ELSEVIER

Gene 217 (1998) 83–90

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Gene organization in the *trxA/B–oriC* region of the *Streptomyces coelicolor* chromosome and comparison with other eubacteria[1]

Ohad Gal-Mor [a], Ilya Borovok [a], Yossi Av-Gay [b], Gerald Cohen [a], Yair Aharonowitz [a,*]

[a] *Department of Molecular Microbiology and Biotechnology, The George S. Wise Faculty of Life Sciences, Tel-Aviv University, Ramat Aviv 69978, Israel*

[b] *The Faculty of Medicine, Division of Infectious Diseases, University of British Columbia, 2733 Heather Street, Vancouver, BC V5Z 3J5, Canada*

Received 15 May 1998; accepted 26 June 1998; Received by H. Cedar

Abstract

The gene organization was determined in the *trxA/B–rnpA* region of the *Streptomyces coelocolor* chromosome, near to the origin of replication, *oriC*. Previously, we showed that the *trxA* and *trxB* genes, coding for thioredoxin and thioredoxin reductase, respectively, occur in *S. coelicolor* as a gene cluster and are contained on a cosmid H24 that carries *oriC* and several genes involved in DNA replication. Here we show that the *trxA/B* locus is positioned approx. 9.4 kb from *oriC*, present the nucleotide sequence of the *trxA/B–rnpA* region and use sequence analysis to identify the nature of the intervening genes. Seven open reading frames were found, all oriented in the same direction, five of which were identified as the *S. coelicolor* homologs of SpoIIIJ, Jag, GidB, Soj and SpoOJ in *Bacillus subtilis* and which have been ascribed different functions in this and other bacteria for either DNA replication, chromosomal partitioning or morphological development. The arrangement of the genes coding for the above five proteins in the *trxA/B–rnpA* region in *S. coelicolor* resembles that in *Mycobacterium leprae*, *Mycobacterium tuberculosis, B. subtilis* and *Pseudomonas putida*, and supports the view that many of the genes necessary for development and cell division in bacteria are organized in a similar fashion. In *B. subtilis* and *P. putida*, however, the *trxA/B* genes are not present in the above gene arrangement.

Keywords: Origin of DNA replication; Sporulation and chromosomal partitioning genes

1. Introduction

Streptomyces are filamentous soil bacteria that possess a complex morphological life-cycle and are remarkable for the production of an extraordinary variety of secondary metabolites (Chater, 1993; Berdy, 1984; Miyadoh, 1993). Much of our current understanding of the interrelationship between these processes has come from knowledge of the structure of the *Streptomyces* genome and, in particular, that of *Streptomyces coelicolor*, the most studied member of this genus from a genetic standpoint (Hopwood and Kieser, 1990; Hopwood et al., 1994). It has an 8 Mb linear chromosome that has been physically mapped (Kieser et al., 1992). These studies, in conjunction with those in other eubacteria, notably *Escherichia coli* and *Bacillus subtilis* (Blattner et al., 1997; Kunst et al., 1997), whose genomes have been entirely sequenced, and those of *Mycobacterium leprae* and *Mycobacterium tuberculosis* which are near to completion, now permit a far greater insight into the physical organization and behavior of genes involved in key cellular activities than could have been imagined just a few years ago.

Our studies on the role of thioredoxin in *Streptomyces* led us to clone and sequence the two genes, *trxA* and

* Corresponding author. Tel: 972-3-6409833; Fax: 972-3-6409407; e-mail: F59yair@ccsg.tau.ac.il

[1] The EMBL/GenBank/DDBJ accession number for the sequence reported in this paper is Y16311.

Abbreviations: aa, amino acid(s); *dnaA*, gene encoding DnaA initiator protein; *gidB*, gene encoding GidB cell division inhibitor protein; *jag*, gene encoding *Bacillus subtilis* Jag sporulation protein; ORF (*orf*), open reading frame; *oriC*, origin of replication; PCR, polymerase chain reaction; RBS, ribosomal binding site; *rnpA*, gene encoding protein sub unit of ribonuclease P; *soj*, gene encoding *B. subtilis* Soj sporulation protein; *spoOJ*, gene encoding *B. subtilis* SpoOJ sporulation and chromosomal partitioning protein; *spoIIIJ*, gene encoding *B. subtilis* SpoIIIJ sporulation protein; *thdF*, gene encoding protein for thiophen and furan oxidation; *trxA*, gene encoding thioredoxin; *trxB*, thioredoxin reductase.

