J. Mol. Microbiol. Biotechnol. (2001) 3(4): 585-600.

JMMB Research Article

Comparative Genomics and Evolution of Genes Encoding Bacterial (p)ppGpp Synthetases/Hydrolases (the Rel, RelA and SpoT Proteins)

Gerhard Mittenhuber*

Institut für Mikrobiologie und Molekularbiologie, Ernst-Moritz-Arndt-Universität, F.-L. Jahnstr. 15, D-17487 Greifswald, Germany

Abstract

In the gram-negative model organism *Escherichia coli*, the effector molecule of the stringent response, (p)ppGpp, is synthesized by two different enzymes, RelA and SpoT, whereas in the gram-positive model organism *Bacillus subtilis* only one enzyme named Rel is responsible for this activity. Rel and SpoT also possess (p)ppGpp hydrolase activity. BLAST searches were used to identify orthologous genes in databases. The construction and bootstrapping of phylogenetic trees allowed classification of these orthologs. Four groups could be distinguished: With the exception of *Neisseria* and *Bordetella* (β subdivision), the RelA and SpoT groups are exclusively found in the γ subdivision of proteobacteria. Two Rel groups representing the actinobacterial and the *Bacillus*/*Clostridium* group were also identified. The SpoT proteins are related to the gram positive Rel proteins. RelA proteins carry substitutions in the HD domain (Aravind and Koonin, 1998, TIBS 23: 469-472) responsible for ppGpp degradation. A theory for the evolution of the specialized, paralogous *relA* and *spoT* genes is presented: After gene duplication of an ancestral *rel*-like gene, the *spoT* and *relA* genes evolved from the duplicated genes. The distribution pattern of the paralogous RelA and SpoT proteins supports a new model of linear bacterial evolution (Gupta, 2000, FEMS Microbiol. Rev. 24: 367-402). This model postulates that the γ subdivision of proteobacteria represents the most recently evolved bacterial lineage. However, two paralogous, closely related genes of *Porphyromonas gingivalis* (*Cytophaga-Flavobacterium-Bacteroides* phylum) encoding proteins with functions probably identical to the RelA and SpoT proteins do not fit in this model. Completely sequenced genomes of several obligately parasitic organisms (*Treponema pallidum*, *Chlamydia* species, *Rickettsia prowazekii*) and the obligate aphid symbiont *Buchnera* sp. APS as well as archaea do not contain *rel*-like genes but they are present in the *Arabidopsis* genome. In crosslinking experiments using different analogs of ppGpp as crosslinking reagents and RNA polymerase preparations of *Escherichia coli*, binding of ppGpp to distinct regions at the C-terminus of the β subunit (the RpoB gene product) and/or at the N-terminus of the β' subunit (the RpoC gene product) was observed previously. RpoB and RpoC sequences of the species which do not possess a *rel* like gene do not exhibit specific insertions or deletions in the ppGpp binding regions.

Introduction

The stringent response is an adaptive global mechanism for control of gene expression in bacterial cells subjected to starvation for amino acids or carbon sources (for review, see Cashel, 1996). Initiallly discovered forty years ago in the gram-negative model organism *Escherichia coli* (Stent and Brenner, 1961), this response is characterized by profound changes in the transcriptome of starved cells (*e.g.* cessation of rRNA biosynthesis and activation of genes encoding amino acid biosynthesis genes, activation of the stationary sigma factor σ^S (Gentry *et al.*, 1993) and transcription at σ^S-dependent promoters (Kvint *et al.*, 2000). The stringent response is mediated through the synthesis of the effector molecule (p)ppGpp (Cashel and Gallant, 1969).

In *E. coli,* two different proteins are involved in (p)ppGpp synthesis: RelA is bound to ribosomes, senses the amount of uncharged tRNA's and synthesizes (p)ppGpp following amino acid limitation (Haseltine and Block, 1973). The SpoT is a cytosolic protein (Gentry and Cashel, 1995) and functions as (p)ppGpp synthetase after carbon (Hernandez and Bremer, 1991; Murray and Bremer, 1996) and fatty acid (Seyfzadeh *et al.*, 1993) starvation. SpoT is also a (p)ppGpp hydrolase (Hernandez and Bremer, 1991; Murray and Bremer, 1996). The *relA* and *spoT* genes of *E. coli* are related resulting in the hypothesis that these genes evolved by gene duplication (Metzger *et al.*, 1989). Residual (p)ppGpp synthesis in a *relA* mutant is abolished in a *relA*/*spoT* double mutant (Xiao *et al.*, 1991). *relA*/*spoT* double mutants show a complex phenotype (morphological alterations; loss in the ability to grow on amino acid-free minimal media) (Xiao *et al.*, 1991; Hernandez and Cashel, 1995). The *relA* and *spoT* genes are listed as paralogous in the COG (cluster of orthologous groups) database (Tatusov *et al.*, 1997; Tatusov *et al.*, 2000) as COG0317.

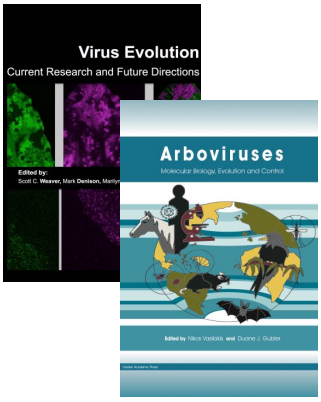
Several gram-positive bacterial species (*i.e. Bacillus subtilis* (Wendrich and Marahiel, 1997); *Corynebacterium glutamicum* (Wehmeier *et al.*, 1998); *Mycobacterium tuberculosis* (Avarbock *et al.*, 1999); *Staphylococcus*

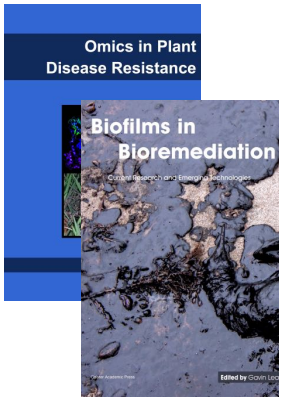
*For correspondence.
Email Gerhard.Mittenhuber@biologie.uni-greifswald.de;
gmittenhuber@paul.de; Tel. +49-3834-864218; Fax. +49-3834-864202.

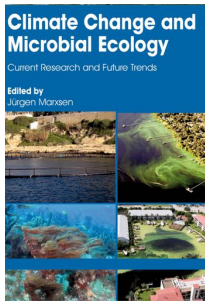
Further Reading

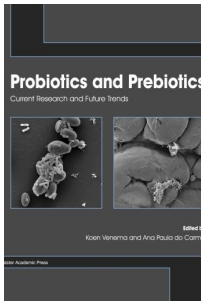
Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at **caister.com**

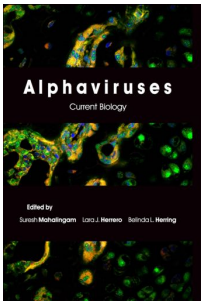
- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof
- ***Aspergillus* and *Penicillium* in the Post-genomic Era**
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2
- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorit, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins
- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr
- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles
- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  www.caister.com/climate
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  www.caister.com/biorem
- **Microalgae: Current Research and Applications**
  Edited by: MN Tsaloglou (2016)
  www.caister.com/microalgae
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma
- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusevol
- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo
- ***Shigella*: Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms
- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha
- **Thermophilic Microorganisms**
  Edited by: F Li (2015)
  www.caister.com/thermophile

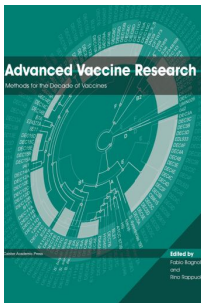


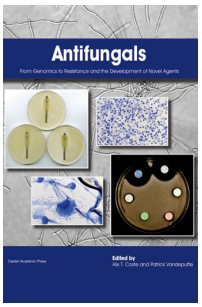












- **Flow Cytometry in Microbiology: Technology and Applications**
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow
- **Probiotics and Prebiotics: Current Research and Future Trends**
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics
- **Epigenetics: Current Research and Emerging Trends**
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015
- ***Corynebacterium glutamicum*: From Systems Biology to Biotechnological Applications**
  Edited by: A Burkovski (2015)
  www.caister.com/cory2
- **Advanced Vaccine Research Methods for the Decade of Vaccines**
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines
- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals
- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
  Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
  www.caister.com/bacteria-plant
- ***Aeromonas***
  Edited by: J Graf (2015)
  www.caister.com/aeromonas
- **Antibiotics: Current Innovations and Future Trends**
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics
- ***Leishmania*: Current Biology and Control**
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2
- ***Acanthamoeba*: Biology and Pathogenesis (2nd edition)**
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2
- **Microarrays: Current Technology, Innovations and Applications**
  Edited by: Z He (2014)
  www.caister.com/microarrays2
- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
  Edited by: D Marco (2014)
  www.caister.com/n2

Order from caister.com/order

Table 1. Alphabetical list of species analyzed for enzymes involved in (p)ppGpp metabolism and information about the phylogenetic status. Abbrevations of names are used in the phylogenetic trees and the Tables.

