The *Bacillus subtilis* genes for ribonucleotide reductase are similar to the genes for the second class I NrdE/NrdF enzymes of *Enterobacteriaceae*

Claudio Scotti,† Angela Valbuzzi,1 Marta Perego,1,2† Alessandro Galizzi1 and Alessandra M. Albertini1

Author for correspondence: Alessandra M. Albertini. Tel: +39 382 505549. Fax: +39 382 528496. e-mail: Albert@pillo.unipv.it

We have cloned and sequenced the *nrd* (nucleotide reductase) locus of *Bacillus subtilis*. The locus seems to be organized in an operon comprising four ORFs. The first three encode polypeptides highly similar to the product of the coding sequences characterizing the *nrdE*/*nrdF* operons of *Enterobacteriaceae*. The sequencing of the conditional lethal mutation *ts-A13*, localized in the *nrdE* cistron, and the lethality of insertional mutations targeted in the internal region of *nrdE* and *nrdF*, demonstrated the essential role of this locus. The fourth ORF, *ymaB*, part of the putative operon, which is not similar to any known protein, is also essential. The regulation of expression of the operon, monitored by *lacZ* transcriptional fusions, is similar to the regulation of the functionally relevant *nrdAB* operon of *Escherichia coli*. The operon was induced by thymidine starvation and its expression was directly or indirectly affected by RecA function. Genetic and functional analysis strongly indicates that in *B. subtilis* the class I ribonucleotide reductase encoded by this *nrd* operon is evolutionarily distant from the homologous class I enzyme of Enterobacteria.

Keywords: *Bacillus subtilis*, genome sequencing, ribonucleotide reductase, *nrdEF* operon, *ts-A13* mutation

INTRODUCTION

Ribonucleotide reductases (RRs, ribonucleoside diphosphate reductases) are enzymes essential for the biosynthesis of deoxyribonucleotides from the corresponding ribonucleotides. In addition they exert an important role in balancing the supply of the four dNTPs and consequently in the control of DNA synthesis. In prokaryotes the presence of three classes of enzymes has been documented (Reichard, 1993). Class I enzymes, the best known from structural studies (Nordlund *et al.*, 1990; Nordlund & Eklund, 1993; Uhlin & Eklund, 1994), are tetrameric, aerobic and use, for the generation of the organic radical necessary in the initial step of the enzymic reaction, an iron-containing protein, represented by the R₄ homodimer. The R₄ homodimer contains the allosteric binding sites and the catalytic site. Class II enzymes, present in different aerobic and anaerobic prokaryotes, are monomeric and use adenosyl cobalamin as radical generator (Booker & Stubbe, 1993) taking over the function of the R₄ protein of the class I enzymes. The amino acid sequence has no similarity with R₄ with the exception of nine amino acids at the C terminus. The class III enzyme was discovered in *Escherichia coli* growing anaerobically. It is a homodimer with an iron–sulfur centre and an oxygen-sensitive protein radical probably represented by a Gly radical close to the C terminus (Sun *et al.*, 1993). A similar function is encoded by bacteriophage T4. Comparison of the amino acid sequences of class I enzymes of different species suggested the existence of three different groups, namely eukaryotic RR, herpesvirus-type enzymes, and *E. coli* and bacteriophage T4 *nrdAB* gene products (Nordlund & Eklund, 1993), with a very low level of similarity between them (at the boundary of significance, between 20 and 30% identity). Recently, the genes encoding a
fourth group of class I enzymes were identified in *Salmonella typhimurium* and *E. coli* and named nrdEF (Jordan *et al.*, 1994a). This new operon, when carried on a multicopy plasmid, can complement nrdAB conditional lethal mutants in standard growth conditions, suggesting that it encodes an aerobic function, usually expressed at low level. The deduced products of nrdEF showed only limited homology with the other groups of class I enzymes (between 19 and 27% identity) but many of the catalytically important amino acid residues were conserved. The products of the two genes, nrdE and nrdF, are homodimeric proteins: the R1p,R2p enzyme is able to catalyse the reduction of CDP to dCDP in the presence of dihydrothreitol or reduced glutaredoxin as an electron donor, instead of thioredoxin, and its allosteric regulation is different from that exhibited by the enzyme encoded by the nrdAB operon (Jordan *et al.*, 1994b).

