Josip Juraj Strossmayer University of Osijek
Ruđer Bošković Institute, Zagreb
University of Dubrovnik
University Postgraduate Interdisciplinary Doctoral Study
of "Molecular Biosciences"

Tina Paradžik

MOLECULAR CHARACTERIZATION, FUNCTIONAL AND PHYLOGENETIC ANALYSIS OF PARALOGOUS SSB PROTEINS FROM *STREPTOMYCES COELICOLOR*

DOCTORAL THESIS

Osijek, 2013

BASIC DOCUMENTATION CARD

Josip Juraj Strossmayer University of Osijek **PhD thesis**
University of Dubrovnik
Ruđer Bošković Institute
University Postgraduate Interdisciplinary Doctoral Study of Molecular biosciences

Scientific Area: Natural sciencies
Scientific Field: Biology

Molecular characterization, functional and phylogenetic analysis of paralogous SSB proteins from *Streptomyces coelicolor*

Tina Paradžik

Thesis performed at: Ruđer Bošković Institute, Zagreb

Supervisor/s: Dušica Vujaklija, PhD, Senior research associate

The linear chromosome of *Streptomyces coelicolor* contains two paralogous *ssb* genes, *ssbA* and *ssbB*. Mutational analysis of *ssb* genes indicated that *ssbA* is essential, whereas *ssbB* has a key role in chromosome segregation during sporulation. The two *ssb* genes were expressed differently, in accordance with their proposed biological roles. Localisation of fluorescently tagged SSB proteins was observed in the vegetative and aerial hyphae. Paralogous SSB proteins bind ssDNA with different affinity. Phylogenetic analysis of eubacterial SSBs indicated frequent duplication, loss of *ssb* genes and polyphyletic origin of well defined groups of SsbB proteins.

Number of pages: 130
Number of figures: 50
Number of tables: 14
Number of references: 158
Original in: English
Key words: SSB, paralogous proteins, *Streptomyces coelicolor*, chromosome segregation, phylogeny

Date of the thesis defense: 6th September, 2013

Reviewers:
1. Davor Zahradka, PhD, Senior Research Associate, Ruđer Bošković Institute
2. Dušica Vujaklija, PhD, Senior Research Associate, Ruđer Bošković Institute
3. Associate Professor Vera Cesar, PhD, Josip Juraj Strossmayer University of Osijek
4.
5.
6. Professor Mirolav Plohl, PhD, Ruđer Bošković Institute (substitute)

Thesis deposited in: City and University Library of Osijek, Europska avenija 24, Osijek; Josip Juraj Strossmayer University of Osijek, Trg sv. Trojstva 3, Osijek

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište Josipa Jurja Strossmayera u Osijeku **Doktorski rad**
Sveučilište u Dubrovniku
Institut Ruđer Bošković
Sveučilišni poslijediplomski interdisciplinarni
doktorski studij Molekularne bioznanosti

Znanstveno područje: Prirodne znanosti
Znanstveno polje: Biologija

Molekularno-biološka karakterizacija i filogenetska analiza paralognih proteina SSB iz bakterije *Streptomyces coelicolor*

Tina Paradžik

Rad je izrađen u: Institut Ruđer Bošković, Zagreb

Mentor/i: dr.sc. Dušica Vujaklija, viša znanstvena suradnica

Na lineranom kromosomu bakterije *Streptomyces coelicolor* nalaze se dva gena *ssb*, *ssbA* i *ssbB*. Metodom transpozonske mutageneze pokazano je kako je gen *ssbA* esencijalan dok *ssbB* ima važnu ulogu u kromosmskoj segregaciji tijekom sporulacije. Dokazana je različita ekspresija gena *ssb*, što je u skladu s predloženim biološkim ulogama. Lokalizacija flourescentno obilježenih proteina SsbA i SsbB detektirana je u vegetativnom i zračnom miceliju bakterije *S. coelicolor*. Pokazano je kako paralogni proteini SSB imaju različite afinitete vezanja jednolančane DNA. Filogenetska analiza bakterijskih proteina SSB ukazala je na učestale duplikacije i gubitke gena *ssb* u svim grupama bakterija te na polifiletsko podrijetlo dobro definiranih skupina proteina SsbB.

Broj stranica: 130
Broj slika: 50
Broj tablica: 14
Broj literaturnih navoda: 158
Jezik izvornika: engleski
Ključne riječi: SSB, paralogni proteini, *Streptomyces coelicolor,* segregacija kromosoma, filogenetska analiza

Datum obrane: 6. rujan 2013.