0378-1119/98/$19.00
PII: S0378-1119(98)00357-6

trxB, that comprise that system and which code for thioredoxin and thioredoxin reductase, respectively (Cohen et al., 1993). Thioredoxins are small heat-stable proteins which, in conjunction with thioredoxin reductase and NADPH, serve as a source of electrons for the reduction of disulfides in low molecular weight compounds and in proteins, and play an important role in the control of DNA synthesis as a cofactor in the reduction of ribonucleotides to deoxyribonucleotide by ribonucleotide reductase. Surprisingly, the *trxA/B* genes in *Streptomyces* spp. occur in a cluster and this arrangement has now also been found in another member of the high G+C branch of the Gram-positive bacteria, namely *Mycobacteria* (Wieles et al., 1995). Physical mapping of the *trx* locus in *S. coelicolor*, using an ordered set of overlapping cosmids, established that the *trx* gene cluster is located near to the chromosomal origin of replication, *oriC* (Av-Gay, 1994). Earlier studies by Calcutt (Calcutt and Schmidt, 1992; Calcutt, 1994) determined the gene organization in the *oriC* region of the *S. coelicolor* chromosome as *rnpA–rpmH–dnaA–oriC–dnaN–recF–gyrB–gyrA*. This gene arrangement markedly resembles that in the *dnaA* region of numerous other Gram-negative and Gram-positive bacteria and is frequently linked to *oriC* (for a review see Ogasawara and Yoshikawa, 1992 and Salazar et al., 1996). Marked conservation of gene order also occurs in the region immediately upstream of *dnaA–oriC* in *B. subtilis, Pseudomonas putida*, *M. leprae* and *M. tuberculosis* (Ogasawara and Yoshikawa, 1992; Fsihi et al., 1996; Philipp et al., 1996). In this work we report the sequence of a 6.7 kb region of *S. coelicolor* DNA that connects *trxA/B* to *rnpA* and which is located directly upstream of *dnaA–oriC*. Sequence analysis reveals this region to contain several genes associated with DNA replication, morphological development and chromosomal partitioning in a gene organization that is similar to that found in *B. subtilis, P. putida, Coxiella burnetii*, *M. leprae* and *M. tuberculosis*.

2. Materials and methods

2.1. Bacterial strains, plasmids and growth conditions

A plasmid-free derivative of *S. coelicolor* A3(2) M145, served as the source of genomic DNA. Cosmid H24 belongs to an ordered overlapping library of *S. coelicolor* A3(2) chromosomal DNA and contains the *trxA/B* gene cluster and the *oriC* replication region (Redenbach et al., 1996). Cosmid H24 was propagated in *E. coli* DH5α; plasmids pUC18 and pBR322 were grown in *E. coli* XL1. Growth of *E. coli* strains was in LB medium (Sambrook et al., 1989).

2.2. Cloning of DNA, and PCR techniques

All standard manipulations of DNA for restriction digestion, ligation and transformation were carried out as described (Sambrook et al., 1989); electroporation was performed as described (Smith et al., 1990). Plasmid and cosmid DNA were isolated using a kit (Qiagen, Boehringer Mannheim) according to the vendor's instructions. PCR was performed using the Expand Long Template PCR system (Boeringer Mannheim). Based on the assumption that the DNA region spanning *trxA/B–oriC* in *S. coelocolor* is likely to be similar to that in *M. leprae* (Fsihi et al., 1996), we anticipated the separation of *trxA/B* and *oriC* to be no more than 10 kb. Primers of 18–22 nucleotides were prepared to the known sequences in the *S. coelicolor trxA* (unpublished results) and *rnpA* genes (Calcutt, 1994) and used in PCR reactions with cosmid H24 as template. Amplified DNA was purified after gel electrophoresis using a Wizard DNA Purification System (Promega). The 6.7 kb PCR fragment obtained was partially sequenced to verify that the terminal nucleotide sequences matched those of the *trxA* and *rnpA* sequences. Restriction mapping of cosmid H24, performed in parallel with that of the PCR fragment, showed that the entire *trxA–rnpA* region was contained in two overlapping 6.7 kb *Bam*HI and 2.3 kb *Hin*dIII fragments which were cloned into pBR322 and pUC18 vectors, respectively.

2.3. DNA sequencing

Double-strand sequencing of the cloned 6.7 kb region was carried out with the ABI Prism 377 Automatic Sequencer using some 31 primers (Biotechnology General, Israel; Gibco, UK) and the ABI Prism Dye Terminator Cycle Sequencing kit (Applied Biosystems).

2.4. Informatics

DNA and protein sequences from predicted ORFs were compared with non-redundant databases (EMBL, SWISS PROT, PIR) at the National Center for Biotechnological Information, Washington, DC, USA, using BLAST programs (Altschul et al., 1994, 1997). Additional computer analysis was performed using version 9 of the GCG package of programs (Genetics Computer Group, University of Wisconsin). Homologous proteins were aligned with PILEUP and BESTFIT, and molecular weights and isoelectric points were calculated with PEPTIDESORT. DNA sequences were aligned with BESTFIT.

3. Results and discussion

Previously we showed by hybridization of an ordered overlapping cosmid library of *S. coelicolor* that the

trxA/B gene cluster is localized in the cosmid designated H24 which contains *oriC* (see Redenbach et al., 1996). To determine the nature and organization of the genes located within the *trxA/B–oriC* region (defined here as upstream of *dnaA*), we sequenced the 6.7 kb region connecting *trxA/B* to *rnpA*. The nucleotide sequence was analysed for the location of ORFs.