In *E. coli* the expression of the nrdAB operon is cell-cycle-regulated (Filpula & Fuchs, 1977, 1978; Sun & Fuchs, 1992) with a cis-acting positive A-T-rich element (Sun *et al.*, 1994) as regulator. In proteins are not required for the coupling of both increased when regulated (Filpula & Fuchs, 1992) with a cis-acting positive A-T-rich element (Sun *et al.*, 1994), but this increase seems to be only indirectly related to DNA damage. The activity of RR and the transcription of the operon are temperature-sensitive mutant (Fuchs, 1983) and after DNA damage (Gibert *et al.*, 1987). The nrd locus, encoding a RR, functionally similar to the *E. coli* and named nrdAB (Fuchs, 1983) and after DNA damage (Gibert *et al.*, 1987). The nrd locus, encoding a RR, functionally similar to the *E. coli* R(A)A was isolated from a plasmid bank of Bacill*ls* subtilis genomic DNA (Perego *et al.*, 1987). The nrd (dnaA) locus was mapped at about 167° on the chromosome, between the *glnR*A operon and the *pks* locus (Fisher *et al.*, 1984; Gianni & Galizzi, 1986; Scotti *et al.*, 1993). In this paper we describe the molecular cloning and sequencing of the entire *B. subtilis* nrd locus, the localization and nature of the *ts-A13* mutation, gene inactivation experiments and preliminary results on the regulation of expression of the locus.

**METHODS**

**Bacterial strains, plasmids and media.** *B. subtilis* strains used for gene inactivation and transcriptional fusion are listed in Table 1. *E. coli* DP500 (supF [F′ tonA53 glnV-44 tyrT-58 dapD-8 lacY-1 supE-44 supF-38 Δ(gal-α2)47; gyrA-29 thyA-57 hisD3]) and LE392 (supE-44 supF-58 hsdR-54 galK-2 galT-22 metB-1 trpR-55 lacY-1) were used as hosts for the *E. coli* operon and the *pks* locus (Fisher *et al.*, 1984; Gianni & Galizzi, 1986; Scotti *et al.*, 1993). In this paper we describe the molecular cloning and sequencing of the entire *B. subtilis* nrd locus, the localization and nature of the *ts-A13* mutation, gene inactivation experiments and preliminary results on the regulation of expression of the locus.

Antibiotics were added at the following concentrations: ampicillin (Ap), 100 μg ml⁻¹; chloramphenicol (Cm), 20 μg ml⁻¹ for *E. coli* and 2.5 or 5 μg ml⁻¹ for *B. subtilis*, kanamycin (Km), 2 μg ml⁻¹. *E. coli* DH5α [supE44 lacU169 (φ80 lacZ AM15) hisDR17 recA1 endA1 gyrA96 thi-1 relA1] was used for plasmid transformation and propagation. The genomic *B. subtilis* λ Charon 4A bank was made with DNA from strain 168 trpC-2 (Ferrari *et al.*, 1983). The insert from the *ΔCH4G621* clone was sub-cloned in the pJM103 integrative vector (Perego, 1993); the integrative vectors pJM783 and pJM115 were used to generate transcriptional fusions to lacZ (Perego, 1993).

Two transcriptional fusions were constructed, one placing the reporter gene, lacZ, downstream of the whole operon to evaluate the expression of the operon. The second fusion was constructed to test the presence of a putative promoter between nrdF and ymbB (Fig. 1). This fusion was inserted into the amyE locus.

A 538 bp fragment corresponding to the 3′ end of ymbB, with artificial EcoRI and BamHI sites, was produced by PCR using the primers 5′ GGAATTCAGTGAAGATAGCAGAGTACGGAT-CA 3′ and 5′ GCCCTAGGATTCAAGAATAACAAACA 3′ (Ferrari *et al.*, 1983). The insert from the *ΔCH4G621* clone was sub-cloned in the pJM783 integrative vector (Perego, 1993).

The plasmid obtained, pM783/538 EB, was used to transform competent cells of strains PB168 or PB1831, selecting for Km. The integration by single crossover in the region of homology, provided by the insert, placed the lacZ reporter gene downstream of the whole operon. The construction was verified by Southern hybridization.

A 487 bp fragment with artificial EcoRI and BamHI ends, corresponding to the intercistronic region between nrdF and ymbB, was synthesized by PCR using the primers 5′ GGAATTCAGTGAAGATAGCAGAGTACGGAT-CA 3′ and 5′ GGGATCC TTATTGTGTTTATCAT 3′ corresponding respectively to nt 3914–3932 and 4400–4381, and cloned in the vector pJM115. The derived plasmid, named pJM115/487 EB was linearized and used to transform competent cells of strain PB168 selecting for Km. Double crossover events caused by the insertion of the nrdF–lacZ fusion at the amyE locus gave an Amy phenotype.

The other plasmids used for sequencing and insertional gene inactivation of the three main coding sequences of the nrd operon were constructed by cloning suitable restriction fragments in the integrative vector pJM103 and are partially listed in Table 2.

**Transformation.** *E. coli* transformation was performed following standard protocols (Hanahan, 1985). *B. subtilis* competent cells (Hoch *et al.*, 1967) were transformed using 0.5–1 μg plasmid DNA or 0–2 μg chromosomal DNA (Albertini & Galizzi, 1985).

