) *J. Bacteriol.* **176,** 3420–3427.
6. Sun, X., Harder, J., Krook, M., Jörnvall, H., Sjöberg, B. M., Reichard, P. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 577–581.
7. Reichard, P. (1993) *J. Biol. Chem.* **268,** 8383–8386.
8. Ollagnier, S., Mulliez, E., Gaillard, J., Eliasson, R., Fontecave, M. & Reichard, P. (1996) *J. Biol. Chem.* **271,** 9410–9416.
9. Sjöberg, B. M., Hahne, S., Mathews, C. Z., Mathews, C. K., Rand, K. N. & Gait, M. J. (1986) *EMBO J.* **5,** 2031–2036.
10. Gott, J. M., Shub, D. A. & Belfort, M. (1986) *Cell* **47,** 81–87.
11. Tomaschewski, J. & Ruger, W. (1987) *Nucleic Acids Res.* **15,** 3632–3633.
12. Shub, D. A., Xu, M. Q., Gott, J. M., Zeeh, A. & Wilson, L. D. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** 193–200.
13. Young, P., Ohman, M., Xu, M. Q, Shub, D. A & Sjöberg, B. M. (1994) *J. Biol. Chem.* **269,** 20229–20232.
14. Quirk, S. M., Bell-Pedersen, D. & Belfort, M. (1989) *Cell* **56,** 455–465.
15. Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J. & Belfort, M. (1994) *Nucleic Acids Res.* **22,** 1125–1127.
16. Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M. & Stevens, T. H. (1990) *Science* **250,** 651–657.
17. Hirata, R., Ohsumk, Y., Nakano, A., Kawasaki, H., Suzuki, K. & Anraku, Y. (1990) *J. Biol. Chem.* **265,** 6726–6733.
18. Gimble, F. S. & Thorner, J. (1992) *Nature (London)* **357,** 301–306.
19. Hodges, R. A., Perler, F. B., Noren, C. J. & Jack, W. E. (1992) *Nucleic Acids Res.* **20,** 6153–6157.
20. Perler, F. B., Olsen, G. J. & Adam, E. (1997) *Nucleic Acids Res.* **25,** 1087–1093.
21. Lambowitz, A. M. & Belfort, M. (1993) *Annu. Rev. Biochem.* **62,** 587–622.
22. Scotti, C., Valbuzzi, A., Perego, M., Galizzi, A. & Albertini A. M. (1996) *Microbiology* **142,** 2995–3004.
23. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
24. Chambers, S. P., Prior, S. E., Barstow, D. A. & Minton, N. P. (1988) *Gene* **68,** 139–149.
25. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
26. Rosenthal, R., Toye, P. A., Korman, R. Z. & Zahler, S. A. (1979) *Genetics* **92,** 721–739.
27. Warner, F. D., Kitos, G. A., Romano, M. P. & Hemphill, H. E. (1977) *Can. J. Microbiol.* **23,** 45–51.
28. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523.
29. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
30. Michel, F. & Westhof, E. (1990) *J. Mol. Biol.* **216,** 585–610.
31. Shub, D. A., Gott, J. M., Xu, M. Q., Lang, B. F., Michel, F., Tomaschewski, J., Pedersen-Lane, J. & Belfort, M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 1151–1155.
32. Cech, T. R. (1988) *Gene* **73,** 259–271.
33. Michel, F., Jacquier, A. & Dujon, B. (1982) *Biochimie* **64,** 867–881.
34. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. (1982) *Nature (London)* **300,** 719–724.
35. Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R. J., Shub, D. A., Szostak J. W. & Tabak, H. F. (1987) *Nucleic Acids Res.* **15,** 7217–7221.
36. Michel, F., Hanna, M., Green, R., Bartel, D. P. & Szostak, J. W. (1989) *Nature (London)* **342,** 391–395.
37. Goodrich-Blair, H. & Shub, D. A. (1994) *Nucleic Acids Res.* **22,** 3715–3721.
38. Cooper, A. A., Chen, Y. J., Lindorfer, M. A. & Stevens, T. H. (1993) *EMBO J.* **12,** 2575–2583.
39. Davis, E. O., Jenner, P. J., Brooks P. C., Colston, M. J. & Sedgwick, S. G. (1992) *Cell* **71,** 201–210.
40. Hirata, R. & Anraku, Y. (1992) *Biochem. Biophys. Res. Commun.* **188,** 40–47.
41. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B. & Xu, M. Q. (1996) *J. Biol. Chem.* **271,** 22159–22168.
42. Pietrokovski, S. (1994) *Protein Sci.* **3,** 2340–2350.
43. Hensgens, L. A., Bonen, L., de Haan, M., van der Horst, G. & Grivell, L. A. (1983) *Cell* **32,** 379–389.
44. Bechhofer, D. H., Hue, K. K. & Shub, D. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11669–11673.
45. Eddy, S. R. & Gold, L. (1991) *Genes Dev.* **5,** 1032–1041.