3.1. Assignment of ORFs

Seven ORFs were identified in the 6.7 kb region. The directions of all seven ORFs are oriented towards the *trx* locus and away from *oriC* (see Fig. 1). Five of the ORFs show strong sequence similarity to genes involved in sporulation and cell division in *B. subtilis* and to genes involved in the regulation of cell division in other bacteria. Fig. 1 presents the overall organization of the genes in the *trxA/B–rnpA* region of *S. coelicolor* and compares their organization with that in other bacteria and indicates their putative functions. Table 1 lists the ORFs, their positions within the 6.7 kb region and the size of the expected protein products. All the structural genes start with an ATG, GTG or TTG translational initiation codon, are oriented towards *trxA/B* and are preceded by recognizable *Streptomyces* Shine Delgarno (SD) sequences (Strohl, 1992). Table 2 shows the similarity of the putative ORF products to proteins in the databases.

(1) *orf*124 (10 kDa-like) is located directly upstream of *rnpA* and is predicted to code for a 124 aa protein with a deduced molecular mass of 13.6 kDa. Its start codon overlaps by one nucleotide the termination codon of *rnpA*. Database analysis showed that the predicted protein of *orf*124 possesses significant sequence similarity to a group of 9–10 kDa bacterial proteins that are encoded by genes that map near the chromosomal origin of replication, between *rnpA* and a gene coding for a 60 kDa protein (see below). Sequence identity among these ORFs and the predicted protein of *orf*124 is given in Table 2.