Organism	Phylogenetic position	Abbreviation
Actinobacillus actinomycetemcomitans	Proteobacteria; γ subdivision; Pasteurellaceae	Aac
Arabidopsis thaliana	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales; Brassicaceae	Ath
Aquifex aeolicus	Aquificales; Aquificaceae	Aae
Bacillus anthracis	Firmicutes; Bacillus/Clostridium group;Bacillus/Staphylococcus group; Bacillus cereus group	Ban
Bacillus halodurans	Firmicutes; Bacillus/Clostridium group;Bacillus/Staphylococcus group	Bha
Bacillus stearothermophilus	Firmicutes; Bacillus/Clostridium group;Bacillus/Staphylococcus group	Bst
Bacillus subtilis	Firmicutes; Bacillus/Clostridium group;Bacillus/Staphylococcus group	Bsu
Bordetella bronchiseptica	Proteobacteria; β subdivision; Alcaligenaceae	Bbr
Bordetella pertussis	Proteobacteria; β subdivision; Alcaligenaceae	Bpe
Borrelia burgdorferi	Spirochaetales; Spirochaetaceae	Bbu
Bradyrhizobium japonicum	Proteobacteria; α subdivision; Bradyrhizobium group	Bja
Buchnera sp. APS	Proteobacteria; γ subdivision	Bsp
Campylobacter jejuni	Proteobacteria; ε subdivision; Campylobacter group	Cje
Caulobacter crescentus	Proteobacteria; α subdivision; Caulobacter group	Ccr
Chlamydia trachomatis	Bacteria; Chlamydiales; Chlamydiaceae	Ctr
Chlamydia pneumoniae	Bacteria; Chlamydiales; Chlamydiaceae	Cpn
Chlorobium tepidum	Green sulfur bacteria	Cte
Clostridium acetobutylicum	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae	Cac
Clostridium difficile	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae	Cdi
Corynebacterium diphteriae	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae	Cdip
Corynebacterium glutamicum	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Corynebacteriaceae	Cgl
Dehalococcoides ethenogenes	Green non-sulfur bacteria; Dehalococcoides group	Det
Deinococcus radiodurans	Thermus/Deinococcus group; Deinococcales	Dra
Desulfovibrio vulgaris	Proteobacteria; δ subdivision	Dvu
Enterococcus faecalis	Firmicutes;Bacillus/Clostridium group; Enterococcaceae	Efa
Escherichia coli	Proteobacteria; γ subdivision; Enterobacteriaceae	Eco
Geobacter sulfurreducens	Proteobacteria; δ subdivision; Geobacteriaceae	Gsu
Haemophilus ducreyi	Proteobacteria; γ subdivision; Pasteurellaceae	Hdu
Haemophilus influenzae	Proteobacteria; γ subdivision; Pasteurellaceae	Hin
Helicobacter pylori	Proteobacteria; ε subdivision; Helicobacter group	Hpy
Legionella pneumophila	Proteobacteria; γ subdivision; Legionellaceae	Lpn
Mycobacterium avium	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium avium complex (MAC)	Mav
Mycobacterium bovis	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex	Mbo
Mycobacterium leprae	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae	Mle
Mycobacterium tuberculosis	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex	Mtu
Mycoplasma genitalium	Firmicutes; Bacillus/Clostridium group; Mollicutes	Mge
Mycoplasma pneumoniae	Firmicutes; Bacillus/Clostridium group; Mollicutes	Mpn
Myxococcus xanthus	Proteobacteria; δ subdivision; Myxobacteria; Myxococcales; Cystobacterinae	Mxa
Neisseria gonorrhoeae	Proteobacteria; β subdivision; Neisseriaceae	Ngo
Neisseria meningitidis	Proteobacteria; β subdivision; Neisseriaceae	Nme
Pasteurella multocida	Proteobacteria; γ subdivision; Pasteurellaceae	Pmu
Porphyromonas gingivalis	CFB group; Bacteriodaceae	Pgi
Pseudomonas aeruginosa	Proteobacteria; γ subdivision; Pseudomonas group	Pae
Pseudomonas putida	Proteobacteria; γ subdivision; Pseudomonas group	Ppu
Salmonella typhi	Proteobacteria; γ subdivision; Enterobacteriaceae	Sty
Shewanella putrefaciens	Proteobacteria; γ subdivision; Alteromonadaceae	Spu
Spiroplasma citri	Firmicutes; Bacillus/Clostridium group; Mollicutes; Spiroplasmataceae	Sci
Staphylococcus aureus	Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group	Sau
Streptococcus equi	Firmicutes; Bacillus/Clostridium group; Streptococcaceae	Seq
Streptococcus equisimilis	Firmicutes; Bacillus/Clostridium group; Streptococcaceae	Sequ
Streptococcus mutans	Firmicutes; Bacillus/Clostridium group; Streptococcaceae	Smu
Streptococcus pneumoniae	Firmicutes; Bacillus/Clostridium group; Streptococcaceae	Spn
Streptococcus pyogenes	Firmicutes; Bacillus/Clostridium group; Streptococcaceae	Spy
Streptomyces antibioticus	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae	San
Streptomyces coelicolor	Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycineae; Streptomycetaceae	Sco
Synechocystis sp. PCC6803	Cyanobacteria; Chroococcales	Ssp
Thermotoga maritima	Thermotogales	Tma
Thiobacillus ferrooxidans	Proteobacteria; γ subdivision	Tfe
Treponema pallidum	Spirochaetales; Spirochaetaceae	Tpa
Ureaplasma urealyticum	Firmicutes; Bacillus/Clostridium group; Mollicutes	Uur
Vibrio cholerae	Proteobacteria; γ subdivision; Vibrionaceae	Vch
Vibrio sp.	Proteobacteria; γ subdivision; Vibrionaceae	Vsp
Xylella fastidiosa	Proteobacteria; γ subdivision; Xanthomonas group	Xfa
Yersinia pestis	Proteobacteria; γ subdivision; Enterobacteriaceae	Ype

Table 2. RelA orthologs. The sequence of the RelA protein of *E. coli* was used as query sequence in the BLAST search. Individual proteins are sorted according to ascending E-values. For designations of species, see Table 1. The asterisk marks sequence fragments which were not included in the alignments because of their shortness. Designations of proteins: SpoT Pfr: SpoT of *Phlomobacter fragariae* (Foissac *et al.*, 2000); SpoT ?: SpoT sequence from an unidentified bacterium (Foissac *et al.*, 2000); CsrS Vsp: CsrS of *Vibrio* sp. (Östling *et al.*, 1995).

Designation	Length	Accession number	E-value
RelA Eco	744	J04039	0.0
RelA Vch	738	Q9S3S3	1.1e-263
RelA Vsp	744	P55133	6.7e-262
RelA Hin	743	P44644	3.5e-256
RelA Pae	747	AAG04323	3.6e-183
RelA Bha	728	BAB04961	1.1e-140
RelA Bsu	734	U86377	2.4e-138
RelA Xfa (XF1316)	718	AE003964	1.2e-136
RelA Nme	769	CAB85211	7.6e-135
Rel Ssp	760	P74007	8e-131
Rel Sau	736	O32419	1.9e-129
Rel Sco	847	X92520	7.5e-128
Rel Cgl	760	O87331	4.1e-127
Rel Mtu	790	Q50638	6.0e-126
Rel Mle	787	Q49640	3.3e-125
Rel Mxa	757	O52177	4.2e-125
Rel Seq	739	Q54089	8.8e-125
Rel San	841	O85709	2.3e-124
Rel Aae	696	O67012	1.2e-123
Rel Dra (DR1838)	787	Q9RTC7	2.6e-111
SpoT Vch (VC2710)	705	AAF95850	3e-108
SpoT Eco	702	P17580	1.7e-101
SpoT Pae	701	AAG08723	1.7e-101
SpoT Nme	725	CAB85138	2.3e-94
SpoT Hin	677	P43811	8.7e-94
Rel Tma (TM0729)	751	Q9WZI8	5.4e-93
Rel Bja	779	Q9RH69	7.2e-84
Rel Bbu	667	O51216	1.1e-80
SpoT Xfa (XF0352)	735	AE003887	8.7e-77
Rel Cje	731	AL139077	5.9e-74
Rel Sci	749	O34098	6.2e-69
RshA Sco (SC4H2.15)	725	O69970	8.6e-67
Rel HpyJ99	776	Q9ZL68	7.6e-66
F15I1.23 Ath	715	Q9SYH1	3.3e-59
RSH2 Ath	710	AAF37282	3.4e-52
SpoT Pfr *	388 (fragment)	AAG00076	3.0e-45
T2H3.9 Ath	615	O81418	2.1e-40
RSH1 Ath	883	AAF37281	2.4e-38
Rel Uur	718	AE002125	3.3e-37
Rel Mpn	733	P75386	7.1e-32
Rel Mge	720	P47520	2.0e-29
SpoT ? *	283 (fragment)	AAF73865	1.3e-27
CsrS Vsp *	119 (fragment)	Q56730	1e-20

aureus (Gentry *et al.*, 2000); *Streptococcus equisimilis* (Mechold *et al.*, 1996; Mechold and Malke, 1997); *Streptomyces coelicolor* (Chakraburtty *et al.*, 1996; Martinez-Costa *et al.*, 1996, 1998), *Streptomyces antibioticus* (Hoyt and Jones, 1999)) however, possess only one *relA*/*spoT* paralog which exhibits both (p)ppGpp synthetase and hydrolase activity. In these different species, ppGpp fulfills a variety of regulatory functions: In *B. subtilis*, mutant strains defective in the *relA*/*spoT* gene exhibit also pleiotropic phenotypes (amino acid auxotrophies (Wendrich and Marahiel (1997); defects in the expression of several genes controlled by the sporulation sigma factor σ^H and reduced sporulation efficiency (Eymann *et al.*, 2001)). A *rel* mutant of *Mycobacterium tuberculosis* displays slower aerobic growth rate and is less able to survive extended anaerobic incubation (Primm *et al.,* 2000). The *relA*/*spoT* gene is even essential in *S. aureus* (Gentry *et al.*, 2000). ppGpp plays a role in antibiotic and pigment production (Martinez-Costa *et al.*, 1996; Chakraburtty and Bibb, 1997; Hoyt and Jones, 1999; Hesketh *et al.*, 2001) and morphological differentiation (Chakraburtty and Bibb, 1997) in *Streptomyces* species.

In the genomes of the gram-negative species *Myxococcus xanthus* (Harris *et al.*, 1998) and *Porphyromonas gingivalis* (Sen *et al.*, 2000) both *relA* and *spoT* genes seem to be present: Harris *et al.* (1998) cloned and sequenced the *relA* gene of *M. xanthus* and demonstrated that its gene product is essential for fruiting body formation. The other paralogous gene is not yet cloned. Sen *et al.* (2000) performed a search of the "unfinished microbial genome" sequence database of *P. gingivalis* at TIGR, identified two homologues ORF's, and showed that ORF1 encodes RelA of *P. gingivalis*. The other ORF is not yet characterized experimentally.

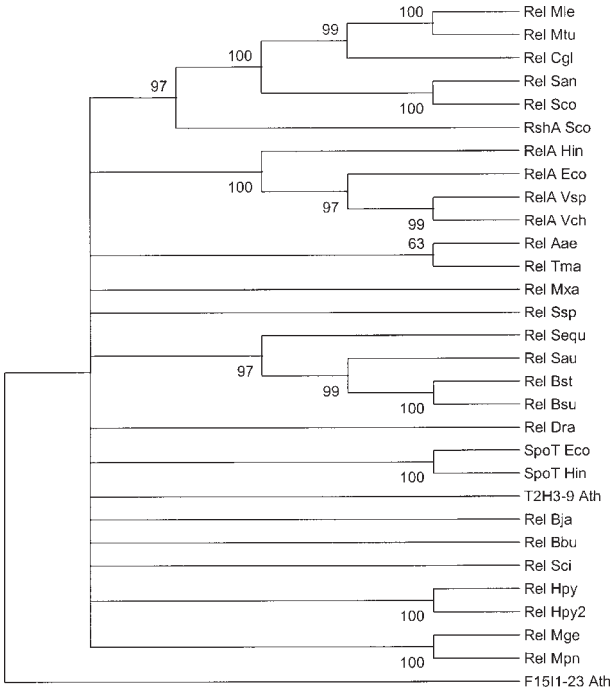


Figure 1. Rectangular cladogram of 30 RelA, SpoT and Rel proteins. Numbers on the branches are percentages of 1,000 bootstrap samples that support the branch; only values >50% are shown. Bootstrapping was performed by the neighbor-joining (NJ) method (Saitou and Nei, 1987) on the basis of the Poisson-corrected amino acid distance (d_{aa}). The sequence F15I1-23 of *Arabidopsis* was defined as outgroup.

RelA/SpoT homologs have also been discovered in the model plant *Arabidopsis thaliana* (van der Biezen *et al*., 2000). These authors used the yeast two-hybrid assay in order to identify proteins that interact with RPP5. This *Arabidopsis* protein which possesses a protein-protein interaction module called "NB-ARC" (van der Biezen and Jones, 1998; Aravind *et al*., 1999) is part of a signal transduction cascade involved in sensing the fungal pathogen *Peronospora parasitica* (Noel *et al*., 1999). In this two-hybrid assay, the RSH1 (RelA/SpoT homolog) protein was identified, a predicted plasma membrane-anchored cytoplasmic molecule with significant homology to RelA/SpoT. In a BLAST search two other unlinked *rel*-like genes were detected in the *Arabidopsis* genome, the *RSH2* and *RSH3* genes (van der Biezen *et al*., 2000). Functional complementation of appropriate *E. coli* and *S. coelicolor* mutants by the *RSH1* gene was also achieved in this study.