Stručno povjerenstvo za obranu:
1. dr. sc. Davor Zahradka, viši znanstveni suradnik Instututa Ruđer Bošković, predsjednik
2. dr. sc. Dušica Vujaklija, viša znanstvena suradnica Instituta Ruđer Bošković, mentorica i član
3. prof. dr.sc. Vera Cesar, izvanredna profesorica Sveučilišta J.J. Strossamyera u Osijeku, član
4.
5.
6. prof. dr. sc. Miroslav Plohl, znanstveni savjetnik Instituta Ruđer Bošković, zamjena člana

Rad je pohranjen u: Gradskoj i sveučilišnoj knjižnici Osijek, Europska avenija 24, Osijek; Sveučilištu Josipa Jurja Strossmayera u Osijeku, Trg sv. Trojstva 3, Osijek

Research for this PhD thesis was performed in the Laboratory of Molecular Genetics, Division of Molecular Biology, Ruđer Bošković Institute under supervision of dr. sc. Dušica Vujaklija. This work was supported by the Ministry of Science, Education and Sport of the Republic of Croatia [grant number 098-0982913-2877]. Fluorescence microscopy analysis was mainly performed at the Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, Glasgow, UK, during T. Paradžik three months stay supported by FEMS, short-term fellowship.

Table of contents

1. Introduction

1.1. Streptomycetes are an important model system for the study of bacterial development

Streptomycetes are Gram positive aerobic bacteria that belong to one of the largest phyla among Bacteria – Actinobacteria. This bacterial group includes microorganisms exhibiting a wide spectrum of morphologies, from coccoid to fragmenting hyphal forms, as well as those possessing highly variable physiological and metabolic properties. Furthermore, Actinobacteria members have adopted different lifestyles, from pathogens (e.g. *Corynebacterium*, *Mycobacterium*, *Nocardia*, *Tropheryma*, and *Propionibacterium*), soil inhabitants (*Streptomyces*), plant commensals (*Leifsonia*), to gastrointestinal commensals (*Bifidobacterium*), (Ventura, Canchaya et al. 2007).

Over a long period of time our group has been focused on studying streptomycetes, members of Actinobacteria, the best known for their ability to produce a vast array of antibiotics. *Streptomyces* species are multicellular bacteria that exhibit a complex developmental programme and morphological differentiation. Since streptomycetes have a distinctive life cycle that resembles those of multicellular eucaryotes, it was tought originally they belong to eukaryotic rather than prokaryotic organisms until the late 1950s (Hopwood 2006). They are found predominantly in soil, and their ability to colonise this environment is greatly facilitated by growth of aerial hyphae which can differentiate into spores that assist in spread and persistence (Kieser 2000). Spores are resistant to lack of nutrients and water, and many *Streptomyces* species can survive for a long time in the soil in this inactive form. Recent research revealed that the marine environment also contains a wide range of distinct streptomycetes which cannot be found in the terrestrial environment (Dharmaraj and Sumantha 2009).

Much current knowledge of streptomycetes is based on the genetic and genomic studies of *Streptomyces coelicolor* A3(2), which genome sequencing project was completed in 2002 (Bentley, Chater et al. 2002). In the last ten years the sequencing of twelve *Streptomyces* genomes was completed and large numbers are undergoing projects (from NCBI database). They all share common features such as high GC content (70 - 74 %), large linear chromosomes encoding large number of genes, the complex life cycle and production of secondary metabolites. The chromosome of *S. coelicolor* encodes 7825 ORFs (**o**pen **r**eading **f**rames) (Bentley, Chater et al. 2002). For comparison, genome of the model Gram-negative bacterium *Escherichia coli* K-12 has 4.65 Mbp genome and 4444 ORFs. Simple eukaryote, *Saccharomyces cerevisiae*, has 12.16 Mbp genome and 6275 ORFs. Large *Streptomyces* genomes are not surprising considering complexity of their life cycle, differentiation and adaptations to demanding environments such as soil. Great percentage of *S. coelicolor* genome consists of regulatory (12.3 %) and transport genes (7.8 %), as well as genes for degradation of

extracellular nutrients (10.5 %) (Bentley, Chater et al. 2002). In addition, genome analysis has revealed the distribution of different types of genes. Nearly all genes predicted to be essential, such as those for cell division, DNA replication, transcription, translation and amino-acid biosynthesis, are located in the core region whilst exceptions tend to be duplicated genes (Bentley, Chater et al. 2002). Contrary to the core region genes, the genes present in chromosome ends (so-called "chromosomal arms") are conditionally adaptive genes that are easily deleted in the laboratory. The tendency of streptomycetes to suffer large deletions and amplifications near the chromosome ends has been known for many years and is thought to be related to the chromosome linearity (Kieser 2000).