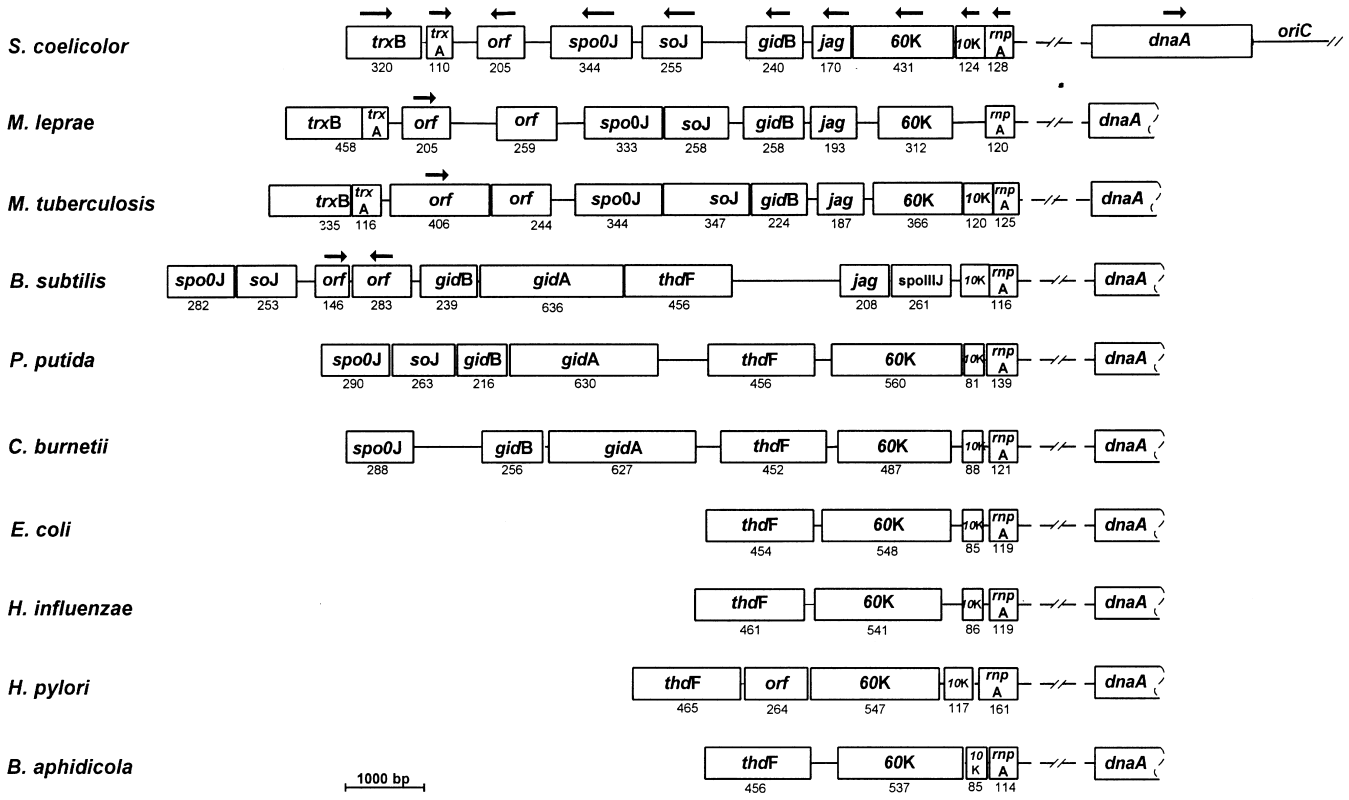

Fig. 1. Comparison of gene organization in the *trxA/B–rnpA* region of eubacterial chromosomes. The chromosomal gene organization is shown for *Streptomyces coelicolor, Mycobacteria leprae, Mycobacteria tuberculosis, Bacillus subtilis, Pseudomonas putida, Coxiella burnetii, Escherichia coli, Haemophilis influenzae, Helicobacteria pylori* and *Buchnera aphidicola*. Arrows indicate gene orientations; unless shown otherwise, homologous genes have the same orientation. Numbers denote the size in amino acids of the predicted proteins. Gene designations and functions are: *dnaA* codes for the initiator protein for chromosomal replication; *rnpA* and *rpmH* code for the protein subunits of ribonuclease P and the ribosomal protein L34, respectively; *S. coelicolor orf*124 and its homologs code for a 10 kDa-like protein of unknown function; *spoIIIJ* codes for a homolog of the *B. subtilis* sporulation stage-III protein; *jag* codes for a homolog of the *B. subtilis* sporulation protein; *gidA*, *gidB* code for glucose inhibition of division proteins; *thdF* codes for a protein determining thiophen and furan oxidation; *soj* codes for a homolog of the *B. subtilis* sporulation protein; *spoOJ* codes for a homolog of the *B. subtilis* protein involved in sporulation and chromosomal partitioning; *S. coelicolor orf*205 and its homologs *orf*259 and *orf*244 from *M. leprae* and *M. tuberculosis*, respectively, code for a protein of unknown function; *trxA* and *trxB* code for thioredoxin and thioredoxin reductase, respectively. Further details of assignment of gene functions are given in the text and in Table 2.

Table 1
Genes and predicted properties of the deduced protein products in the *trxA/B–rnpA* region of the *S.coelicolor* chromosome

ORF	Position (from–to)	Length (aa)	Codon (first…last)	SD sequences and initiation codons[a]	Molecular mass (kDa)	pI
*orf*124	92–467	124	ATG…TGA	GCTACTGGGAGGGGGCGCGCG**ATG**	13.59	10.7
*orf*431	471–1766	431	GTG…TAA	ATGCCCCAAGGAGCATGATTA**GTG**	47.29	10.86
jag	1782–2294	170	GTG…TGA	AAGTAAGAAGGAGTCCATCCC**GTG**	18.32	4.31
gidB	2419–3138	239	GTG…TAG	CCATACGGAAGGACGGTCCCC**GTG**	25.61	8.13
soj	3690–4457	255	ATG…TGA	CGACCCGAGCAGACCCGGGTC**ATG**	27.19	4.79
spo0J	4588–5622	344	TTG…TGA	ACGCCCCGACCGAGAAGTCGG**TTG**	37.61	5.45
*orf*205	5960–6577	205	ATG…TAG	AGGAAGCGAGGAACACCCTTC**ATG**	22.53	7.30

[a]Apparent Shine Delgarno (SD) ribosomal binding sites shown underlined, initiation codons shown in bold.

Unlike the other 9–10 kDa proteins, the *S. coelicolor* putative protein contains an additional C-terminal portion of about 45 aa, accounting for its larger size. Although homologous genes to *orf*124 are present in virtually all the known sequenced *oriC* regions, the function of the group of proteins they encode is not known.

(2) *orf*431 (*spoIIIJ*-like) is located immediately upstream of *orf*124 and is predicted to code for a 431 aa protein with a deduced molecular mass of 47.3 kDa. Its GTG start codon is positioned just three nucleotides in front of the *orf*124 termination codon. Database analysis showed that the predicted protein of *orf*431 resembles a number of related 60 kDa inner membrane proteins of largely unidentified function from a wide range of bacteria (see Table 2). However, in one case, that of *B. subtilis*, the related protein SpoIIIJ is well characterized and plays a key role in sporulation through selective transcription of genes in the prespore and mother cell during stage III development (Errington et al., 1992). In the eubacteria the gene encoding the 60 kDa protein is highly conserved and located near to *oriC*, and in *B. subtilis* forms part of a bicistronic operon containing a second gene, *jag*, that too is involved in morphological development (Errington et al., 1992, and see below).

(3) *orf*170 (*jag*-like) is located upstream of *orf*431 and is predicted to code for a 170 aa protein of molecular mass 18.3 kDa. Its GTG start codon is positioned 15 nucleotides in front of the *orf*431 termination codon. A database search revealed that the predicted protein of *orf*170 is similar to the *B. subtilis* Jag protein, which is associated with SpoIIIJ, and is necessary for the third stage of sporulation (Errington et al., 1992), and to L222 *orf*143 in *M. leprae* (see Table 2).

(4) *orf*239 (*gidB*-like) is located upstream of *orf*170 and is predicted to code for a 239 aa protein of molecular mass 25.6 kDa. Its GTG start codon is positioned 125 nucleotides in front of the *orf*170 termination codon. A database search showed that the predicted protein of *orf*239 is closely related to the *gidB* (glucose inhibited division) gene product of *M. leprae* and other bacteria and is probably involved in the regulation of cell division (Ogasawara and Yoshikawa, 1992, and see Table 2).

(5) *orf*255 (*soj*-like) is located upstream of *orf*239 and is predicted to code for a 255 aa protein of molecular mass 27.2 kDa. Its ATG start codon is, by comparison with the above *orfs*, well separated by some 551 nucleotides from the termination codon of the preceding *orf*239. The predicted protein of *orf*255 was found in database searches to be related to proteins encoded by two groups of genes involved in sporulation and chromosomal partitioning. In the first group, similarity was found with the Soj protein of *B. subtilis* which functions to inhibit initiation of sporulation, whereas SpoOJ (see below) antagonizes that inhibition (Ireton et al., 1994; Sharpe and Errington, 1996), and to related proteins from *M. leprae*, *P. putida* and other bacteria (see Table 2). In the second group, similarity was found to the ParA and RepA family of proteins involved in chromosome and plasmid partitioning (see Table 2). These proteins belong to the ATPase superfamily and are characterized by the presence of three well-conserved amino-acid motifs, A, A′ and B (Koonin, 1993). Inspection of the amino-acid sequence of the predicted protein of *orf*255 showed it to contain all three sequences. Recent studies in *B. subtilis* indicate that *soj* does not appear to be strictly required for chromosome partitioning (Sharpe and Errington, 1996).