Very recently, the cloning and characterization of a gene encoding a *relA*/*spoT* homolog in *Bacillus stearothermophilus* was reported (Wendrich *et al*., 2000). The authors include a phylogenetic tree of 32 proteins homologous to RelA/SpoT of *B. subtilis*; van der Biezen *et al.* (2000) also present a phylogenetic tree. The RelA and SpoT proteins of *E. coli* and closely related bacteria form distinct clusters. A cluster of actinobacterial RelA/SpoT sequences and a cluster of sequences representing the *Bacillus*/*Clostridium* group can be observed. Wendrich *et al*. (2000) also propose a nomenclature for enzymes involved in (p)ppGpp metabolism: (p)ppGpp synthetase I (represented by RelA of *E. coli*); (p)ppGpp synthetase II (represented by SpoT of *E. coli*) and (p)ppGpp synthetase III (represented by synthetases of gram-positive origin; the name Rel is proposed for this type of enzyme). This nomenclature is adopted in this paper, although biochemical evidence has not yet been reported for several of the proteins named "Rel".

However, Wendrich *et al.* (2000) and van der Biezen *et al.* (2000) do not address the interesting question of evolutionary relationships between these three classes of enzymes to some detail. In this paper, these relationships are investigated using a larger number of RelA, SpoT and Rel sequences for construction of phylogenetic trees (for a list of species analyzed, see Table 1) and by performing statistical tests of tree topologies (*i.e.* bootstrapping (Felsenstein, 1985)). During these studies, we noted that genes encoding Rel-like proteins are absent in completely sequenced genomes of bacteria adopted to intracellular lifestyles. Therefore, we also investigated whether the target regions for binding of (p)ppGpp to the β and β' subunits of RNA polymerase (RNAP) in these species exhibit insertions or deletions (indels).

Results

Four General Groups of ppGpp Synthetase/Hydrolase Proteins

Our initial phylogenetic analysis of the dataset of Wendrich *et al.* (2000) showed a different tree topology than the one reported by the authors (data not shown). The reliability of the branching pattern of these tree data was then tested using bootstrapping (Felsenstein, 1985). Bootstrapping involves a randomization process in a multiple sequence alignment where all residues in a given column of the alignment are preserved in that column. Pseudosamples are created by random selection of sites from the data set with replacement. This is repeated a large number of times (*e.g.* 1000) and a phylogenetic tree is derived for each pseudosample. The percentage of pseudosamples which supports a unique clustering topology provides an estimate of the support for the pattern examined. Two short sequence fragments (*i.e.* CsrS from *Vibrio* sp. (GenBank accession number AAA85717; length: 119 amino acids; Östling *et al*., 1995) and a fragment derived from a gene of *Pseudomonas denitrificans* (GenBank acc. No. P29941; length 175 aa; Crouzet *et al*., 1991) which were included in the dataset of Wendrich *et al.* (2000) were excluded from the dataset used for creation of the alignment, because bootstrapping of a dataset containing short sequence fragments is not informative. A phylogenetic tree was constructed by the neighbor-joining (NJ) method (Saitou and Nei, 1987) on the basis of the Poisson-corrected amino acid distance (d_{aa}). Bootstrapping was performed with 1000 replicates (Figure 1). This tree is mainly multifurcating (or "condensed") because many of the interior branches exhibit a bootstrap value lower than 50%. Only branches separating representatives of closely related species as well as the RelA and SpoT proteins are supported by high bootstrap values.

Table 3. SpoT orthologs. The sequence of the SpoT protein of *E. coli* was used as query sequence in the BLAST search. Individual proteins are sorted according to ascending E-values. For designations of species, see Table 1. For explanation of the asterisk, see Figure 2. Designation of the protein: SpoT Pde: SpoT of *Pseudomonas denitrificans* (Crouzet *et al.*, 1991).

Designation	Length	Accession number	E-value
SpoT Eco	702	P17580	0.0
SpoT Vch (VC2710)	705	AAF95850	1.9e-280
SpoT Pae	701	AAG08723	6.6e-207
SpoT Hin	677	P43811	2.3e-204
SpoT Pfr *	388 (fragment)	AAG00076	1.1e-165
SpoT Xfa (XF0352)	735	AE003887	3.7e-165
SpoT Nme	725	CAB85138	1.3e-164
Rel Bha	728	BAB04961	5.8e-148
Rel Bsu	734	U86377	2.5e-147
Rel Seq	739	Q54089	4.2e-143
Rel Ssp	760	P74007	3.8e-142
Rel Sau	736	O32419	8e-140
Rel Mxa	757	O52177	1.9e-138
Rel Aae	696	O67012	1.6e-134
Rel San	841	O85709	1.7e-133
Rel Mtu	790	Q50638	2e-132
Rel Cgl	760	O87331	2.3e-131
Rel Mle	787	Q49640	1.6e-130
Rel Sco	847	X92520l	3.8e-127
Rel Dra (DR1838)	787	Q9RTC7	5.5e-126
Rel Tma (TM0729)	751	Q9WZI8	1.3e-125
SpoT ? *	283 (fragment)	AAF73865	9.1e-123
Rel Bja	779	Q9RH69	4.7e-116
RelA Pae	747	AAG04323	9.5e-112
RelA Nme	769	CAB85211	1e-105
RelA Hin	743	P44644	9.4e-103
Rel Bbu	667	O51216	5.1e-102
RelA Eco	744	J04039	5.2e-100
RelA Vsp	744	P55133	7.6e-97
RelA Xfa (XF1316)	718	AE003964	2.9e-94
Rel Cje	731	AL139077	3.4e-91
Rel Sci	749	O34098	1.6e-90
SC4H2.15 Sco	725	O69970	1e-88
RelA Vch (VC249-51)	738	Q9S3S3	2.5e-88
Rel Hpy99	776	Q9ZL68	1.5e-85
RSH1 Ath	883	AAF37281	1.3e-73
F15I1.23 Ath	715	Q9SYH1	9.6e-72
Rel Uur	718	AE002125	1.2e-67
RSH2 Ath	710	AAF37282	1.5e-67
T2H3.9 Ath	615	O81418	8.1e-62
Rel Mge	720	P47520	4.1e-56
Rel Mpn	733	P75386	6e-55
CsrS Vsp *	119 (fragment)	Q56730	1.1e-45
SpoT Pde *	175 (fragment)	P29941	2.4e-10

We assumed that a larger set of sequences used as input file for the alignment might result in a more informative tree. Therefore BLAST searches using the RelA (Table 2) and SpoT (Table 3) sequences of *E. coli* and the Rel sequence of *B. subtilis* (Table 4) as query sequences were performed. The Rel sequence of *B. subtilis* was also used to search the "unfinished microbial genomes" database at TIGR (see Table 5). A dataset comprising 80 related sequences was obtained in this way. An alignment of these sequences was performed and an unrooted tree was obtained (Figure 2). Several designations have been replaced by numbers in this tree. A clear separation of the RelA and SpoT families on one hand and of two Rel families of gram-positive origin (*Bacillus*/*Clostridium* group and a actinobacterial group) can be observed. Bootstrapping of this dataset also resulted in a largely condensed tree with internally supported branches (Figure 3). Interestingly, as also indicated by Wendrich *et al.* (2000) and as depicted in Figure 1, mycoplasmal Rel proteins form a clear outgroup.

It seemed likely that exclusion of certain sequences from the dataset used to produce the alignment might increase the resolution of the condensed part of the tree. In a first attempt, the outgroup sequences of *Arabidopsis* (RSH2 Ath, F15I1-23 Ath) and of the *Mycoplasma* (Rel Mpn, Rel Mge, Rel Uur) group as well as the *Arabidopsis* sequences T2H3-9 and RSH1 which are located in the condensed part of the tree were removed from the dataset (total: 73 sequences) subjected to the alignment. The resulting bootstrapped tree did not exhibit increased resolution in the condensed part of the tree (data not shown). In a second attempt, internal sequences of the condensed region of the tree in Figure 3 (*i.e.* Rel Dra, Rel Ssp, Rel Det, Rel Aae, Rel Tma, Rel Mxa, Rel Gsu, Rel

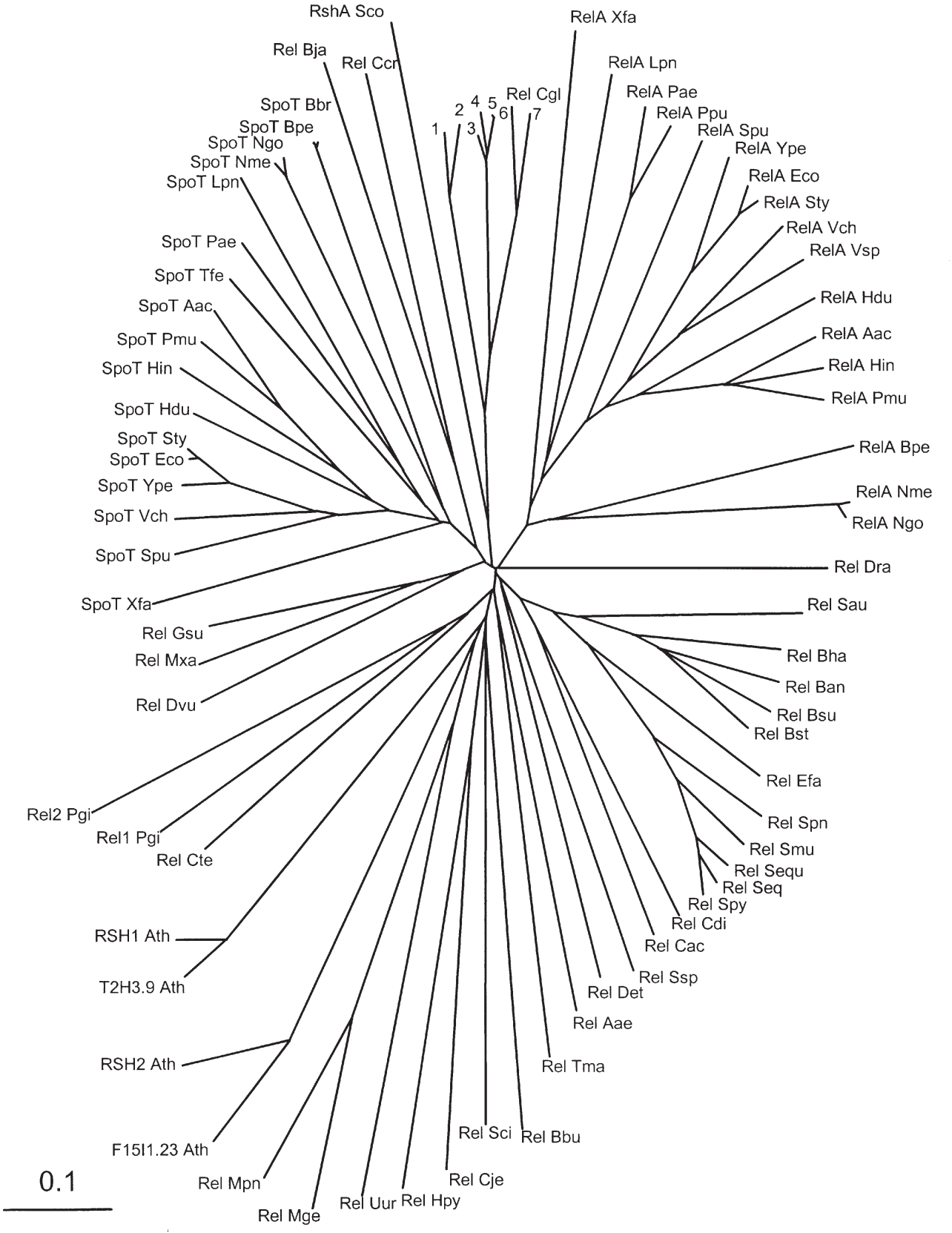


Figure 2. Radial phylogenetic tree of 80 RelA, SpoT and Rel proteins. Due to space limitations, the designation of individual Rel proteins was replaced by numbers. 1 refers to Rel Sco, 2 to Rel San, 3 to Rel Mav, 4 to Rel Mle, 5 to Rel Mtu, 6 to Rel Mbo and 7 to Rel Cdiph.