From a human perspective, the most important biochemical feature of streptomycetes is their capacitiy to produce antibiotics and other bioactive compounds (such as antitumor and immunosuppressive substances). Actinomycetes produce around 2/3 antibiotics in use, and 80 % comes from the genus *Streptomyces*. From the organism's point of view, this is an adaptation to competition in the soil. Antibiotics are made at the transition between vegetative and aerial stage of growth, and as a response to nutritional stresses sensed by the colony as it communicates with its environment (Hopwood 2007). This is achieved by the temporal and spatial control of gene expression, morphogenesis, metabolism and the flux of metabolites. Since these are mostly attributes associated to more complex biological systems extensive knowledge of streptomycetes developmental biology is of basic scientific interest. Moreover, their great biotechnological potential makes them even more significant for research.

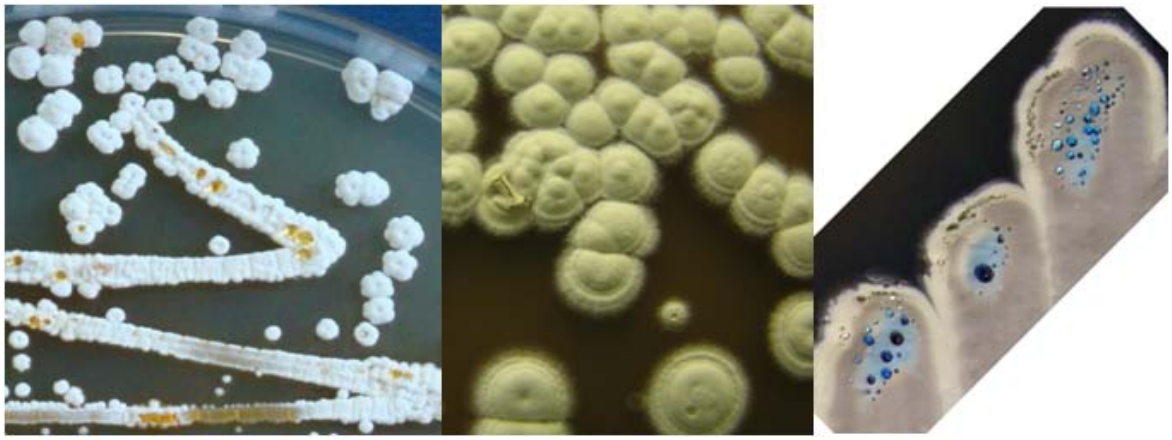


Figure 1.1. Antibiotic secretion on the top of sporulated colonies from several *Streptomyces* species (kindly provided by A. Mikoč)

1.1.1 Streptomyces life cycle

The main factors contributing to streptomycetes being considered as exceptionally interesting model organisms are their complex life cycle and morphology. In favourable life conditions *Streptomyces* spores start to germinate by swelling and produce aseptate germ tubes that grow by tip extension (Figure 1.2). Branched vegetative mycelia are formed by further growth and lateral tip branching.