(6) *orf*344 (*spoOJ*-like) is located upstream of *orf*255 and is predicted to code for a 344 aa protein of molecular mass 37.6 kDa. Its TTG start codon is separated by 130 nucleotides from the *orf*255 termination codon. In analogy to *orf*255, database searches showed that the predicted protein product of *orf*344 possesses similarity to two groups of proteins involved in morphological development and partitioning. In the first group, similarity was found with SpoOJ which functions in *B. subtilis* for the initiation of sporulation and for normal chromosomal segregation. Null mutations in *spoOJ* cause

Table 2
Similarity of the predicted ORFs present in the *trxA/B–rnpA* region of the *S. coelicolor* chromosome with proteins in the database

orf	Sequence similarity	BLAST score (*p*-value)[a]
*orf*124	● 9–10 kDa protein in rnpA 3′ region from:	
	Coxiella burnetii 9.9 kDa (sp/P45649)[b]	2.4e-23
	Proteus mirabilis 9.6 kDa (sp/P22834)[b]	1.2e-22
	Bacillus subtilis (gb/AF008220)[c]	1.0e-22
	Escherichia coli 9.4 kDa (sp/P22847)[b]	2.1e-21
	Pseudomonas putida 9.2 kDa (sp/P25753)[b]	7.0e-19
	Haemophilus influenzae 9.5 kDa (sp/P44972)[b]	1.8e-19
	Buchnera aphidicola 9.9 kDa (sp/P29432)[b]	3.5e-11
	Mycobacterium tuberculosis MTV028.13c (gb/AL021426)[b]	6.0e-10
*orf*431	● 60 kDa inner-membrane protein from:	
	Mycobacterium leprae (gb/L39923)[b]	5.1e-39
	Mycobacterium tuberculosis MTV028.12c (gb/AL021426)[b]	3.0e-38
	Haemophilus influenzae (sp/P44973)[b]	5.4e-28
	Pseudomonas putida (sp/P25754)[b]	1.7e-26
	Coxiella burnetii (sp/P45650)[b]	2.7e-26
	Escherichia coli (sp/P25714)[b]	1.3e-17
	Helicobacter pylori (gi/2314626)[b]	4.0e-27
	Proteus mirabilis (Chi-Yung and Baumann, 1992)[b]	
	Stage III sporulation protein J from *Bacillus subtilis* (sp/Q01625)[b]	5.5e-19
jag	Jag protein (SpoIIIJ associated protein) from *Bacillus subtilis* (pir/I40438)[b]	1.0e-14
	*Mycobacterium leprae orf*193; L222-*orf*12 (gb/L39923)[b]	1.1e-38
	Mycobacterium tuberculosis MTV028.11c (gb/AL021426)[b]	4.0e-36
	*Leptospira interrogans orf*3 (gb/AB001721)	3.3e-12
gidB	● Glucose inhibited division protein B (GidB) from:	
	Mycobacterium leprae (gb/L39923)[b]	3.6e-58
	Mycobacterium tuberculosis MTV028.10c (gb/AL021426)[b]	1.0e-45
	Bacillus subtilis (sp/P25813)[b]	3.4e-12
	Pseudomonas putida (sp/P25757)[b]	5.8e-10
	Escherichia coli (sp/P17113)	2.5e-13
	Helicobacter pylory (gi/2314206)	2.4e-11
	Haemophilus influenzae (sp/P44728)	3.6e-09
soj	*Mycobacterium leprae orf*278 L-222 (gb/L39923)[b]	2.5e-101
	Mycobacterium tuberculosis MTV028.09c (gb/AL021426)[b]	2.0e-80
	● Soj protein from:	
	Bacillus subtilis (sp/P37522)[b]	3.1e-84
	Helicobacter pylori (gb/AE000620)	4.2e-52
	Mycoplasma pneumoniae (pir/S62837)	3.9e-09
	Pseudomonas putida (sp/P31856)[b]	3.7e-71
	Methanococcoccus jannaschii (pir/G64512)	2.1e-25
	Mycoplasma genitalium (sp/P47706)	1.0e-07
	● ParA chromosome partitioning protein from:	
	Caulobacter crescentus (gb/U87804)	1.8e-69
	Alcaligenes eutrophus ParA-like partitioning protein (pir/S60670)	1.4e-11
	● RepA protein from:	
	Rhizobium leguminosarum (gb/ X89447)	3.6e-14
	Agrobacterium rhizogenes (sp/P05682)	2.7e-14
	Agrobacterium tumefaciens (pir/A32812)	5.4e-14
	Paracoccus versutus (gb/U60522)	1.1e-11
	Rhizobium etli (gb/U80928)	1.1e-11
spo0J	*Mycobacterium leprae orf*333 (sp/L39923)[b]	1.2e-94
	Mycobacterium tuberculosis MTV028.08c (gb/AL021426)[b]	7.0e-70
	Spo0J stage 0 sporulation protein from *Bacillus subtilis* (sp/P26497)[b]	2.0e-65
	*Coxiella burnetii orf*288 (gb/Y10436)[b]	3.0e-51
	Streptococcus pneumoniae SPSpoJ (gb/AF000658)	5.2e-47
	Pseudomonas putida 32.4 kDa protein (sp/P31857)[b]	3.6e-43
	Bacillus subtilis 32.8 kDa protein (*orf*283) (sp/P37524)	1.2e-40
	Par B chromosome partitioning protein from *Caulobacter crescentus* (gb/U87804)	2.2e-50
	● Plasmid replication–partition related protein from:	
	Helicobacter pylori (gb/AE000620)	4.6e-37
	Coxiella burnetii (pir/S58487)	4.0e-34
	Rep B protein from *Paracoccus versutus* (gb/U60522)	0.00051
orf205	L-222 *orf*1 from *Mycobacterium leprae* (sp/L39923)[b]	5.0e-55
	Mycobacterium tuberculosis MTV028.07c (gb/AL021426)[b]	1.0e-33