**Sequencing.** Sequencing was performed on plasmid subclones in pJM103 (Table 2) obtained from the EcoRI–SalI fragment of *ΔCH4G621* as indicated in Fig. 1. The overlapping HindIII fragment was amplified by reverse PCR, recovering a fragment of 1.8 kb (dashed in Fig. 1) from chromosomal DNA digested with HindIII, ligated and cut at the unique SacI site at nt 1782. Amplification was obtained with primers NRDA3, 5′ CCGA-CTCATGAGAAGAAAACGA 3′, and NRDA4, 5′ GGAATTCAGTGAAGATAGCAGAGTACGGAT-CA 3′. The plasmid subclones were sequenced with Sequenase (USB) or T7 (Pharmacia) sequencing kits. The reverse PCR product was sequenced with the Circumvent kit (Biolabs). In all cases the sequence was determined on both strands by primer walking using custom oligonucleotides synthesized with a Cyclon Plus Synthesizer (Milligen Biosearch).
Table 1. *B. subtilis* strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype*</th>
<th>Reference or source†</th>
</tr>
</thead>
<tbody>
<tr>
<td>PB168</td>
<td>trpC-2</td>
<td>J. Lederberg</td>
</tr>
<tr>
<td>PB566</td>
<td>trpC-2 tbyA tbyB</td>
<td></td>
</tr>
<tr>
<td>PB1679</td>
<td>ilvA-1 metB-5 ts-A13</td>
<td>E3113 (D. Karamata)</td>
</tr>
<tr>
<td>PB1783</td>
<td>tbyA tbyB glmA-100</td>
<td>JH 949 (J. Hoch)</td>
</tr>
<tr>
<td>PB1831</td>
<td>trpC-2 phe-1</td>
<td>JH642 (J. Hoch)</td>
</tr>
<tr>
<td>PB1843</td>
<td>trpC-2 recA-260 metB-5 SPJ(−) xin-1 erm</td>
<td>BGSC 1A746</td>
</tr>
<tr>
<td>PB1848</td>
<td>trpC-2 phe-1 nrd::pJM783/538 EB nrd-lacZ cat</td>
<td></td>
</tr>
<tr>
<td>PB1849</td>
<td>trpC-2 nrd::pJM783/538 EB nrd-lacZ cat</td>
<td></td>
</tr>
<tr>
<td>PB1850</td>
<td>trpC-2 tbyA tbyB nrd::pJM783/538 EB nrd-lacZ cat</td>
<td></td>
</tr>
<tr>
<td>PB1851</td>
<td>trpC-2 phe-1 ΔcompP-A nrd::pJM783/538 EB nrd-lacZ kan cat</td>
<td></td>
</tr>
<tr>
<td>PB1852</td>
<td>trpC-2 phe-1 recA-260 nrd::pJM783/538 EB nrd-lacZ erm cat</td>
<td></td>
</tr>
<tr>
<td>PB1854</td>
<td>trpC-2 tbyA tbyB ΔcompP-A nrd::pJM783/538 EB nrd-lacZ kan cat</td>
<td></td>
</tr>
<tr>
<td>PB1855</td>
<td>trpC-2 tbyA tbyB recA-260 nrd::pJM783/538 EB nrd-lacZ erm cat</td>
<td></td>
</tr>
<tr>
<td>PB1874</td>
<td>trpC-2 phe-1amE::pJM115/487 EB nrd-lacZ kan</td>
<td></td>
</tr>
<tr>
<td>PB5091</td>
<td>trpC-2 ΔcompP-A kan</td>
<td>QB4714 (T. Msadek)</td>
</tr>
<tr>
<td>PB5137</td>
<td>trpC-2 phe-1 spoOA-12</td>
<td>JH646 (J. Hoch)</td>
</tr>
</tbody>
</table>

* sat, kan and erm are the determinants for resistance to chloramphenicol, kanamycin and erythromycin, respectively.
† BGSC, Bacillus Genetic Stock Center, Ohio, USA; J. Hoch, La Jolla, USA; D. Karamata, Lausanne, Switzerland; J. Lederberg, New York, USA; T. Msadek, Paris, France.