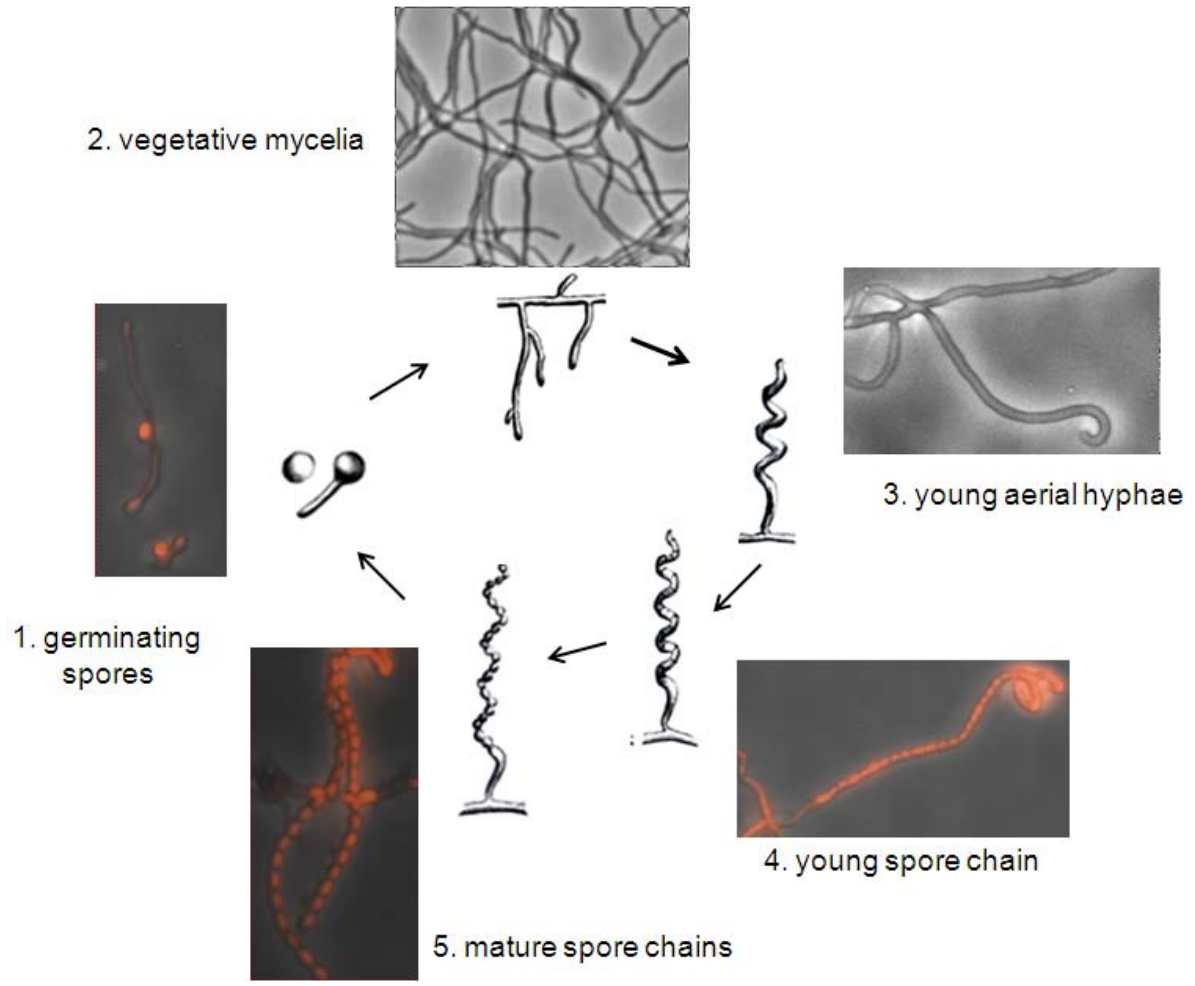


Figure 1.2. *Streptomyces coelicolor* life cycle. The life cycle starts with the spore germination and the formation of a germ tube (on the left, 1.). Hyphal growth and lateral branching produce branched vegatetive (substrate) mycelium containing multiple genome copies within single compartments (on the top, 2.). Aerial mycelium starts to emerge at the time of nutrient depletion (top right, 3.). These sporogenic structures are eventually septated (bottom right, 4.), and transformed into the chains of unigenic spores (bottom, 5.). The spores are dispersed and germinate when life conditions are favourable again.

Cross walls (septae) in vegetative hyphae are formed only occasionally (Flardh 2003), and are more frequent in older parts of mycelium, as well as branches, and never close to the tip. Therefore, exponential growth of the vegetative mycelium is achieved by the combination of tip growing and lateral branching, rather than by division. Cell division during vegetative growth results in formation of irregularly spaced hyphal cross walls, where each compartment contains multiple genome copies. This is in contrast with *E. coli,* which is a typical example of symmetric binary fission, where cell division occurs at the midd-cell, after daughter nucleoids have been segregated (Donachie 1993). In a response to nutrient limitation or other physiological stresses, aerial hyphae start to emerge. Aerial hyphae extend by the tip growth. Complex signalling network regulates aerial growth (Willey, Schwedock et al. 1993). Some changes that precede or coincide with the formation of aerial mycelia are: increased production of extracellular enzymes, onset of secondary metabolism, lysis of some compartments of substrate mycelium. Aerial hyphae differ from vegetative mycelium for their lack of branches, width and fast elongation rate. In the aerial mycelia extensive DNA replication takes place. The apical compartment of individual aerial hyphae contains many tens of genome copies. When aerial growth stops, multiple septae are formed that subdivide the apical compartment into unigenomic pre-spore compartments (Figure 1.2). The aerial hyphae give the colonies their characteristic white fluffy appearance. At the final stage of spore formation pre-spore compartments are turned into spores by wall thickening and deposition of grey pigment (Jakimowicz and van Wezel 2012).

1.1.2. DNA replication in Streptomyces

The biochemistry of replication is similar in all bacteria, but linearity of the *Streptomyces* chromosome brings with it the need for a special replication strategy. Similar to other bacteria, the DNA synthesis begins at the region called origin of replication (*ori*C). In *E. coli*, the *ori*C consists of three A-T rich 13-mer repeats and four 9-mer repeats (Baker and Wickner 1992). In streptomycetes there are 19 copies of nine-base-pair sequence (DnaA boxes) in 800 bp around *oriC*, providing extended target for DnaA, an initiator protein (Jakimowicz, Majkadagger et al. 2000). Unwinding of DNA at the origin, and synthesis of new strands, forms a replication fork. In addition to DNA polymerase, the enzyme that synthesizes the new DNA by adding nucleotides matched to the template strand, a number of other proteins are associated with the replication fork and assist in the initiation and continuation of DNA synthesis. The DNA polymerase can work only in one direction; 5' to 3' (see Figure 1.4). This means that replication progress is different in two DNA strands. One strand, called continuous strand (i.e. leading strand), is progressing in the direction in which