Table 4. Rel orthologs. The sequence of the Rel protein of *B. subtilis* was used as query sequence in the BLAST search. Individual proteins are sorted according to ascending E-values. For designations of species, see Table 1. For explanation of the asterisk, see Figure 2.

Designation	Length	Accession number	E-Value
Rel Bsu	734	U86377	0.0
Rel Bha	728	BAB04961	2e-278
Rel Sau	736	O32419	1.2e-241
Rel Seq	739	Q54089	9.9e-197
Rel Ssp	760	P74007	7.5e-167
Rel Mxa	757	O52177	7.5e-167
Rel Sco	847	X92520	2.6e-164
Rel San	841	O85709	3.1e-161
Rel Mtu	790	Q50638	2.9e-158
Rel Mle	787	Q49640	2.3e-156
Rel Cgl	760	O87331	2.4e-154
Rel Aae	760	O67012	9.5e-148
SpoT Eco	702	P17580	3e-140
SpoT Pae	701	AAG08723	3.1e-138
RelA Pae	747	AAG04323	3.5e-137
Rel Tma (TM0729)	751	Q9WZI8	9.3e-137
SpoT Vch (VC2710)	705	AAF95850	4.5e-135
RelA Eco	744	J04039	6.8e-134
RelA Vch	738	Q9S3S3	1.8e-133
RelA Vsp	744	P55133	1.1e-131
RelA Hin	743	P44644	5.7e-130
SpoT Nme	725	CAB85138	3.7e-129
SpoT Hin	677	P43811	2.2e-122
RelA Nme	769	CAB85211	1.3e-121
SpoT Xfa (XF0352)	735	AE003887	1.7e-120
Rel Bbu	667	O51216	1.2e-104
Rel Sci	749	O34098	1.4e-103
RelA Xfa (XF1316)	718	AE003964	6e-103
Rel Bja	779	Q9RH69	6.8e-102
Rel Cje	731	AL139077	4.8e-93
Rel Dra (DR1838)	787	Q9RTC7	2.7e-91
RshA Sco (SC4H2.15)	725	O69970	9.8e-88
Rel Hpy	776	Q9ZL68	1.3e-82
Rel Uur	718	AE002125	2.8e-78
T2H3.9 Ath	615	AAC28176	1.9e-71
Rel Mge	720	P47520	5.3e-69
Rel Mpn	733	P75386	2.3e-68
F15I1.23 Ath (identical to RSH3)	715	AAD25787	1.2e-67
RSH3	712	AAF37283	e-67
SpoT Pfr *	388 (fragment)	AAG00076	3.3e-65
RSH1 Ath	883	AAF37281	8e-64
RSH2 Ath	710	AAF37282	1.1e-62
SpoT ? *	283 (fragment)	AAF73865	3.9e-43
CsrS Vsp *	119 (fragment)	Q56730	1.3e-25
SpoT Pde *	175 (fragment)	P29941	1.8e-8

Dvu) were removed from the dataset (total: 65 sequences) used for production of the alignment. However, use this dataset in the construction of a bootstrapped tree also did not result in increased resolution of the consensed part on the tree (data not shown). In a third attempt, seven other outgroup sequences (*i.e.* Rel Cte, Rel1 Pgi, Rel2 Pgi, Rel Bbu, Rel Sci, Rel Cje and Rel Hpy) were removed from the dataset. The alignment of this dataset (58 sequences) finally resulted in improved resolution of the resulting bootstrapped tree (Figure 4).

A statistically insignificant branching (53%) separates the RelA proteins from the SpoT proteins and two groups of Rel proteins from gram-positive species. A statistically significant branching (99 %) separates the Rel proteins of the *Bacillus*/*Clostridium* group from the SpoT proteins and the actinobacterial Rel proteins. The SpoT proteins are separated from the actinobacterial Rel proteins by a statistically less significant branching (78%). An unrooted, radial tree originating from this alignment is presented in Figure 5. It should be noted that on the other hand, when another dataset comprising only the sequences representing the condensed part of the tree of Figure 2 was used as input file for an alignment, this procedure also did not increase the resolution in the condensed part of the resulting bootstrapped tree (data not shown).

Substitutions in the HD Domain of RelA Proteins

It was recognized by Aravind and Koonin (1998) that the HD domain defines a new superfamily of metal-dependent phosphohydrolases. Aravind and Koonin (1998) define five motifs within the HD domain. Three of them (motifs I, II, and V) contain highly conserved histidine or aspartate residues. Site-directed mutagenesis experiments of genes encoding cGMP-phosphodiesterases have shown that mutations of the histidine residue in the HD signature in motif II or of the aspartate residue in motif V resulted in the

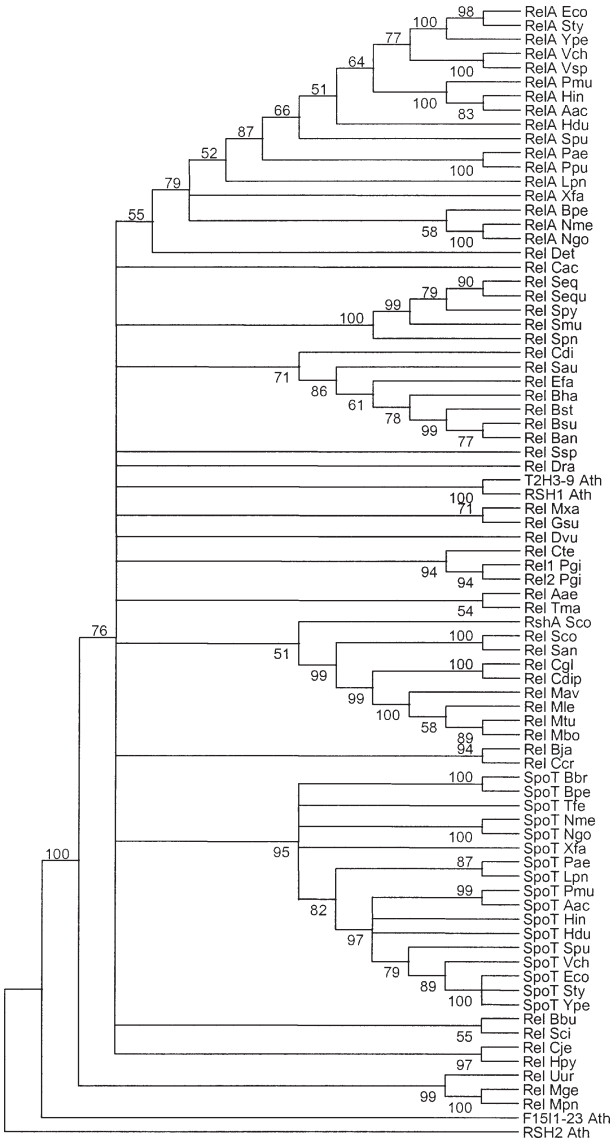


Figure 3. Rectangular cladogram of 80 RelA, SpoT and Rel proteins. Numbers on the branches indicated are bootstrap values as described in the legend to Figure 1.

most severe effect on the catalytic activity (Turko *et al.*, 1996). Aravind and Koonin (1998) showed that the RelA and SpoT proteins are also members of this superfamily. But the RelA protein of *E. coli* contains substitutions in the predicted catalytic residues of this domain. Aravind and Koonin (1998) suggest that the HD domain of RelA is inactivated but still retains its native structure. An inactive phosphohydrolase domain helps to explain why RelA proteins are not able to degrade ppGpp. On the other hand, the N-terminus of SpoT encompassing the HD domain is suffient for ppGpp hydrolase activity (Gentry and Cashel, 1996). An alignment of the N-termini of the SpoT, Rel and RelA proteins is presented in Supplementary Material (http://jmmb.net/supplementary). This alignment shows the region around conserved motifs I and II and the region around conserved motif V. Importantly, all the proteins identified in this study as putative RelA proteins carry substitutions in all conserved regions whereas Rel and SpoT proteins and the *Arabidopsis* paralogs possess a functional HD domain (http://jmmb.net/supplementary). Interestingly, in most substitutions of the HD signature sequence, insertion of a proline residue took place, an amino acid known to influence protein folding (*e.g.* Eyles and Gierasch, 2000). Nevertheless, the SpoT, Rel and RelA proteins share some overall similiarity in this region. This analysis also suggests that the protein designated RelA of *M. xanthus* by Harris *et al.* (1998) is in fact a Rel or SpoT protein.

Comparison of Putative ppGpp Binding Sites

During this analysis, we noted that genes encoding (p)ppGpp synthetases/hydrolases are absent in the completely sequenced genomes of the obligate parasites *Chlamydia pneumoniae* and *C. trachomatis* (Kalman *et al.*, 1999), *Rickettsia prowazekii* (Andersson *et al.*, 1998), and *Treponema pallidum* (Fraser *et al.*, 1998). Similarly, the genome of an intracellular symbiont of aphids, *Buchnera* sp. APS (Shigenobu *et al.*, 2000) does not contain a RelA or SpoT gene, although this extremely adapted bacterium is a close relative of *E. coli.* It seems likely that the *rel* genes were deleted during the adaptation to the intracellular environment and the establishment of specialized lifestyles in a process called “reductive evolution” (Andersson and Kurland, 1998).

Genetic studies (Glass *et al.*, 1986) using strains carrying amber mutations in *rpo* genes and biochemical investigations (Reddy *et al.*, 1995; Chatterji *et al.*, 1998; Toulokhonov *et al.*, 2001) using various ppGpp analogs as crosslinking agents identified RNAP of *E. coli* as target for ppGpp. Glass *et al.* (1986) concluded that ppGpp binds to the β subunit encoded by *rpoB* and also identified a common motif in purine-nucleotide binding proteins (GXXXXGK; la Cour *et al.*, 1985) in the amino acid sequence of RpoB at residues 880 to 886. By a combination of a variety of methods (SDS-PAGE analysis of the [^{32}P]N_3ppGpp (azido-ppGpp)-bound enzyme, tryptic digestion and Western blots), Chatterji *et al.* (1998) identified a 45 kDa fragment of the β subunit as binding site for ppGpp. This fragment is located at the C-terminus and consists of residues 802-1211, 1216 or 1223 (Chatterji *et al.*, 1998). Using the smaller ppGpp analog 6-thio-ppGpp, however, the β' subunit encoded by *rpoC* was very recently identified as target for ppGpp (Toulokhonov *et al.*, 2001). The binding site for ppGpp was determined at the N-terminus (between residues 29-102) of the β' subunit. The authors explain these conflicting results by the properties of the different crosslinking reagents and by the 3D model of RNAP (Zhang *et al.*, 1999). This model shows that the N-terminus of β' and the C-terminus of β are spatially close and constitute an intertwined interface. Toulokhonov *et al.*, (2001) postulate that binding of ppGpp to RNAP is allosteric, that the binding site is modular and is located close to the intersubunit interface of the N-terminal and C-terminal ends of β' and β.