Fig. 1. Physical map and gene organization of the *nrd* locus of *B. subtilis*. (a) Clone λCh4G621. (b) Map of the sequenced region. The hatched bar represents the sequence obtained by reverse PCR; arrows indicate the extent of the coding sequences; hairpins indicate putative terminator or attenuator structures; boxes indicate the localization of the fragments cloned in relevant plasmids. H, *HindIII*; E, *EcoRI*; S, *SalI*; P, *PstI*; S2, *Sall*. △ indicates the position of the ts-A13 mutation.
Table 2. Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description (source or reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vectors</td>
<td></td>
</tr>
<tr>
<td>pJH101</td>
<td>Ap Tc Cm; integrative vector (Ferrari et al., 1983)</td>
</tr>
<tr>
<td>pJM103</td>
<td>Ap Cm; integrative vector (Perego, 1993)</td>
</tr>
<tr>
<td>pJM783</td>
<td>Ap Cm; vector used to construct a lacZ transcriptional fusion</td>
</tr>
<tr>
<td></td>
<td>(Perego, 1993)</td>
</tr>
<tr>
<td>pJM115</td>
<td>Ap Km; vector used to construct lacZ transcriptional fusion and</td>
</tr>
<tr>
<td></td>
<td>recombine into the chromosome at amyE, derivative of pDH32</td>
</tr>
<tr>
<td></td>
<td>(Shimotsu &amp; Henner, 1986)</td>
</tr>
<tr>
<td>Other plasmids</td>
<td></td>
</tr>
<tr>
<td>pMP25.8</td>
<td>Ap Tc Cm; pJH101 derivative (Perego et al., 1987)</td>
</tr>
<tr>
<td>pJM780</td>
<td>Ap Cm; 0-7 kb EcoRI-SalI fragment in pJM103</td>
</tr>
<tr>
<td>pJM3000</td>
<td>Ap Cm; 3 kb SalI fragment in pJM103</td>
</tr>
<tr>
<td>pJM606</td>
<td>Ap Cm; 0-6 kb HindIII fragment subcloned from pJM3000 in pJM103</td>
</tr>
<tr>
<td>pJM244</td>
<td>Ap Cm; 0-2 kb HindIII fragment subcloned from pJM3000 in pJM103</td>
</tr>
<tr>
<td>pJM231</td>
<td>Ap Cm; 0-2 kb HindIII–SalI fragment subcloned from pJM3000 in pJM103</td>
</tr>
<tr>
<td>pJM783/538 EB</td>
<td>Ap Cm; 0-5 kb EcoRI–BamHI linked PCR fragment in pJM783</td>
</tr>
<tr>
<td>pJM115/487 EB</td>
<td>Ap Cm; 0-4 kb EcoRI–BamHI linked PCR fragment in pJM115</td>
</tr>
</tbody>
</table>

Sequencing of the ts-A13 mutation. The products obtained from two independent experiments of PCR amplification of chromosomal DNA of mutant strain PB1679 using the primers at nt 1872 (5' CTTGGTGGCGGGTCTCT 3') and nt 2685 (5' ACCGCGAGGGCGATTGA 3') were digested with HindIII and cloned in pJM103 to be sequenced on both strands. We considered as mutations only the base substitutions verified on clones obtained from both the independent PCR reactions.

β-Galactosidase assay. β-Galactosidase activity in strains harbouring lacZ transcriptional fusions was measured as described previously (Albertini & Galizzi, 1985).

DNA manipulations. Plasmid and phage DNA preparations and Southern hybridization were performed according to standard protocols (Sambrook et al., 1989). B. subtilis chromosomal DNA was prepared by phenol extraction (Albertini & Galizzi, 1985).

RESULTS AND DISCUSSION

Cloning and sequencing of the nrd locus

Plasmid clone pMP25.8 (Perego et al., 1987) was able to marker rescue the temperature-sensitive mutant ts-A13 in transformation experiments and direct selection to ts+. None of the transformants was CmR. This observation indicated that integration by single crossover is probably lethal and that the insert of the plasmid was internal to the nrd locus and not inclusive of the 5' or 3' boundaries. For this reason we started screening a Charon strain A bank (Ferrari et al., 1981) by plaque hybridization using labelled pMP25.8 as probe and identified one positive clone, AG621, with an insert of about 13 kb (Fig. 1). The restriction map of the clone was similar to that of clone ΛA-2T-4 isolated from the same bank and bearing the sequence corresponding to the tr-A13 and glnA locus (Fisher et al., 1984). DNA from phage ΛG621 did in fact transform glnA-100 to prototrophy, supporting the idea that the two clones are identical.

We subcloned the 780 bp EcoRI-SalI fragment of pJM780 and the 3283 bp SalI fragment of pJM3000 in pJM103 (Perego, 1993) and sequenced by primer walking. Analysis of the sequence revealed the existence at the 5' end of an incomplete ORF showing similarity to the major subunit of class I RRs. To extend the cloning in the 5' direction, we tried, unsuccessfully, different strategies, i.e. chromosome walking with various integrative plasmids and screening of different B. subtilis λ or plasmid libraries. The sequence was obtained by reverse PCR and direct sequencing of the PCR product (see Methods).

Organization of the nrd locus

Analysis of the sequence comprising the nrd locus and the surrounding region, performed using the BLAST (Altschul et al., 1990) and FASTA (Pearson, 1990) collections of programs, revealed the following features. The first 865 bp are characterized by short, less than 279 bp ORFs on both strands, whose deduced products did not show any similarity to known proteins. Nevertheless, we found the consensus for a putative sA promoter at nt 51-80 and a hairpin-loop structure (AG−114 kJ mol⁻¹) at nt 745-776, suggesting that the sequence corresponds to a transcription unit without coding capacities.