replication fork is travelling, while the other has to be constantly reinitiated and extended back to join onto the segment of new DNA just being made. The segments of discontinuous strand (i.e. lagging strand) are called Okazaki fragments. Since DNA polymerase can make new DNA only by extending an existing DNA strand, the linear ends of *Streptomyces* chromosome represent potential problem in replication. *Streptomyces* have a unique solution to this universal "end problem". They have a special protein which is bonded to the free 5' end of the chromosome, primes DNA synthesis, allowing the DNA polymerase to make the terminal Okazaki fragment from the very end of the chromosome without the need for an RNA primer (Bao and Cohen 2003).

Replication of *Streptomyces* chromosomes seems asynchronous both in vegetative and aerial hyphae. Chromosome replication takes place in many compartments of both types of hyphae, with the apical compartments of the aerial mycelium exhibiting the highest replication activity. Within a single compartment, the number of ongoing DNA replications seems to be lower than the chromosome number, indicating that this process is asynchronous and that only selected chromosomes undergo replication at any one time (Ruban-Osmialowska, Jakimowicz et al. 2006). In the unicellular bacteria, chromosome replication is strictly coupled with the cell division: when a cross wall forms in the centre of the cell, one chromosome copy ends up in each new cell, in process called chromosome partitioning. Replication in *Streptomyces* is not directly coupled with the cell division, since it has multiple chromosome copies in each hyphal compartment (Hopwood 2007).

1.1.3. Streptomyces development and regulation of chromosomal segregation

The mycelial life style of streptomycetes imposes novel requirements for the cell division and its control mechanisms. As nutrients in the colony's surroundings become exhausted, the aerial mycelium formation must be initiated. This involves genetically regulated microengineering of amazing complexicity (Hopwood 2007). Since the cell division is not essential for growth of the hyphae (McCormick 2009), *Streptomyces* are ideal organisms for the analysis of cell division proteins.

First streptomycetes developmental mutants were described a decades ago (Hopwood 1967). Those mutants were called *bld* mutants (for bald - "hairless") as they had serious defects in development of aerial mycelia. No clear regulatory role has been established for any *bld* gene until the discovery that *bldD* encodes a repressor of several developmental genes during growth (Kelemen, Viollier et al. 2001). Many other *bld* genes are now recognized to be involved in the complex regulatory network that govern the aerial growth (Chater 2001).

Sporulation events in *Streptomyces* require several *Streptomyces*-specific mechanisms, not found in the other bacteria, particularly in the terms of regulation. An important set of sporulation

regulators are those encoded by the *whi* genes. *Whi* mutants are defective in the synthesis of the grey polyketide spore pigment. The *whi* genes control the expression of a number of sporulation associated genes such as *ftsZ*, *parAB* and others. Transformation of sporogenic aerial hyphae into chains of mature unigenomic spores requires two synchronous processes: formation of sporulation septae and chromosome segregation (Flardh and Buttner 2009). Although a number of genes are known to be involved in these processes, there are still many gaps in understanding how the septum site localisation is controlled and how septum formation and chromosome segregation are co-ordinated. Analysis of the spores of several mutants, such as *ssgG* (Noens, Mersinias et al. 2005) have shown that chromosome condensation may take place independently of the divisome (cell-division machinery). There is normal DNA condensation and segregation in *ssgG* mutants which frequently skip the septae formation during sporulation, thus producing large spores containing multiple well-condensed chromosomes (Noens, Mersinias et al. 2005).

The inventory of the cell division machinery in *Streptomyces* highlights many similarities to the cell mechanism found in most other bacteria. The cell division protein FtsZ polymerises into protofilaments, which assemble to structures called Z-rings (Erickson 1995). The Z-ring mediates the recruitment of the divisome to the mid-cell position (Adams and Errington 2009) and provides driving force for the constriction during cell division (Erickson and Osawa 2010). Because unicellular bacteria can propagate only by making daughter cells, FtsZ is essential for their viability. Contrary to this, *S. coelicolor* mutant lacking *ftsZ* gene is viable, even though it made no cross walls in any part of the colony (McCormick, Su et al. 1994). It could not make spores but otherwise was normal in the appearance, despite the fact that the colony was a single giant cell. While FtsZ ring formation is required for cross-wall formation, the other divisome components are not (Jakimowicz and van Wezel 2012). SsgA-like proteins (SALPs) are a group of actinomycete-specific family that play important role in morphogenesis and control of cell division. *S. coelicolor* contains seven SALPs. Fluorescent microscopy analysis of labelled Ssg proteins revealed that SsgA localise to the sites where the cell wall remodelling takes place and it was shown that it activates germination and cell division (Noens, Mersinias et al. 2007). The SsgB protein is required for the sporulation, and co-localise with FtsZ during the process of sporulation-specific cell division (Willemse, Borst et al. 2011), (Figure 1.3). Interestingly, during early aerial growth SsgB foci are formed while FtsZ foci are still diffused (Figure 1.3). Formation of FtsZ protofilaments *in vitro* are stimulated by SsgB protein (Willemse, Borst et al. 2011). This point to the fact that FtsZ is not the first protein to arrive at the septum site, contrary to other systems, such as those studied in the *E. coli* and *B. subtilis*. Therefore, this explains the absence of mid-cell reference point in streptomyces, in which different way of