In order to reveal the presence or absence of the different identified ppGpp binding regions in the β' and β sequences, alignments of RpoB (Table 6; http://jmmb.net/supplementary) and RpoC (Table 7; http://jmmb.net/

Table 5. Rel orthologs. The sequence of the Rel protein of *B. subtilis* was used as query sequence in the BLAST search of the "unfinished microbial genomes" database at TIGR. For designations of species, see Table 1. Individual proteins are sorted according to ascending E-values.

Designation	Length	Contig designation	E-value
Rel Ban	727	gnl|TIGR_1392|banth_1489	0.0
Rel Bst	707 (fragment)	gnl|UOKNOR_1422|bstear_Contig408	0.0
Rel Efa	737	gnl|TIGR|gef_10492	0.0
Rel Cdi	735	gnl|Sanger_1496|cdifficile_Contig876	0.0
Rel Spn	740	gnl|TIGR|S.pneumoniae_3836	0.0
Rel Smu	740	gnl|UOKNOR_1309|S.mutans_Contig98	0.0
Rel Spy	739	gnl|OUACGT_1315|Spyogenes_Contig1	0.0
Rel Cac	740	gnl|GTC|C.aceto_gnl	0.0
Rel Sequ	719 (fragment)	gnl|Sanger_1336|sequi_Contig474	0.0
Rel Gsu	716	gnl|TIGR_35554|gsulf_57	0.0
Rel Det	728	gnl|TIGR_61435|deth_1541	0.0
Rel Mbo	738	gnl|Sanger_1765|mbovis_Contig403	e-171
Rel Cdiph	759	gnl|Sanger_1717|cdiph_Contig2	e-166
Rel Dvu	717	gnl|TIGR_881|dvulg_159	e-166
Rel Mav	647	gnl|TIGR|M.avium_223	e-162
SpoT Tfe	759	gnl|TIGR|t_ferrooxidans_6160	e-149
SpoT Ype	707 (fragment)	gnl|Sanger_632|Y.pesits_Contig1008	e-148
SpoT Sty	709 (fragment)	gnl|Sanger_601|S.typhi_CT18	e-148
RelA Sty	744	gnl|Sanger_601|S.typhi_CT18	e-142
RelA Ype	744	gnl|Sanger_632|Y.pesits_Contig854	e-146
RelA Ppu	746	gnl|TIGR|pputida_10724	e-145
Rel Cte	731	gnl|TIGR|C.tepidum_3499	e-143
SpoT Spu	712 (fragment)	gnl|TIGR_24|sputre_6408	e-142
RelA Spu	735	gnl|TIGR_24|sputre_6426	e-142
RelA Lpn	734	gnl|CUCGC_446|lpneumo_MF.18.110397	e-141
RelA Pmu	739	gnl|CBCUMN_747|Pmultocida	e-140
SpoT Pmu	711 (fragment)	gnl|CBCUMN_747|Pmultocida	e-139
SpoT Aac	712 (fragment)	gnl|OUACGT_714|A.actin_Contig253	e-139
SpoT Bbr	759	gnl|Sanger_518|bbronchi_Contig2526	e-140
SpoT Bpe	759	gnl|Sanger_520|B.pertussis_Contig236	e-140
RelA Aac	743	gnl|OUACGT_714|A.actin_Contig233	e-136
SpoT Ngo	718	gnl|OUACGT_485|Ngon_Contig1	e-134
RelA Hdu	735	gnl|HTSC_730|ducreyi	e-133
RelA Ngo	737	gnl|OUACGT_485|Ngon_Contig1	e-129
RelA Bpe	737	gnl|Sanger_520|B.pertussis_Contig301	e-120
SpoT Hdu	569 (fragment)	gnl|HTSC_730|ducreyi	e-111
Rel1 Pgi	762	gnl|TIGR|P.gingivalis_GPG.con	e-114
Rel Ccr	743	gnl|TIGR|C.crescentus_12574	e-113
SpoT Lpn	530 (fragment)	gnl|CUCGC_446|lpneumo_MF.91.102397	e-94
Rel2 Pgi	746	gnl|TIGR|P.gingivalis_GPG.con	e-81

supplementary) amino acid sequences were performed. Our Supplementary Material (http://jmmb.net/supplementary) shows that the common motif in purine nucleotide-binding proteins described above is present also in RpoB sequences of obligately parasitic bacteria and in the symbiont *Buchnera* sp. APS. Large insertions around residues 1140 to 1260 of the RpoB alignment are present in the proteobacterial and chlamydial sequences and again around residues 1340 to 1420 in the proteobacterial, mycobacterial and mycoplasmal sequences. Most interestingly, the RpoB and RpoC sequences of the species which do not possess a *rel* like gene do not exhibit specific

Table 6. RpoB sequences compared in this study. For designations of species, see Table 1.

Designation	Length	Accession number
RpoB Eco	1342	RNEBC
RpoB Bsu	1193	P37870
RpoB Tpa	1178	O83269
RpoB Bsp	1342	BAB12761
RpoB Rpr	1374	O52271
RpoB Ctr	1252	O84317
RpoB Cpn	1252	BAA98291
RpoB Hin	1343	P43738
RpoB Sau	1182	P47768
RpoB Mpn	1391	P78013
RpoB Bbu	1155	AAB91501
RpoB Mtu	1172	CAB09390

Table 7. RpoC sequences compared in this study.

Designation	Length	Accession number
RpoC Eco	1407	RNECC
RpoC Bsu	1199	P37871
RpoC Sau	1057	P47770
RpoC Mtu	1316	P47769
RpoC Bsp	1407	BAB12760
RpoC Hin	1415	P43739
RpoC Rpr	1372	Q9ZE20
RpoC Ctr	1396	O84316
RpoC Cpn	1393	Q9Z999
RpoC Tpa	1416	O83270
RpoC Bbu	1377	O51349
RpoC Mpn	1290	P75271

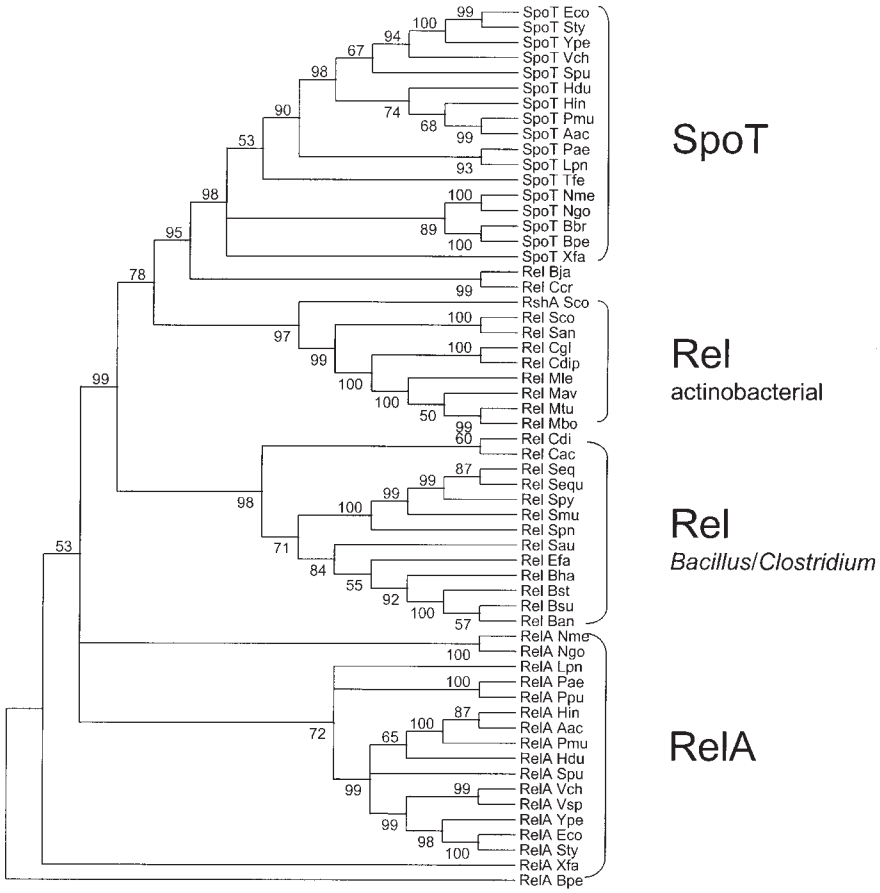


Figure 4. Rectangular cladogram of 58 RelA, SpoT and Rel proteins. Four different groups can be distinguished. Numbers on the branches indicated are bootstrap values as described in the legend to Figure 1.

alterations. This might indicate that the binding sites for ppGpp are still present in the β and β' subunits of these species.

Discussion

Absence of ppGpp in Archaea, Presence of ppGpp in Plants

Genes encoding (p)ppGpp synthetases are absent in all completely sequenced genomes of Archaea. Since eubacterial RNA polymerase is the target for the regulatory action of (p)ppGpp and since archaea possess a transcriptional apparatus similar to that of eukaryotes instead (Langer *et al.*, 1995), this difference between bacteria and archae might explain this absence. Starvation studies in halobacteria demonstrated stringent control of stable RNA biosynthesis but both growth rate control and stringent control are probably governed by mechanisms that operate in the absence of ppGpp (Cimmino *et al.*, 1993; Scoarughi *et al.*, 1995).

The finding that *Arabidopsis* possesses RelA/SpoT homologs (van der Biezen *et al.*, 2000) might indicate that plants have retained some aspects of the bacterial transcription apparatus. Indeed, in *Arabidopsis* proteins similar to bacterial sigma factors (σ^{70}) have been identified as products of genes specifically expressed in leafs and destined for chloroplasts (Isono *et al.*, 1997). Also, chloroplast genomes encode subunits of bacterial-type RNA polymerases (Sugiura *et al.*, 1998; Turmel *et al.*, 1999; Allison, 2000). Furthermore, ppGpp was detected in the alga *Chlamydomonas reinhardtii* (Heizmann and Howell, 1978).