A second consensus for a putative sA promoter could be recognized between nt 865 and 891, followed at nt 921-926 by a putative ribosome binding site just upstream of the GTG start codon of an ORF, named ymaA (Fig. 1). The predicted polypeptide of 130 aa (expected M, 14600) showed significant similarity to the deduced product of
short ORFs of unknown function just upstream of the\n
\textit{nrdEF} operon of \textit{S. typhimurium} (ORFB) and \textit{E. coli} (ORFC) (Jordan \textit{et al.}, 1996). A similar polypeptide was also predicted in \textit{Lactococcus lactis} (GenBank accession no. X92690) and in \textit{Mycoplasma genitalium} (predicted coding region MG230, Fraser \textit{et al.}, 1995). In all cases the ORFs seem to be part of an \textit{nrdEF} operon.

A ribosome binding site consensus at nt 1276–1283 precedes an ORF encoding a polypeptide of 700 aa, with a predicted \( M_r \) of 80700, showing high similarity (identity score 44–53\%) to the second group of enterobacterial class I RRs described by Jordan \textit{et al.} (1994a, b). The same type of RR was isolated from \textit{Mycoplasma tuberculosis} (Yang \textit{et al.}, 1994) and deduced to be present in \textit{Myc. genitalium} (Fraser \textit{et al.}, 1995). The high similarity is evident in the multiple alignment shown in Fig. 2, and extends beyond the conserved three cysteines of the essential thiol, characterizing all the large RR subunits of class I. As in the case of the NrdE subunits of \textit{Enterobacteriaceae}, the similarity with the prokaryotic NrdA subunit is low (22\% identity), even lower than the similarity observed with the eukaryotic RR major subunits (25\% identity).

A ribosome binding site consensus at nt 3397–3401 precedes a sequence (nt 3411–4400) potentially encoding a polypeptide of 329 aa, with an expected \( M_r \) of 38400, with high similarity to the second group of class I minor subunits of RR, namely NrdF, found in \textit{Enterobacteriaceae} (Jordan \textit{et al.}, 1994a, b) and \textit{Myc. genitalium} (Fraser \textit{et al.}, 1995). Fig. 2 shows the alignment between the NrdF polypeptides sharing identity scores between 39 and 54\%.

In this case as well, the similarity extends beyond the conservation of residues involved in defined functions such as the major subunit binding surface, the environment of the tyrosyl residue, the hydrogen bonding network, the iron ligands and the four \( \alpha\)-helix domains surrounding them. The similarity with other class I minor subunits is lower, ranging from 23\% identity with the eukaryotic ones and 21\% with NrdB of \textit{E. coli} or bacteriophage T4.

No terminator site or recognizable promoter was identified between \textit{ymaA} and \textit{nrdE}, or between \textit{nrdE} and \textit{nrdF}.

In addition, the general organization of the coding sequences is very similar to that of the characterized \textit{nrdEF} operon of \textit{S. typhimurium} and \textit{E. coli} (Jordan \textit{et al.}, 1996). We thus propose that the three ORFs belong to the same operon.

It should be pointed out that the \textit{B. subtilis nrd} operon, even though highly similar to the enterobacterial \textit{nrdEF} operon, lacks a gene encoding a polypeptide similar to glutaredoxin, the proposed electron donor for the NrdEF enzyme (Jordan \textit{et al.} 1996). Such an ORF is not present even in the proximity of the \textit{nrd} locus.

The stop codon of the putative NrdF coding region is part of the loop of a stem–loop structure with a stability of \(-114 \text{kJ mol}^{-1}\), followed by a putative ribosome binding sequence (nt 4432–4439) which precedes an ATG codon. The following ORF (nt 4450–5070), named \textit{ymaB}, could encode a polypeptide of 206 aa, with a predicted \( M_r \) of 23400. It should be noted that a partial consensus to a \( rA \) promoter, which could drive the transcription of \textit{ymaB}, is present between nt 4345 and 4372.

The hypothetical product YmaB is devoid of similarity to known proteins or to hypothetical products present in data banks. The ORF is followed by a hairpin–loop structure (Fig. 1) that could act as a strong (\( \Delta G = -169.9 \text{kJ mol}^{-1}\)) terminator for the \textit{nrd} operon and for the transcription unit on the complementary strand (ORF from nt 5559 to 5134), that corresponds to the 3' end of the \textit{cv/c} gene described by Kuroda \textit{et al.} (1993).

\textbf{Inactivation of the three main coding sequences}

The three subclones pJM606, pJM244 and pJM231 (described in Table 2) were used as insertion vectors to specifically interrupt, by insertion of the plasmid, the three major ORFs, i.e. \textit{nrdE}, \textit{nrdF} and \textit{ymaB}, respectively. In all instances the restriction fragments used to target the insertion of the plasmid in the chromosome forced the recombination event in an internal region of the gene as indicated in Fig. 1.