control of Z-ring formation evolved. Similarly to other bacteria, the chromosome segregation in *Streptomyces* involves ParA and ParB proteins. ParA homologues are Walker A ATPases, which can form filaments and are showing dynamic localisation in the cells (Figure 1.3). ParB homologues are DNA binding proteins that form high-order nucleoprotein complex with the partition sites (*parS*). The *par* genes are expressed even during vegetative growth, although in the vegetative hyphae the chromosomes are not precisely segregated. This suggests their participation possibly in the regulation of replication role even at this growth stage (Jakimowicz and van Wezel 2012). In the aerial hyphae, ParAB proteins ensure proper partitioning of the chromosomes into prespore compartments before septation (Kim, Calcutt et al. 2000). When ParB binds *parS* sites around origin of replication, ParA mediates efficient assembly of these complexes and provides scaffolding for their proper segregation (Jakimowicz, Zydek et al. 2007). Mutants that lack these genes produce 19 % (*parA*) or 24 % (*parB*) of anucleate spores (Kois, Swiatek et al. 2009). Furthermore, novel interacting partner of ParA was repored recently. ParJ negatively regulates ParA polymer assembly (Ditkowski, Troc et al. 2010) and if deleted, 8 % of prespores lack DNA.

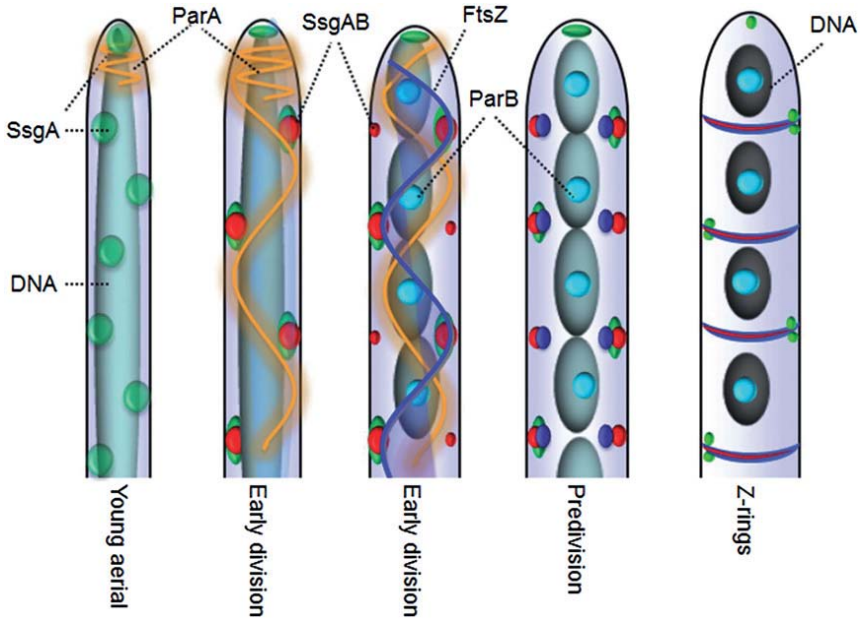


Figure 1.3. The proposed model of the control of cell division in *Streptomyces*. In young aerial hyphae, SsgA form foci, while SsgB and FtsZ are diffuse. ParA is constrained to the hyphal tip. In the next step (early division), SsgA and SsgB colocalise temporarily at typically alternating sides of the aerial hyphae, while ParA extends as a filament along the hyphae. In the presence of ParA, ParB nuceloprotein complexes are formed over the yet uncondensed DNA. Prior to sporulation-specific cell division, FtsZ is seen as long spiral filaments through the hyphae. In the predivision hyphae FtsZ and SsgB continue to colocalise, and SsgB tethers FtsZ to the membrane. Z-rings are formed at the sporulation stage, followed by chromosome condensation and segregation and formation of the sporulation septa. Adapted from (Jakimowicz and van Wezel 2012).