Syntheny Considerations

Synteny studies (comparison of the relative positions of genes in the genome of different organisms) might be a source for auxiliary information to identify orthologous genes in genome sequencing projects (Huynen and Bork, 1998). In a systematic comparison using 256 operons of *E. coli* and 100 operons of *B. subtilis* as query sequences, Itoh *et al.* (1999) tried to identify operons showing a conserved order of orthologous genes in 11 complete genome sequences. These authors found that operon structures are very rarely conserved. The sequences of the *relA* and *spoT* operons of *E. coli* (but not the sequence of the corresponding *rel* gene of *B. subtilis*) were also used as query sequences. This analysis showed that no genome contains operons showing exactly the same organization as the *spoT* and *relA* operons of *E. coli*. However, due to the complexity of the *spoT* operon, the function of this

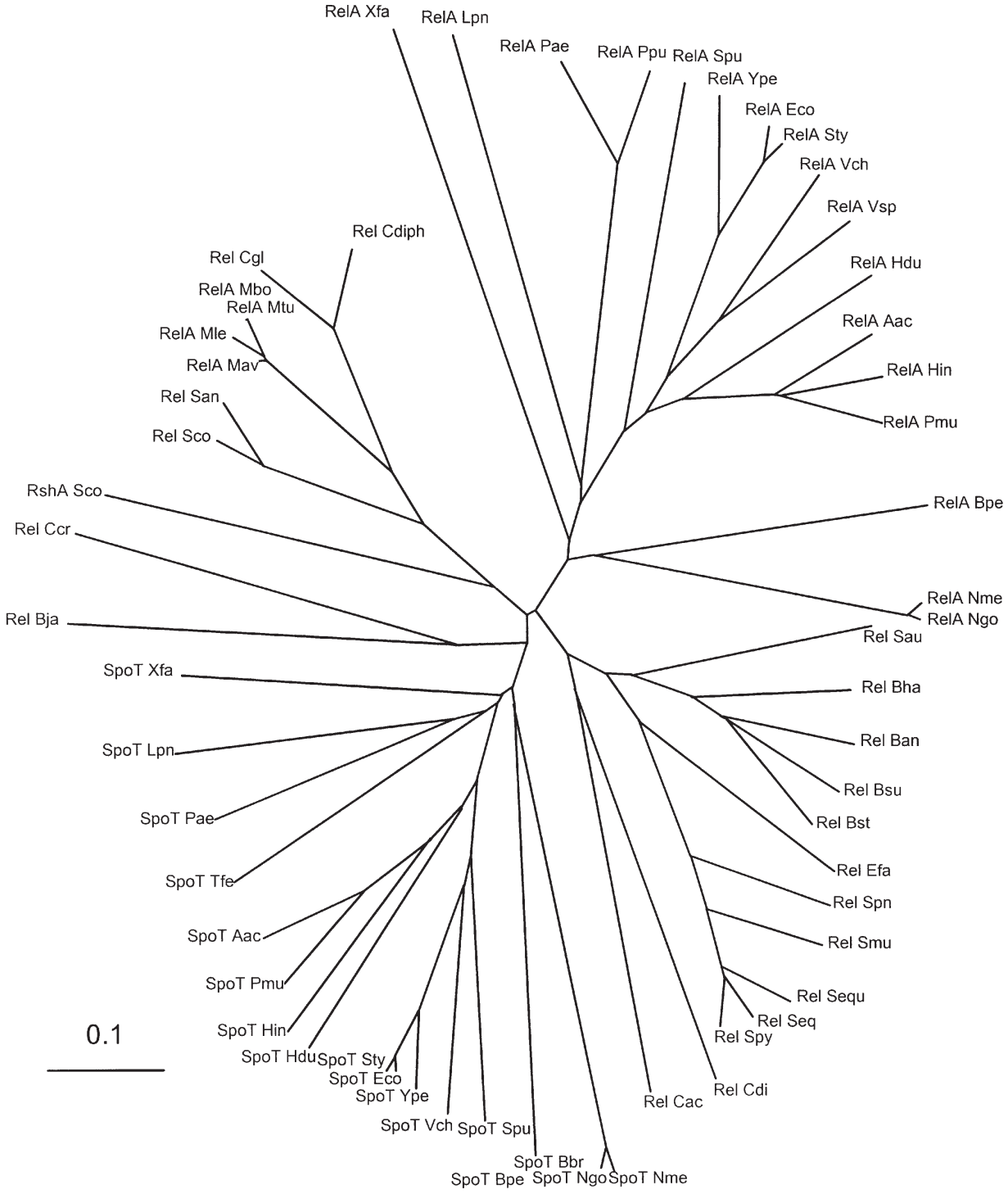


Figure 5. Radial phylogenetic tree of 58 RelA, SpoT and Rel proteins.

operon was misclassified as “DNA/RNA processing” in this study. For this reason, a short description of the *relA* and *spoT* operons of *E. coli* and the *rel* gene region of *B. subtilis* is given.

In *E. coli*, the *relA* gene is the first gene in an operon containing also the genes encoding the MazEF (ma-ze, hebrew for “what is it”) antitoxin/toxin module (Aizenman *et al*., 1996). Two promoters of the *mazEF* genes are negatively autoregulated by MazE and MazF and expression of the module is positively regulated by FIS (Marianovsky *et al*., 2001). In *B. subtilis* and other gram-positive bacteria, the putative *mazEF* locus is located upstream of the *sigB* operon (Mittenhuber, 1999), encoding the general stress response (Hecker and Völker, 1998) sigma factor σ^B and its regulatory proteins (Wise and Price, 1995). A third gene in the *relA* operon, *mazG*, encodes a

protein with unknown function. A BLAST search revealed that similar genes are present in many bacterial genomes (data not shown). They are designated as similiar to a hypothetical 26.1 kDa protein named YBL1 (accession number P33653) of *Streptomyces cacaoi* (Urabe and Ogawara, 1992).

The *spoT* gene of *Escherichia coli* is the third gene in an operon consisting of five genes. The order is *gmk-rpoZ-spoT-trmH-recG. Gmk* encodes guanylate kinase (Gentry *et al.*, 1993), *rpoZ* the Ω-subunit of RNAP (Gentry and Burgess, 1989), *trmH* a tRNA modifying enzyme (tRNA (Gm18) 2'-*O*-methyltransferase) (Persson *et al.,* 1997) and *recG* a DNA helicase (Lloyd and Sharples, 1991; Kalman *et al.*, 1992). Interestingly, the first three genes are linked to guanosine metabolism whereas the *trmH* and *recG* gene products play a role in RNA and DNA processing. Lloyd and Sharples (1991) identified a putative promoter near the end of *spoT*, indicating that *trmH* and *recG* might be transcribed as separate unit.

The organization of the *rel* loci of *B. subtilis*, *M. leprae*, *M. tuberculosis, S. equisimilis* and *S. coelicolor* is similar (Wendrich and Marahiel, 1997; Avarbock *et al.*, 1999). Upstream of the *rel* gene, the *apt* genes (encoding adenine phosphoryltransferase catalyzing the formation of adenosine monophosphate from adenine and phosphoribosyl pyrophosphate in a salvage reaction) and genes encoding components of the protein export machinery (*secDF*) are located in the same direction and downstream of *rel*, the *cypH* genes encoding cyclophilin (a peptidyl-prolyl *cis-trans* isomerase) are located in the opposite direction.

Proposal for the Evolution of the Rel, RelA and SpoT Proteins

The Mollicutes (*Mycoplasma* and relatives) which are normally associated with the *Bacillus*/*Clostridium* group of gram-positive bacteria form a distinct outgroup in the trees. These organisms undergo faster rates of evolution and quite regularly form outgroups in phylogenetic trees of orthologous protein sequences (Eisen, 1998).

With the exception of *Neisseria* and *Bordetella* species both belonging to the β-proteobacteria, two different genes encoding enzymes involved in (p)ppGpp synthesis and degradation, namely RelA and SpoT are only found in the ß and γ subdivision of proteobacteria. The SpoT genes of this group are related to genes encoding the actinobacterial group and the *Bacillus*/*Clostridium* group of Rel proteins (Figures 4 and 5; see also the E-values in Table 3). Since complete protein sequences were compared, the separation of the RelA proteins from the Rel and SpoT proteins is most probably due to the fact that the Rel and SpoT proteins possess ppGpp hydrolase activity, whereas the RelA proteins lack this activity. It is also interesting to note that the paralogous genes of individual species are not closely related to each other. Multiple, parallel gene duplications in individual species as origin of the *relA* and *spoT* genes can therefore be excluded.

Facilitated by the knowledge of the enzymatic activities of the Rel, RelA and SpoT proteins and based on some logical speculation, a model for evolution of the RelA and SpoT proteins within the β and γ subdivision of proteobacteria is proposed:

(1) The common ancestor of the β- and γ-proteobacteria possessed a *rel* gene of actinobacterial origin. The *spoT* gene evolved from this gene under adaptation of the gene product to carbon and fatty acid starvation.
(2) Following gene duplication, the additional copy of this ancestral gene is able to adopt to new functions. In this case, this adoption of new function was most probably inactivation of the HD domain, loss of (p)ppGpp hydrolyzing activity and adaptation to amino acid starvation. This copy evolved to the *relA* gene of β- and γ-proteobacteria.

The Special Case of *P. gingivalis*

P. gingivalis, however represents an interesting exception from the scenario described in the previous paragraph: This organism possesses a *relA* and a *spoT* gene (Sen *et al.*, 2000; http://jmmb.net/supplementary) which are however not related to the proteobacterial *relA* and *spoT* genes (Figure 3). Both paralogs which are named Rel1 Pgi and Rel2 Pgi in this paper are closely related to each other and to Rel of *Chlorobium tepidum* (Figure 3). In *P. gingivalis*, the evolution of the paralogous *rel1* (*spoT*-related) and *rel2* (*relA*-related) occured independently of the proteobacterial *relA* and *spoT* genes. At the moment, it is impossible to decide whether this paralogy represents an example of convergent evolution. Alternatively, lateral gene transfer of short, catalytically active domains of ppGpp synthetases/hydrolases might have been involved in the evolution of the *rel1* and *rel2* genes of *P. gingivalis*. In order to solve this question, more sequences encoding ppGpp synthetases/hydrolases from representatives of the CFB (*Cytophaga-Flavobacterium-Bacteroides*) phylum and green sulfur bacteria should be determined and compared (see also next paragraph). Alternatively, a careful phylogenetic analysis of the individual domains of the Rel, RelA and SpoT proteins (Aravind and Koonin, 1998), which is however beyond the scope of this paper might help to solve this issue. Such an analysis has been performed for the separate domains of the conserved HSP70 (DnaK) family: Striking differences in the relative rate of amino acid replacement in different rates were observed and were interpreted as evidence for functional divergence (Hughes, 1993).

ppGpp Synthetases/Hydrolases and Bacterial Evolution

A drastically different view of bacterial evolution has been recently postulated by Gupta (2000a, b). This hypothesis places the γ subdivision of proteobacteria at the end of a evolutionary linear scheme. According to this theory, the γ subdivision of proteobacteria constitutes the most recently evolved bacterial group. Therefore, specialized, paralogous *spoT* and *relA* genes might be a relatively new invention of bacterial evolution which might explain this limited distribution pattern among the β- and γ-proteobacteria. Similarly, the (maybe more efficient) vitamin B_6 biosynthesis pathway of *E. coli* is mainly restricted to the γ subdivision, whereas another widely conserved, biochemically uncharacterized pathway is presumably operating in other species (Mittenhuber, 2001). Other unique features of the γ-proteobacteria were recently recognized by Margolin (2000) in his study on prokaryotic cell division: Genes encoding the ZipA, FtsL and FtsN proteins are only found

in genomes of completely sequenced γ-proteobacteria. Very recently, Eisen (2000) noted that the currently available datasets of completely sequenced genomes are not representative of evolutionary diversity. It might be predicted that the availability of completely sequenced genomes of proteobacterial species belonging to subdivisions other than β and γ might help to identify the precursor of the proteobacterial *relA* and *spoT* genes.

ppGpp is Not Present in Obligately Intracellular Species

The species lacking a *rel*-like gene require living cells for their replication and growth and cannot be cultivated *in vitro*. The absence of *rel*-like genes in genomes of obligately parasitic organisms was also detected in a comparative genome analysis of *B. burgdorferi* and *T. pallidum* (Subramanian *et al*., 1999). *R. prowazekii* possesses a few short pieces of a pseudogene which show strong sequence similarity to the *relA*/*spoT* homologs (Andersson *et al*., 1998; Zomorodipour and Andersson, 1999), whereas no *rel*-like gene is present in the *C. trachomatis* genome (Stephens *et al*., 1998). Interestingly, some cultivable intracellular pathogens (*e.g. B. burgdorferi* and *Mycoplasma* species among others) possess *rel*-like genes. In most cases, cultivation of bacteria implies that the inoculum is able to form colonies. Investigations on vertical sections of *E. coli* colonies show that the colony is composed of different layers of bacteria, containing also nonviable bacteria (Shapiro, 1994). It is very likely that many cells in a developed colony are starved for nutrients and that the stringent response is switched on in these cells. It might be extremely interesting to investigate whether a functional connection between *in vitro* cultivability of bacterial species and the absence of genes encoding (p)ppGpp synthetases/hydrolases in obligate parasites can be established experimentally.