We used these plasmids to transform competent cells of \textit{B. subtilis} to \textit{Cm}R and we never obtained transformants. This result is similar to that obtained with plasmid pMP25.8. Since we had no difficulty in obtaining transformants by a single crossover event using plasmid pJM3000, bearing the \( C \) terminus of \textit{nrdE} and the entire coding regions of \textit{nrdF} and \textit{ymaB}, we deduced that the interruption of the operon was probably lethal and that at least the further downstream ORF \textit{ymaB} encodes an essential function.

\textbf{Mapping and sequencing of the ts-A13 mutation}

Bazill & Karamata (1972) analysed the conditional lethal mutant, \textit{ts-A13}, impaired in ribonucleotide reductase activity. It is the only conditional lethal mutant reported for this function in \textit{B. subtilis}, suggesting that the mutation impairs the functionally relevant class I enzyme of this organism.

Perego \textit{et al.} (1987) isolated pMP25.8, an integrative plasmid able to marker rescue to ts\textsuperscript{+} the temperature-sensitive mutant \textit{ts-A13}. To define more precisely the region of the mutation, we subcloned a number of DNA fragments and tested them in marker rescue experiments.

Among the subclones, plasmid pJM606, unable to transform cells to \textit{Cm}R, was able to give ts\textsuperscript{+} transformants when used to transform strain PB1679. The same \textit{HindIII} fragment was obtained by PCR amplification from chromosomal DNA of the mutant strain PB1679 and was sequenced. We found three mutations: two in close proximity (\textit{GC} \rightarrow \textit{AT} transition at nt 2292; \textit{AT} \rightarrow \textit{GC} transition at nt 2295) were silent substitutions in the third position of the codons for Lys333 and Glu334 of NrdE and the third mutation was an ACG \rightarrow TA transition, changing His255 of NrdE to Tyr. This histidine residue is conserved in all class I major subunits of RR so far examined, and corresponds to the last residue (His308) of
Fig. 2. For legend see facing page.
Fig. 2. Phylogenetic analysis of the class I major subunits (a) and minor subunits (b) of RR graphically displayed using the NIPLOT software of M. Gouy (University of Lyon, France). The scale bars indicate the distances as calculated from the multiple alignment (Thompson et al., 1994). Species abbreviations and protein or DNA sequence accession numbers used for (a) are: NrdEB.s., B. subtilis (Z68500); NrdES.t., S. typhimurium (Q08698); NrdEM.t., Myco. tuberculosis (L34407); NrdEM.g., Myc. genitalium (U39701); NrdAE.c., E. coli (P00452); NrdAH.i., H. influenzae (P43754); NrdAS.t., S. typhimurium (P37426); NrdAT4, bacteriophage T4 (P32282); Rir1S.c., H. sapiens (P23921); Rir1C.e., Caenorhabditis elegans (Q036304); Rir1V.v., Vaccinia virus (P20503); Rir1P.f.; Plasmodium falciparum (A48412); Rir1S.c., Saccharomyces cerevisiae (P21524). For (b) : NrdFS.t., B. subtilis (Z68500); NrdFS.t., S. typhimurium (P17424); NrdFE.c., E. coli (P37146); NrdFM.g., Myc. genitalium (P47471); NrdFE.c., E. coli (P00453); NrdBS.t., S. typhimurium (P37427); NrdBH.i., H. influenzae (P43755); NrdBT4, bacteriophage T4 (P11156); Rir2P.f., P. falciparum (B49412); Rir2H.s., H. sapiens (P31359); Rir2He.s., Herpes simplex virus (P28847); Rir2S.s., Spisula solidissima (P07201); Rir2N.t., Nicotiana tabacum (X92443); Rir2C.e., C. elegans (P42170); Rir2S.p., Schiz. pombe (P36603); Rir2S.c., Sacch. cerevisiae (P09938); Rir2D.d., Dictyostelium discoideum (P42521).

Fig. 2. Multiple alignments (a) of NrdE polypeptides, major subunits of RR, and (b) of NrdF polypeptides, minor subunits of RR, found in B. subtilis (B.s.), E. coli (E.c.), S. typhimurium (S.t.), Myc. genitalium (M.g.) and Myco. tuberculosis (M.t.). GenBank accession numbers are given in the legend to Fig. 3. (a) In the NrdE alignment, • indicates the Cys residues proposed for the 3' hydrogen abstraction of the substrate and ▼ indicates the two Cys residues of the active site. (b) In the NrdF alignment, the functionally relevant conserved residues are indicated by the following symbols: •, hydrogen bonding network sites; ◦, residues of the tyrosyl radical environment; ◇, residues of the major subunit binding surface; *, iron ligand residues.

the βC sheet of the α/β barrel domain described in the three-dimensional structure of the NrdA subunit (Uhlin & Eklund, 1994). We infer that this mutation is the one responsible for the temperature-sensitive phenotype of the ts-A73 mutant allele.