Several other genes were identified to be involved in chromosome organisation and distribution during S*treptomyces* sporulation. SMC (structural maintenance of chromosomes) protein participates

in the condensation and organisation of the nucleoid (Kois, Swiatek et al. 2009). FtsK is a motor protein that removes each genome copy before septae closure. Various *ftsK* mutants frequently have irregular DNA content due to large terminal deletion (Wang, Yu et al. 2007). Double mutants of these genes, *smc/ftsK'*-truncation alleles produced from 7 % to 15 % of anucleate spores, while segregation defect of 24 % was observed for *smc/parB* mutant (Dedrick, Wildschutte et al. 2009).

In summary, different combinations of double and triple mutant genes reported and mentioned here still produced viable spores, suggesting that additional genes were responsible for proper developmental genome segregation.

1.2. SSB protein overview

SSB (**s**ingle-**s**tranded DNA **b**inding) proteins are found in all kingdoms of life and are indispensable for the cell survival (Szczepankowska, Prestel et al. 2011). These specialized class of proteins have evolved in cells to protect single-stranded DNA (ssDNA) intermediates and to prevent the formation of secondary structures during processes of DNA recombination, replication and repair (Glassberg, Meyer et al. 1979). To accomplish these obligatory processes essential for the maintenance of genome integrity the double-stranded DNA (dsDNA) has to be unwound, but resulting ssDNA is prone to the chemical and nucelolytic attacks that can produce breaks or lesions. SSB proteins bind to single stranded DNA with high affinity and in a sequence independent manner, thus protecting it from degradation and formation of unproductive secondary structures (Figure 1.4). This allows undisturbed functioning of other proteins involved in processes of DNA metabolism. The SSB from *E. coli* (SSB_{Eco}) has been extensively studied over 40 years (Meyer and Laine 1990) and consequently it has become a prototypic model protein for studying ssDNA-SSB interactions (Shereda, Kozlov et al. 2008).

The vast majority of eubacterial SSB proteins display two distinct domains: an N-terminal domain containing a conserved oligonucleotide-oligosaccharide fold (OB-fold) responsible for ssDNA binding, and a C-terminal domain (Ct) enriched in glycine and acidic amino-acids. ssDNA binding domain of SSB proteins is topologically identical to a number of proteins known to bind oligonucleotides or oligosaccharides and its fold belongs to the well-characterized OB (oligomer-binding) fold (Murzin 1993). Although these proteins do not have sequence homology, their structures share a common oligosaccharide/oligonucleotide-binding fold (OB-fold), suggesting that they have a similar mode of action.

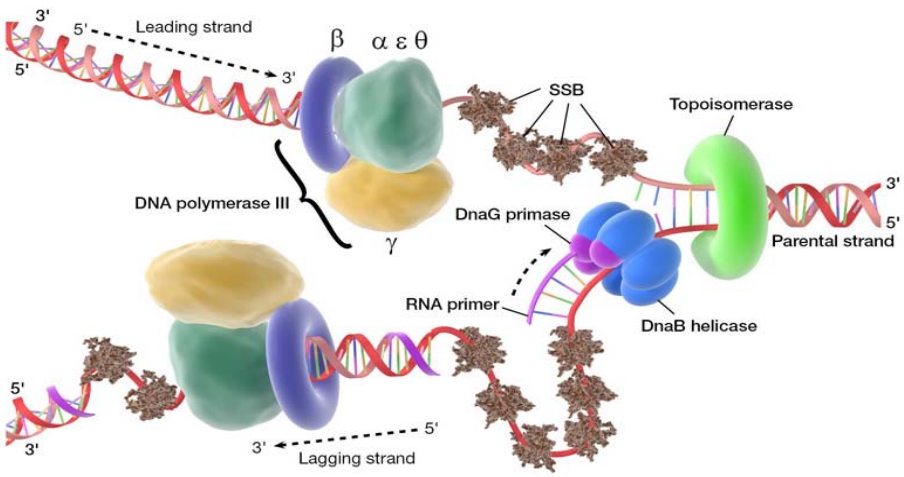


Figure 1.4. SSB protein in replication fork (adapted from www.pdbj.org/eprots/index_en.cgi?PDB%3A3BEP). SSB is bound to ssDNA, assuring proper activity of other proteins involved in this process.

The OB-fold is comprised of antiparallel β-sheets forming a barrel with a well defined cleft (Figure 1.5). OB-fold binds to ssDNA through a combination of electrostatic and base-stacking interactions with the phosphodiester backbone and nucleotide bases (Raghunathan, Kozlov et al. 2000). Most bacterial SSBs function as homotetramers in which four OB-folds act together to bind to ssDNA. There are some rare exceptions to this type of domain arrangement, such as SSBs from Deinococcus –Thermus group, which contain 2 OB-folds per monomer and assemble as homodimers (Bernstein, Eggington et al. 2004; Fedorov, Witte et al. 2006). Nevertheless, these SSBs still have 4 OB-folds in total.