Experimental Procedures

Using the deduced protein sequences of the *relA* and *spoT* genes of *E. coli* and of the *rel* gene of *B. subtilis* as query sequences, BLAST searches (Altschul *et al*., 1997) of the non-redundant protein database nrdb95 (Holm and Sander, 1998) were performed at http://dove.embl-heidelberg.de/Blast2/ using the blastp program. BLAST searches of unfinished microbial genomes using the deduced protein sequence of the *rel* gene of *B. subtilis* as query were performed at http://www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html using the program tblastn. Raw DNA sequences were translated using the program "Translate tool" at http://www.expasy.ch/tools/dna.html. The COG database can be assessed at http://www.ncbi.nlm.nih.gov/COG. Alignments were generated using CLUSTALW (Thompson *et al*., 1994) at http://www.ebi.ac.uk/clustalw/. Radial phylogenetic trees were constructed using the data from the alignments with the help of the program TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html; Page, 1996) and edited using the program Metafile Companion (http://www.companionsoftware.com). Bootstrapping of datasets was performed by the neigbor-joining (NJ) method on the basis of the Poisson-corrected amino acid distance (d_{aa}) (Saitou and Nei, 1987; for a discussion of different statistical methods and their applications see Hughes, 1999; Nei and Kumar, 2000) using MEGA 2.0 (http://www.megasoftware.net; Kumar *et al*., 2001).

Acknowledgements

This work was performed in the lab of Michael Hecker, whose support is gratefully acknowledged. I thank Michael Hecker, Ulrich Zuber and an anonymous referee for comments on the manuscript. Sequences of unfinished microbial genomes were obtained at the website of "The Institute of Genomic Research" (TIGR) at http://www.tigr.org. Sequencing of unfinished genomes is in progress at several institutions funded by a variety of organizations: This listing includes "name of organism (institution(s); source of funding)" and can be also viewed at http://www.tigr.org/tdb/mdb/mdbinprogress.html. *A. actinomycetemcomitans* (University of Oklahoma; National Institute of Dental Research (NIDR)); *B. anthracis* (TIGR; Office of Naval Research (ONR)/Department of Energy (DOE)/National Institute of Allergy and Infectious Diseases (NIAD)); *B. halodurans* (Japan Marine Science and Technology Center); *B. stearothermophilus* (University of Oklahoma; National Science Foundation (NSF); *B. bronchiseptica*, *B. pertussis*, *C. difficile*, *N. meningitidis*, *Y. pestis* (Sanger Centre; Beowulf Genomics); *C. crescentus*, *C. tepidum*, *D. ethenogenes*, *D. vulgaris*, *S. putrefaciens*, *T. ferrooxidans* (TIGR/DOE); *C. acetobutylicum* (Genome Therapeutics; DOE); *C. diphteriae* (Sanger Centre/World Health Organization (W.H.O.)/Public Health Laboratory; Beowulf Genomics); *G. sulfurreducens* (TIGR/University of Massachusetts, Amherst/DOE); *H. ducreyi* (NIAID); *L. pneumophila* (Columbia Genome Center; NIAID); *M. avium*, *V. cholerae* (TIGR/NIAID); *M. bovis* (Sanger Centre/Institut Pasteur/VLA Weybridge; MAFF/Beowulf Genomics); *M. leprae* (Sanger Centre; The New York Community Trust); *N. gonorrhoeae* (University of Oklahoma; NIAID); *P. multocida* (University of Minnesota; US Department of Agriculture - National Reserach Initiative (USDA-NRI)/Minnesota Turkey Growers Association); *P. gingivalis* (TIGR/Forsyth Dental Center; NIDR); *P. aeruginosa* (University of Washington/PathoGenesis; Cystic Fibrosis Foundation/PathoGenesis); *P. putida* (TIGR/German Consortium; DOE/Bundesministerium für Bildung und Forschung (BMBF); *S. typhi*, (Sanger Centre; Wellcome Trust); *S. aureus* (TIGR; NIAID/Merck Genome Research Institute (MGRI)); *S. pneumoniae* (TIGR; TIGR/NIAID/MGRI); *S. coelicolor* (Sanger Centre/John Innes Centre; Biotechnology and Biological Sciences Research Council (BBSRC)/Beowulf Genomics).

References

Aizenman, E., Engelberg-Kulka, H., and Glaser, G. 1996. An *Escherichia coli* chromosomal "addiction module" regulated by guanosine [corrected] 3',5'-bispyrophosphate: a model for programmed bacterial cell death. Proc. Natl. Acad. Sci. USA 93: 6059-6063.

Allison, L.A. 2000. The role of sigma factors in plastid transcription. Biochimie 82: 537-548.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped

BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

Andersson, S.G., and Kurland, C.G. 1998. Reductive evolution of resident genomes. Trends Microbiol. 6: 263-268.

Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H., and Kurland, C.G. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 296: 133-140.

Aravind, L., and Koonin, E.V. 1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem. Sci. 23: 469-472.

Aravind, L., Dixit, V.M., and Koonin, E.V. 1999. The domains of death: evolution of the apoptosis machinery. Trends Biochem. Sci. 24: 47-53.

Avarbock, D., Salem, J., Li, L.S., Wang, Z.M., and Rubin, H. 1999. Cloning and characterization of a bifunctional *relA*/*spoT* homologue from Mycobacterium tuberculosis. Gene 233: 261-269.

Cashel, M., and Gallant, J. 1969. Two compounds implicated in the function of the RC gene of *Escherichia coli*. Nature 221: 838-841.

Cashel, M., Gentry D.R., Hernandez V.J. , and Vinella, D. 1996. The stringent response. In: *Escherichia coli* and *Salmonella:* Cellular and Molecular Biology. F.C. Neidhardt (Editor-in- Chief). ASM Press, Washington D.C. p. 1458-1496.

Chakraburtty, R., White, J., Takano, E., and Bibb, M. 1996. Cloning, characterization and disruption of a (p)ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3. Mol. Microbiol. 19: 357-368.

Chakraburtty, R., and Bibb, M. 1997. The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. J. Bacteriol. 179: 5854-5861.

Chatterji, D., Fujita, N., and Ishihama, A. 1998. The mediator for stringent control, ppGpp, binds to the β-subunit of *Escherichia coli* RNA polymerase. Genes to Cells 3: 279-287.

Cimmino, C., Scoarughi, G.L., and Donini, P. 1993. Stringency and relaxation among the halobacteria. J. Bacteriol. 175: 6659-6662.

Crouzet, J., Levy-Schil, S., Cameron, B., Cauchois, L., Rigault, S., Rouyez, M.C., Blanche, F., Debussche, L. and Thibaut, D. 1991. Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair *Pseudomonas denitrificans* DNA fragment containing five *cob* genes and identification of structural genes encoding Cob(I)alamin adenosyltransferase, cobyric acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. J. Bacteriol. 173: 6074-6087.

Eisen, J.A. 1998. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. Genome Res. 8: 163-167.

Eisen, J.A. 2000. Assessing evolutionary relationships among microbes from whole-genome analysis. Curr. Opin. Microbiol. 3: 475-480.

Eyles, S.J., and Gierasch, L.M. 2000. Multiple roles of prolyl residues in structure and folding. J. Mol. Biol. 301: 737-747.

Eymann, C., Mittenhuber, G., and Hecker, M. 2001. The stringent response, σ^H-dependent gene expression and sporulation in *Bacillus subtilis*. Mol. Gen. Genet. 264: 913-923.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Foissac, X., Danet, J.L., Zreik, L., Gandar, J., Nourrisseau, J.G., Bove, J.M., and Garnier, M. 2000. Cloning of the *spoT* gene of "Candidatus *Phlomobacter fragariae*" and development of a PCR-restriction fragment length polymorphism assay for detection of the bacterium in insects. Appl. Environ. Microbiol. 66: 3474-3480.

Fraser, C.M., Norris, S.J., Weinstock, G.M., White O., Sutton, G.G., Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A., Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J., Khalak, H., Richardson, D., Howell, J.K., Chidambaram, M., Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M.D., Fujii, C., Garland, S., Hatch, B., Horst, K., Roberts, K., Watthey, L., Weidman, J., Smith, H.O., and Venter, J.C. 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. Science 281: 375-388.

Gentry, D., Bengra, C., Ikehara, K., and Cashel, M. 1993. Guanylate kinase of *Escherichia coli* K-12. J. Biol. Chem. 268: 14316-14321.

Gentry, D., Li, T., Rosenberg, M., and McDevitt, D. 2000. The *rel* gene is essential for *in vitro* growth of *Staphylococcus aureus*. J. Bacteriol. 182: 4995-4997.

Gentry, D.R., and Burgess, R.R. 1989. *rpoZ*, encoding the omega subunit of *Escherichia coli* RNA polymerase, is in the same operon as *spoT*. J. Bacteriol. 171: 1271-1277.

Gentry, D.R., Hernandez, V.J., Nguyen, L.H., Jensen, D.B., and Cashel, M. 1993. Synthesis of the stationary-phase sigma factor σ^S is positively regulated by ppGpp. J. Bacteriol. 175: 7982-7989.

Gentry, D.R., and Cashel, M. 1995. Cellular localization of the *Escherichia coli* SpoT protein. J. Bacteriol. 177: 3890-3893.

Gentry, D.R., and Cashel, M. 1996. Mutational analysis of the *Escherichia coli spoT* gene identifies distinct but overlapping regions involved in ppGpp synthesis and degradation. Mol. Microbiol. 19: 1373-1384.

Glass, R.E., Jones, S.T., and Ishihama, A. 1986. Genetic studies on the β subunit of *Escherichia coli* RNA polymerase VII. RNA polymerase *is* a target for ppGpp. Mol. Gen. Genet. 203: 265-268.

Gupta, R.S. 2000a. The natural evolutionary relationships among prokaryotes. Crit. Rev. Microbiol. 26: 111-131.

Gupta, R.S. 2000b. The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. FEMS Microbiol. Rev. 24: 367-402.

Harris, B.Z., Kaiser, D., and Singer, M. 1998. The guanosine nucleotide (p)ppGpp initiates development and A-factor production in *Myxococcus xanthus*. Genes Dev. 12: 1022-1035.

Haseltine, W.A., and Block, R. 1973. Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. Proc. Natl. Acad. Sci. USA 70: 1564-1568.