The identification of three mutations, two neutral and one that is presumably responsible for the conditional lethal phenotype, in a short stretch of DNA sequence can be explained by the fact that the mutant ts-A73 was isolated after NTG treatment, a mutagen which acts on the DNA replicative fork and is known for inducing multiple mutations (Karamata & Gross, 1970).

The localization in nrdE of the mutation that impairs the activity of the class I RR (Bazil & Karamata, 1972) and the evidence that the nrdEF operon cannot be inactivated led us to infer that this operon exerts a role essential for the biosynthesis of deoxyribonucleotides. It seems safe to conclude that the NrdER1/NrdFR2 enzyme plays the

3001
same role in *B. subtilis* that the NrdAR\_1/NrdBR\_2 enzyme plays in *Enterobacteriaceae*. In *E. coli* and in *S. typhimurium*, the nrdEF operon, even if expressed, is not essential, and in standard conditions is unable to replace the absence of NrdAR\_1/NrdBR\_2 activity (Jordan *et al.*, 1996). Although at present we cannot rule out the hypothesis that an nrdAB-type operon is present in the *B. subtilis* genome, it appears that *B. subtilis* made a functional choice concerning the RR enzyme different from that made by the *Enterobacteriaceae*. The same choice was made by *M. genitalium*: in the whole genome only one locus is devoted to deoxyribonucleotide biosynthesis and it appears to be organized in an operon with three ORFs, the *nrdF* homologue, the homologue of *B. subtilis* *ymaA* and the *nrdE* homologue. With the exception of the relative order of genes, the genetic tools present in *Myc. genitalium* are the same as found in *B. subtilis*. *Haemophilus influenzae*, the other bacterium for which the entire genome has been sequenced (Fleischmann *et al.*, 1995), made the opposite choice: only an nrdAB operon is present.

Prokaryotes evolved different strategies for deoxyribonucleotide production (Reichard, 1993), namely class I and class II aerobic and class III anaerobic enzymes. The three classes are not related at the level of primary structure. In addition, if we examine the variety of aerobic Fe(II1)-tyrosyl radical-based class I enzymes, a category comprising the eukaryotic enzymes, we observe a wide diversification at the level of primary structure. Fig. 3 represents the phylogenetic analysis of RR major and minor subunits: the NrdER\_1/NrdFR\_2 RR diverged from NrdAR\_1/NrdBR\_2 enzymes at least as much as both diverged from the eukaryotic enzymes. In the case of the major subunits, the NrdA-type polypeptides are more similar to the eukaryotic counterpart than to the NrdE prokaryotic polypeptide.

New questions arise from these observations. Why did *Enterobacteriaceae* and *H. influenzae* make a functional choice (in the case of *Enterobacteriaceae*, maintaining both the genetic equipment and the production at low levels of the NrdER\_1/NrdFR\_2 enzyme) different from *B. subtilis* and *Myc. genitalium*? Is it possible that an ancestor of *Enterobacteriaceae* acquired the nrdAB genes from a bacteriophage, such as T4, which has an RR highly similar to the NrdAR\_1/NrdBR\_2 enzyme, and the newly acquired enzyme has been raised to a higher functional role in the course of evolution? Is the choice of the *Enterobacteriaceae* in some way related to the presence of the anaerobic nrdD-encoded class III enzyme, proposed as the closest relative of the common ancestor of all RRs (Reichard, 1993)? In more general terms, why do we find class II enzymes in Eubacteria (e.g. *Lactobacillus leichmanii*; Booker & Stubbe, 1993) and so many types of class I enzymes?

The completion of the sequencing projects of the entire genome of *B. subtilis* and other bacteria will answer some of these questions. At present the finding that the nrd genes of *Myc. genitalium* are phylogenetically closer to the nrdEF genes of *B. subtilis* than to the nrdAB genes of *Enterobacteriaceae* confirms other observations indicating an evolutionary relationship between *Mycoplasma* and Gram-positive bacteria (Fraser *et al.*, 1995).

### Table 3. Expression of the nrd operon

The indicated values represent the mean of at least two experiments and were the highest obtained from a series of samples assayed at various times during growth. ND, not determined.

<table>
<thead>
<tr>
<th>Strain</th>
<th>SM medium</th>
<th>SM medium</th>
<th>TM medium</th>
<th>TM medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>PB1848 wild-type</td>
<td>13</td>
<td>14</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>PB1849 wild-type</td>
<td>12</td>
<td>ND</td>
<td>14:2</td>
<td>ND</td>
</tr>
<tr>
<td>PB1850 Thy^-</td>
<td>28:5</td>
<td>35</td>
<td>23:1</td>
<td>50:6</td>
</tr>
<tr>
<td>PB1851 Com^-</td>
<td>11:5</td>
<td>ND</td>
<td>8:5</td>
<td>ND</td>
</tr>
<tr>
<td>PB1852 Rec^-</td>
<td>10</td>
<td>ND</td>
<td>8</td>
<td>ND</td>
</tr>
<tr>
<td>PB1854 Com^- Thy^-</td>
<td>22</td>
<td>33:5</td>
<td>28:5</td>
<td>52</td>
</tr>
<tr>
<td>PB1855 Rec^- Thy^-</td>
<td>21</td>
<td>28:2</td>
<td>21:5</td>
<td>16:5</td>
</tr>
</tbody>
</table>

* SM, Schaeffer Medium; TM, competence minimal medium (Anagnostopoulos & Spizizen, 1961). Both media were supplemented with 50 μg thymidine ml\(^{-1}\). (—) indicates thymidine deprivation.
suggesting the absence of a functional promoter in the intercistronic sequence (data not shown).