[a]The *p*-value, calculated with the BLAST program, represents the probability that two proteins are related. The lower the *p* value, the greater the degree of similarity.
[b]The location of the individual gene in the bacteria L chromosome, is the same as the location of this homolog in the *S. coelicolor* chromosome.
[c]Found only by using the TBLASTN program, but not annotated, at present, as a complete gene in the databases.

a defect in chromosomal partitioning and result in anucleated cells (Ireton et al., 1994). In the second group, similarity was found to the ParB and RepB family of proteins which, together with ParA and RepA, are implicated in plasmid and bacterial chromosomal partitioning (Koonin, 1993).

(7) *orf*205 is located upstream of *orf*344 and is predicted to code for a 205 aa protein of molecular mass 22.5 kDa. Its ATG start codon is separated by 339 nucleotides from the *orf*344 termination codon. An intergenic gap of 336 nucleotides connects that codon to the termination codon of *trxA*. *orf*205 resembles L222-*orf*1 of *M. leprae* and to an *orf* in *M. tuberculosis*, but no known function has been ascribed to the putative gene.

3.2. Other features in the trxA/B-rnpA region

A 14 bp perfect palindrome – GTTTCACGTGAAAC – occurs in four sites in the *trxA/B–rnpA* region, twice in the region separating the putative genes *gidB* and *soj*, once at the beginning of the *spoOJ* gene and once, as noted previously (Cohen et al., 1993), at the end of the *trxA* gene. In this respect, it is of interest that several 7 bp repeats occur downstream of the *B. subtilis* and *P. putida spoOJ* genes which have been proposed as potential binding sites for SpoOJ (Ireton et al., 1994).

3.3. Comparison of the chromosomal organization of genes in the trxA/B-oriC region in Streptomyces with other bacteria

In many bacteria, such as *P. putida, B. subtilis, S. coelicolor* and *M. leprae*, the region around *oriC* contains a cluster of highly conserved genes, *rnpA–rpmH–dnaA–dnaN–recF–gyrB*, several of which play a key role in DNA replication (Yoshikawa and Ogasawara, 1991; Ogasawara and Yoshikawa, 1992; Calcutt, 1994; Salazar et al., 1996). Much less is known about the nature and order of the genes located upstream of *rnpA* and extending to *trxA/B*. Fig. 1 compares the organization of these genes in *S. coelicolor* with that of other bacteria. In this report we show that in *S. coelicolor* four of the genes located in this region – *orf*431 (*spoIIIJ*-like), *orf*170 (*jag*-like), *orf*255 (*soj*-like) and *orf*344 (*spoOJ*-like) – encode putative proteins that are homologs of proteins in *B. subtilis, P. putida*, *M. leprae* and *M. tuberculosis* and in a few other bacteria. In *B. subtilis* all four genes are implicated in the regulation of sporulation, while *spoOJ* is also necessary for control of chromosomal partitioning during prespore formation. (Ireton et al., 1994; Sharpe and Errington, 1996). Possibly, the above *Streptomyces* genes, and related genes in other sporogenic bacteria, play a similar role to that in *B. subtilis* despite some fundamental differences in their sporulation. Alternatively, these genes may be involved in DNA partitioning or may act in cooperation with other *Streptomyces* genes known to be necessary for morphological development, namely *whi* and *bld* (for a review see Chater, 1993). In non-sporogenic bacteria, such as *M. leprae, M. tuberculosis, P. putida* and *C. burnetii, soj* and *spoOJ* are presumed, on the basis of their sequence relatedness to *parA/repA* and *parB/repB*, respectively, to be involved in plasmid and chromosomal partitioning. These findings support the view that the presence of spore-like genes in clusters in non-sporogenic bacteria may reflect an evolutionary process that took place with the development of spore-forming bacteria in which primordial genes, associated in this case with chromosomal segregation, have acquired new functions needed for spore formation (Ogasawara and Yoshikawa, 1992). Thus, spore formation in *B. subtilis*, which depends on accurate chromosomal segregation in order that the prespore develops from the mother cell and acquires a single copy of the chromosome, employs genes that are required in vegetative growth. In this context, it is noteworthy that each of the non-sporogenic, as well as the sporogenic bacteria, listed in Table 2 contain a gene coding for a 60 kDa inner membrane protein which, in *B. subtilis*, is designated *spoIIIJ* for the role it plays in spore maturation. Furthermore, another gene encoding a 9–10 kDa protein, with as yet no known function, is also conserved, with the possible exception of *M. leprae*, in all of the above bacteria.