Hecker, M., and Völker, U. 1998. Non-specific, general and multiple stress resistance of growth-restricted *Bacillus subtilis* cells by the expression of the σ^B regulon. Mol. Microbiol. 29: 1129-1136.

Heizmann, P., and Howell, S.H. 1978. Synthesis of ppGpp and chloroplast RNA in *Chlamydomonas reinhardi*. Biochem. Biophys. Acta 517: 115-124.

Hesketh, A., Sun, J., and Bibb, M. 2001. Induction of ppGpp synthesis in *Streptomyces coelicolor* A3(2) grown under conditions of nutritional sufficiency elicits *actII-orf4* transcription and actinorhodin biosynthesis. Mol. Microbiol. 39: 136-144.

Hernandez, V.J., and Bremer, H. 1991. *Escherichia coli* ppGpp synthetase II activity requires *spoT*. J. Biol. Chem. 266: 5991-5999.

Hernandez, V.J., and Cashel, M. 1995. Changes in conserved region 3 of *Escherichia coli* σ^{70} mediate ppGpp-dependent functions *in vivo*. J. Mol. Biol. 252: 536-549.

Holm, L., and Sander, C. 1998. Removing near-neighbour redundancy from large protein sequence collections. Bioinformatics 14: 423-429.

Hoyt, S., and Jones, G.H. 1999. RelA is required for actinomycin production in *Streptomyces antibioticus*. J. Bacteriol. 181: 3824-3829.

Hughes, A.L.. 1993. Nonlinear relationships among evolutionary rates identify regions of functional divergence in heat-shock protein 70 genes. Mol. Biol. Evol. 10: 243-255.

Hughes, A.L. 1999. Adaptive evolution of genes and genomes. Oxford University Press, Oxford, UK.

Huynen, M., and Bork, P. 1998. Measuring genome evolution. Proc. Natl. Acad. Sci. USA 95: 5849-5856.

Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A., and Kobayashi, H. 1997. Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of σ^{70} factors of bacterial RNA polymerase in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 94: 14948-14953.

Itoh, T., Takemoto, K., Mori, H. and Gojobori, T. 1999. Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes. Mol. Biol. Evol. 16: 332-346.

Kalman, M., Murphy, H., and Cashel, M. 1992. The nucleotide sequence of *recG*, the distal *spo* operon gene in *Escherichia coli* K-12. Gene 110: 95-99.

Kalman, S., Mitchell, W., Maranthe, R., Lammel, C., Fan, J., Hyman, R.W., Olinger, L., Grimwood, J., Davis, R.W., and Stephens, R.S. 1999. Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. Nature Genet. 21: 385-389.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. 2001. MEGA 2: Molecular evolutionary genetics analysis software. Bioinformatics (submitted).

Kvint, K., Farewell, A., and Nyström, T. 2000. RpoS-dependent promoters require guanosine tetraphosphate even in the presence of high levels of σ^S. J. Biol. Chem. 275: 14795-14798.

la Cour, T.F., Nyborg, J., Thirup, S., and Clark, B.F. 1985. Structural details of the binding of guanosine diphosphate to elongation factor Tu from *E. coli* as studied by X-ray crystallography. EMBO J. 4: 2385-2388.

Langer, D., Hain, J., Thuriaux, P., and Zillig, W. 1995. Transcription in archaea: similarity to that in eucarya. Proc. Natl. Acad. Sci. USA 92: 5768-5772.

Lloyd, R.G., and Sharples. 1991. Molecular organization and nucleotide sequence of the *recG* locus of *Escherichia coli* K-12. J. Bacteriol. 173: 6837-6843.

Margolin, W. 2000. Themes and variations in prokaryotic cell division. FEMS Microbiol. Rev. 24: 531-548.

Marianovsky, I., Aizenman, E., Engelberg-Kulka, H., and Glaser, G. 2001. The regulation of the *Escherichia coli mazEF* promoter involves an unusual alternating palindrome. J. Biol. Chem. 276: 5975-5984.

Martinez-Costa, O.H., Arias, P., Romero, N.M., Parro, V., Mellado, R.P., and Malpartida, F. 1996. A *relA*/*spoT* homologous gene from *Streptomyces coelicolor* A3(2) controls antibiotic biosynthesis genes. J. Biol. Chem. 271: 10627-10634.

Martinez-Costa, O.H., Fernandez-Moreno, M.A., and Malpartida, F. 1998. The *relA*/*spoT* homologous gene in *Streptomyces coelicolor* encodes both ribosome-dependent (p)ppGpp-synthesizing and -degrading activities. J. Bacteriol. 180: 4123-4132.

Mechold, U., Cashel, M., Steiner, K., Gentry, D., and Malke, H. 1996. Functional analysis of a *relA*/*spoT* gene homolog from *Streptococcus equisimilis*. J. Bacteriol. 178: 1401-1411.

Mechold, U., and Malke, H. 1997. Characterization of the stringent and relaxed responses of *Streptococcus equisimilis*. J. Bacteriol. 179: 2658-2667.

Metzger, S., Sarubbi, E., Glaser, G., and Cashel, M. 1989. Protein sequences encoded by the *relA* and the *spoT* genes of *Escherichia coli* are interrelated. J. Biol. Chem. 264: 9122-9125.

Mittenhuber, G. 1999. Occurence of MazEF-like antitoxin/toxin systems in bacteria. J. Mol. Microbiol. Biotechnol. 1: 295-302.

Mittenhuber, G. 2001. Phylogenetic analyses and comparative genomics of vitamin B_6 (pyridoxine) and pyridoxal phosphate biosynthesis pathways. J. Mol. Microbiol. Biotechnol. 3: 1-20.

Murray, K.D., and Bremer, H. 1996. Control of *spoT*-dependent ppGpp synthesis and degradation in *Escherichia coli*. J. Mol. Biol. 259: 41-57.

Nei, M., and Kumar, S. 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford, UK.

Noel, L., Moores, T.L., van der Biezen, E.A., Parniske, M., Daniels, M.J., Parker, J.E, and Jones, J.D. 1999. Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of *Arabidopsis*. Plant Cell 11: 2099-2112.

Östling, J., Flardh, K., and Kjelleberg, S. 1995. Isolation of a carbon starvation regulatory mutant in a marine *Vibrio* strain. J. Bacteriol. 177: 6978-6982.

Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. CABIOS 12: 357-358.

Persson, B.C., Jäger, G., and Gustafsson, C. 1997. The *spoU* gene of *Escherichia coli*, the fourth gene of the *spoT* operon, is essential for tRNA (Gm18) 2'-*O*-methyltransferase activity. Nucleic Acids Res. 25: 4093-4097.

Primm, T.P., Andersen, S.J., Mizrahi, V., Avarbock, D.,

Rubin, H., and Barry, C. E. III. 2000. The stringent response of *Mycobacterium tuberculosis* is required for long-term survival. J. Bacteriol. 182: 4889-4898.

Reddy, P.S., Raghavan, A., and Chatterji, D. 1995. Evidence for a ppGpp-binding site on *E. coli* RNA polymerase: proximity relationship with the rifampicin binding domain. Mol. Microbiol. 15: 255-265.

Seyfzadeh, M., Keener, J., and Nomura, M. 1993. *spoT*-dependent accumulation of guanosine tetraphosphate in response to fatty acid starvation in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 90: 11004-11008.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Tabata, S. 1999. Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res. 6: 283-290.

Shapiro, J.A. 1994. Pattern and control in bacterial colony development. Sci. Prog. 76: 399-424.

Scoarughi, G.L., Cimmino, C., and Donini, P. 1995. Lack of production of (p)ppGpp in *Halobacterium volcanii* under conditions that are effective in the eubacteria. J. Bacteriol. 177: 82-85.

Sen, K., Hayashi, J., and Kuramitsu, H.K. 2000. Characterization of the *relA* gene of *Porphyromonas gingivalis*. J. Bacteriol. 182: 3302-3304.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407: 81-86.

Stephens, R.S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R.L., Zhao, Q., Koonin, E.V., and Davis, R.W. 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science 282: 754-759.

Stent, G.S., and Brenner, S. 1961. A genetic locus for the regulation of ribonucleic acid synthesis. Proc. Natl. Acad. Sci. USA 47: 2005-2014.

Subramanian, G., Koonin, E.V., and Aravind, L. 2000. Comparative genome analysis of the pathogenic spirochetes *Borrelia burgdorferi* and *Treponema pallidum*. Infect. Immun. 68: 1633-1648.

Sugiura, M., Hirose, T., and Sugita, M. 1998. Evolution and mechanism of translation in chloroplasts. Ann. Rev. Genet. 32: 437-459.

Tatusov, R.L., Koonin, E.V., and Lipman, D.J. 1997. A genomic perspective on protein families. Science 278: 631-637.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E.V. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28: 33-36.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Toulokhonov, I.I., Shulgina, I., and Hernandez, V.J. 2001. Binding of the effector ppGpp to *E. coli* RNA polymerase is allosteric, modular, and occurs near the N-terminus of the β' subunit. J. Biol. Chem. 276: 1220-1225.

Turko, I.V., Francis, S.H., and Corbin, J.D. 1998. Potential roles of conserved amino acids in the catalytic domain of the cGMP-binding cGMP-specific phosphodiesterase. J. Biol. Chem. 273: 6460-6466.

Turmel, M., Otis, C., and Lemieux C. 1999. The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. Proc Natl Acad Sci USA 96: 10248-10253.

Urabe, H., and Ogawara H. 1992. Nucleotide sequence and transcriptional analysis of activator-regulator proteins for β-lactamase in *Streptomyces cacaoi*. J. Bacteriol. 174: 2834-2842.

van der Biezen, E.A., and Jones, J.D. 1998. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr. Biol. 8: R226-R227.

van der Biezen, E.A., Sun, J., Coleman, M.J., Bibb, M.J., and Jones, J.D. 2000. *Arabidopsis* RelA/SpoT homologs implicate (p)ppGpp in plant signaling. Proc. Natl. Acad. Sci. USA 97: 3747-3752.

Wehmeier, L., Schäfer, A., Burkovski, A., Krämer, R., Mechold, U., Malke, H., Pühler, A., and Kalinowski, J. 1998. The role of the *Corynebacterium glutamicum rel* gene in (p)ppGpp metabolism. Microbiology 144: 1853-1862.

Wendrich, T.M., and Marahiel, M.A. 1997. Cloning and characterization of a *relA*/*spoT* homologue from *Bacillus subtilis*. Mol. Microbiol. 26: 65-79.

Wendrich, T.M., Beckering, C.L., and Marahiel, M.A. 2000. Characterization of the *relA*/*spoT* gene from *Bacillus stearothermophilus*. FEMS Microbiol. Lett. 190: 195-201.

Wise, A.A., and Price, C.W. 1995. Four additional genes in the *sigB* operon of *Bacillus subtilis* that control activity of the general stress factor σ^B in response to environmental signals. J. Bacteriol. 177: 123-133.

Xiao, H., Kalman, M., Ikehara, K., Zemel, S., Glaser, G., and Cashel, M. 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. J. Biol. Chem. 266: 5980-5990.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. 1999. Crystal structure of *Thermus aquaticus* RNA polymerase at 3 Å resolution. Cell 98: 811-824.

Zomorodipour, A., and Andersson, S.G.E. 1999. Obligate intracellular parasites: *Rickettsia prowazekii* and *Chlamydia trachomatis*. FEBS Lett. 452: 11-15.