We propose that ymaB is co-transcribed with ymaA, nrdE and nrdF, and thus is part of the nrd operon.

To study the level of expression and to perform a preliminary analysis of the transcriptional control of the nrd operon, we used a lacZ transcriptional fusion, inserted into the chromosome downstream of ymaB. The fusion by means of plasmid pJ783/538 EB gave the results summarized in Table 3 and represented in Fig. 4. The level of transcription was low, but significant in rich and minimal media, and the presence of the mutations Δcomp-A and recA-260 did not change the basal level of transcription significantly.

As in the case of the nrdAB operon of E. coli (Filpula & Fuchs, 1978), the transcription of RR genes increased as a consequence of a block of DNA synthesis obtained by thymidine starvation in the thymidine-requiring strain PB1850 (Fig. 4). Since this effect was higher in TM medium, a medium for the induction of competence in B. subtilis (Anagnostopoulos & Spizizen, 1961), we tested if it was part of the competence response, and thus positively controlled by the response regulators ComP-A (Dubnau, 1993), or part of the RecA-controlled responses (Love & Yasbin, 1984).

In the Δcomp-A mutant strain PB1854, the peak of induction by thymidine starvation was comparable to that reached by the parental strain (Table 3). In contrast, the recA mutant PB1855 was severely affected in its ability to induce the operon after thymidine starvation (Fig. 4).

We can explain these results by the hypothesis that in B. subtilis, as in E. coli for nrdAB, the nrdEF operon is under the control of cell cycle and as a consequence under the influence of the SOS regulator RecA. In B. subtilis, RecA synthesis is induced both by thymidine starvation (Love & Yasbin, 1984) and by the attainment of the competence phase (Yasbin et al., 1993). Since in the Δcomp-A mutant, thymidine starvation was still effective in inducing transcription of the nrd operon, we conclude that the increased level of expression of the operon is not due to the attainment of the competent state. The nrd induction can be tentatively classified as a Type I SOS phenomenon following the indication of Yasbin (1993).

These preliminary observations reinforce the suggestion that this is the functionally relevant operon in the synthesis of deoxyribonucleotides and is subject to fine control of expression to modulate their synthesis.

ACKNOWLEDGEMENTS

This work was supported in part by the European Commission under the Biotechnology programme (contract Bio2-CT93-0272) and by Consiglio Nazionale delle Ricerche (Grants 92.02183.CT174 AND 93.02009.CT14).

REFERENCES


Fig. 4. β-Galactosidase assays of the nrd-lacZ fusion monitoring the induction of operon transcription by thymidine starvation. Half of the cultures in TM medium supplemented with thymidine (50 μg ml⁻¹) were collected at the indicated times (arrows), washed by centrifugation and resuspended in the same medium without thymidine. ●, wild-type strain; ○, wild-type strain starved of thymidine; ▲, recA strain; ▽, recA strain starved of thymidine.

In the Δcomp-A mutant strain PB1854, the peak of induction by thymidine starvation was comparable to that reached by the parental strain (Table 3). In contrast, the recA mutant PB1855 was severely affected in its ability to induce the operon after thymidine starvation (Fig. 4).

We can explain these results by the hypothesis that in B. subtilis, as in E. coli for nrdAB, the nrdEF operon is under the control of cell cycle and as a consequence under the influence of the SOS regulator RecA. In B. subtilis, RecA synthesis is induced both by thymidine starvation (Love & Yasbin, 1984) and by the attainment of the competence phase (Yasbin et al., 1993). Since in the Δcomp-A mutant, thymidine starvation was still effective in inducing transcription of the nrd operon, we conclude that the increased level of expression of the operon is not due to the attainment of the competent state. The nrd induction can be tentatively classified as a Type I SOS phenomenon following the indication of Yasbin (1993). These preliminary observations reinforce the suggestion that this is the functionally relevant operon in the synthesis of deoxyribonucleotides and is subject to fine control of expression to modulate their synthesis.

ACKNOWLEDGEMENTS

This work was supported in part by the European Commission under the Biotechnology programme (contract Bio2-CT93-0272) and by Consiglio Nazionale delle Ricerche (Grants 92.02183.CT174 AND 93.02009.CT14).

REFERENCES