A number of differences in gene organization in the *trxA/B–oriC* region among the bacteria are worth mentioning. Thus, the *jag* gene is present in this region only in the Gram-positive bacteria. Also, the *gidA* gene (presumed to be involved in cell division) and another *orf*, termed *thdF* (thiophen and furan oxidation) that codes for a putative 50 kDa GTPase, are found near to *oriC* in many bacteria but are conspicuously absent in the *S. coelicolor*, *M. leprae* and *M. tuberculosis oriC* regions. Genome analysis reveals that in *M. leprae* and *M. tuberculosis* the *thdF* gene maps in an entirely different region of the chromosome from that of *oriC*. Among the Gram-negative bacteria, *P. putida* shows a noticeable similarity in gene organization in the *trxA/B–rnpA* region to that of the Gram-positive bacteria, and contains *orf*s homologous to the *B. subtilis spoOJ*, *soj* and *gidB* genes. In contrast, *E. coli* and some other Gram-negative bacteria, including *Haemophilis influenzae, Helicobacteria pylori* and *Buchnera aphidicola*, lack all three of the above genes in this region of the chromosome. The fact that in *E. coli oriC* is linked to *gidA*, which is located about 40 kb from the *dnaA* region, suggests that *oriC* and the flanking *gidA* and *gidB* genes have been translocated by a fragment inversion (Ogasawara and Yoshikawa, 1992; Burland et al., 1993; Fsihi et al., 1996). A similar event also appears to have taken place in *Borrelia burgdorferi, Mycoplasma genitalium* and *H. influenzae* (see Fsihi et al., 1996).

Perhaps the most significant difference in gene organization in the chromosomal region upstream of *dnaA* is that only in the actinomycetes, *S. coelicolor*, *M. leprae* and *M. tuberculosis*, are the thioredoxin genes *trxA/B* linked to *dnaA* and *oriC*. It will be of interest to see whether this arrangement is common to all members of the high G+C actinomycetes and, if so, it might reveal important insights into the relationship between genome architecture and some unique redox, thioredoxin-based, metabolism recently discovered in this group of bacteria (Newton et al., 1993, 1995, 1996).

Despite the above differences, marked gene organization clearly occurs in the *trxA/B–rnpA* region of bacterial chromosomes. Although less pronounced than that in the *dnaA–oriC–gyrB* region, it implies the existence of strong evolutionary constraints that preserve this gene organization. Moreover, the close proximity of genes involved in DNA replication, cell division and morphological development points to a complexity of interactions among these genes in the cell cycle.

4. Conclusions

(1) The DNA region upstream of *oriC* in *S. coelicolor* contains several genes homologous to, and organized in a similar fashion to, the genes in the corresponding regions in Gram-positive and Gram-negative bacteria and which are most strongly conserved in *Mycobacteria* species and *B. subtilis*.
(2) Several of the *Streptomyces* genes are homologs of *B. subtilis* genes, *spoIIIJ, jag, soj* and *spoOJ*, and some other bacterial genes that function in morphological development and chromosomal and plasmid partitioning.
(3) In *S. coelicolor* (and some other *Streptomyces* species, unpublished results), *M. leprae* and *M. tuberculosis*, the *trxA/B* gene cluster is located near to *oriC*, whereas in other non-actinomycetes the *trx* genes are far removed from *oriC*.

Acknowledgement

We thank Einat Hazkany and Margarita Mislovati for technical assistance.

References

Altschul, S.F., Boguski, M.S., Gish, W., Wooton, J.C., 1994. Issues in searching molecular sequence databases. Nat. Genet. 6, 119–129.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Av-Gay, Y., 1994. Molecular genetic studies of the *Streptomyces* thioredoxin and cold shock systems. Ph.D. Thesis, Tel-Aviv University, Tel-Aviv, Israel.

Blattner, F.R. et al., 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277, 1453–1461.

Berdy, J., 1984. New ways to obtain antibiotics. Chin. J. Antibiot. 7, 272–290.

Burland, V., Plunkett, III, G., Daniels, D.L., Blattner, F.R., 1993. DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. Genomics 16, 551–561.

Calcutt, M.J., 1994. Gene organization in the *dna*A–*gyr*A region of the *Streptomyces coelicolor* chromosome. Gene 151, 23–28.

Calcutt, M.J., Schmidt, F.G., 1992. Conserved gene arrangement in the origin region of the *Streptomyces coelicolor* chromosome. J. Bacteriol. 174, 3220–3226.

Chater, K.F., 1993. Genetics of differentiation in *Streptomyces*. Annu. Rev. Microbiol. 47, 685–713.

Cohen, G., Yanko, M., Mislovati, M., Argaman, A., Schreiber, R., Av-Gay, Y., Aharonowitz, Y., 1993. Thioredoxin–thioredoxin reductase system of *Streptomyces clavuligerus*; sequences, expression, and organization of the genes. J. Bacteriol. 175, 5159–5167.

Chi-Yung, L., Baumann, P., 1992. Genetic analysis of an aphid endosymbiont DNA fragment homologous to the *rnpA–rpmH–dnaA–dnaN–gyrB* region of eubacteria. Gene 113, 175–181.

Errington, J., Abbleby, L., Daniel, R.A., Goodfellow, H., Partridge, S.R., Yudkin, M.D., Structure and function of the *spo*IIIJ gene of *Bacillus subtilis*: a vegetatively expressed gene that is essential for σ^G activity at an intermediate stage of sporulation. 1992. J. Gen. Microbiol. 138, 2609–2618.

Fsihi, H., De Rossi, E., Salazar, L., Cantoni, R., Labo, M., Riccardi, G., Takiff, H.E., Eiglmeier, K., Bergh, S., Cole, S.T., 1996. Gene arrangement and organization in a ~76 kb fragment encompassing the *ori*C region of the chromosome of *Mycobacterium leprae*. Microbiology 142, 3147–3161.

Hopwood, D.A., Chater, K.F., Bibb, M.J., 1994. Genetics of antibiotic production in *Streptomyces coelicolor* A3(2), a model streptomycete. In: Vining, L.C., Stuttard, C. (Eds.), Genetics and Biochemistry of Antibiotic Production. Butterworth-Heinemann, Philadelphia, PA, pp. 65–102.

Hopwood, D.A., Kieser, T., 1990. The *Streptomyces* genome. In: Drlica, K., Riley, M. (Eds.), The Bacterial Chromosome. American Society for Microbiology, Washington, DC, pp. 147–162.

Ireton, K., Gunther, IV, N.W., Grossman, A.D., 1994. *spo*OJ is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. J. Bacteriol. 176, 5320–5329.

Kieser, H.M., Kieser, T., Hopwood, D.A., 1992. A combined genetic and physical map of the chromosome of *Streptomyces coelicolor* A3(2). J. Bacteriol. 174, 5496–5507.

Koonin, E.V., 1993. A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motifs. J. Mol. Biol. 229, 1165–1174.

Kunst, F. et al., 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature 390, 249–256.

Miyadoh, S., 1993. Research on antibiotic screening in Japan over the last decade: a producing microorganism approach. Actinomycetologica 7, 100–106.

Newton, G.L., Arnold, K., Price, M.S., Sherrill, C., Delcardayre, S.B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R.C., Davis, C., 1996. Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. J. Bacteriol. 178, 1990–1995.

Newton, G.L., Bewley, C.A., Dwyer, T.J., Horn, R., Aharonowitz, Y., Cohen, G., Davies, J., Faulkner, D.J., Fahey, R.C., 1995. The structure of U17 isolated from *Streptomyces clavuligerus* and its properties as an antioxidant thiol. Eur. J. Biochem. 230, 821–825.

Newton, G.L., Fahey, R.C., Cohen, G., Aharonowitz, Y., 1993. Low-molecular-weight thiols in streptomycetes and their potential role as antioxidants. J. Bacteriol. 175, 2734–2742.

Ogasawara, N., Yoshikawa, H., 1992. Genes and their organization in the replication origin region of the bacterial chromosome. Mol. Microbiol. 6 (5), 629–634.

Philipp, W.J., Poulet, S., Eiglmeier, K., Pascopela, L., Balasubramanian, V., Heym, B., Bergh, S., Bloom, B.R., Jacobs Jr., W.R., Cole, S.T., 1996. An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. Proc. Natl. Acad. Sci. USA 93, 3132–3137.

Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, H., Kinashi, H., Hopwood, D.A., 1996. A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. Mol. Microbiol. 21 (1), 77–96.

Salazar, L., Fsihi, H., de Rossi, E., Riccardi, G., Rios, C., Cole, S.T., Takiff, H.E., Organization of the origins of replication of the chromosomes of *Mycobacterium smegmatis, Mycobacterium leprae* and *Mycobacterium* tuberculosis and isolation of a functional origin from *M. smegmatis*. 1996. Mol. Microbiol. 20 (2), 283–293.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sharpe, M.E., Errington, J., 1996. The *Bacillus subtilis soj–spo*OJ locus is required for a centromere-like function involved in prespore chromosome partitioning. Mol. Microbiol. 21 (3), 501–509.

Smith, M., Jesse, J., Landers, T., Jordan, J., 1990. High efficiency bacterial electroporation: 1×10^{10} *E. coli* transformants/μg. Focus 12 (2), 38–40.

Strohl, W.R., 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acids Res. 20 (5), 961–974.

Wieles, B., van Soolingen, D., Holmgren, A., Offringa, R., Ottenhoff, T., Thole, J., 1995. Unique gene organization of thioredoxin and thioredoxin reductase in *Mycobacterium leprae*. Mol. Microbiol. 16, 921–929.

Yoshikawa, H., Ogasawara, N., 1991. Structure and function of DnaA and the DnaA-box in eubacteria: evolutionary relationships of bacterial replication origins. Mol. Microbiol. 5, 2589–2597.