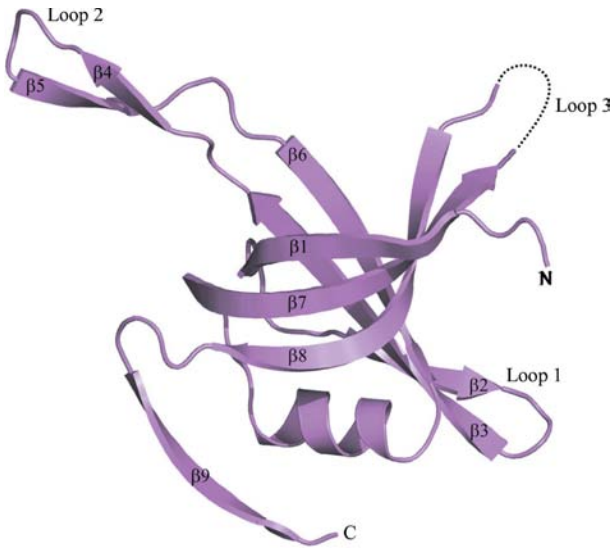


Figure 1.5. Ribbon representation of OB fold that is present in the crystal structure of the *S. coelicolor* SSB. There are seven ß-strands, one α–helix and three loops forming a ß-barrel capped by the α–helix. Adapted from (Stefanic, Vujaklija et al. 2009)

Contrary to eubacterial SSBs, eukaryotic SSB (RPA, replication protein A) has heterotrimeric structure. Majority of ssDNA binding activity of RPA resides in the largest subunit. Similary to eubacterial SSB, the central function of RPA is stabilization of ssDNA, and this protein also associates with the other proteins involved in DNA replication, recombination and repair. These interactions play a central role in the initiations of DNA processes, suggesting RPA as an important regulatory target (Wold 1997). In addition to this, several bacteriophage and viral SSB proteins function as monomers (T4 gp32) and dimers (T7 gene 2.5) (Shamoo, Friedman et al. 1995).

The primary structure of SSB protein displays two distinct domains (Figure 1.6). The Ct region of SSB proteins is enriched in glycine and proline residues with acidic amino acids in a hexapeptide motif (D-D-D-I/L-P-F) which are important for protein interactions (Curth, Genschel et al. 1996) (Figure 1.6). In contrast to the proline- and glycine-rich sequence, the region of the last 10 amino acids is highly conserved among procaryotic SSB proteins.

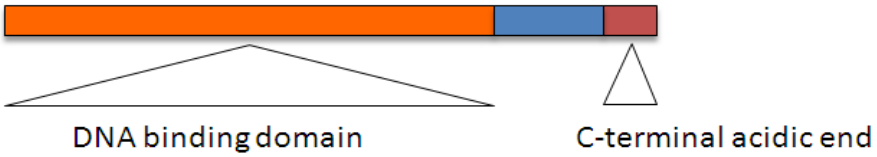


Figure 1.6. Schematic representation of SSB protein domains. This figure shows general scheme for SSB protein primary structure involving N-terminal ssDNA binding domain (OB fold), and variable C terminal region with conserved acidic tail.

In spite of realtively low sequence identity (36 %), the single-stranded DNA binding protein of human mitochondria (HsmtSSB) and SSB_{Eco} display high degree of structural homology and similar ssDNA binding properties. HsmtSSB lacks a region homologous to the C-terminal third of SSB_{Eco}. Although the DNA binding properties of HsmtSSB and SSB_{Eco} are quite similar, HsmtSSB does not function in *E. coli.* Experiments with chimeric proteins have shown that unlike many bacterial SSBs, SSB_{Eco} and HsmtSSB cross-species heterotetramers cannot be formed (Curth, Genschel et al. 1996). This failure cannot be complemented by fusing the C-terminal third of SSB_{Eco} to HsmtSSB, thus differences in the N-terminal parts of both proteins must be responsible for this incompatibility. This reflects structural differences between the N-terminal parts of HsmtSSB and SSB_{Eco} (Curth, Genschel et al. 1996).

Fluorescence titrations and DNA-melting curves showed that the C-terminus of SSB_{Eco} is not essential for DNA-binding *in vitro*. The affinity for single-stranded DNA is even increased by the removal of the last 10 amino acids. Various *E. coli* SSB point mutations in Ct region have effect on temperature sensitivity and sensitivity to DNA damage (Meyer, Glassberg et al. 1980; Johnson 1984).

The importance of the C-terminus is illustrated by the *ssb-113* mutation, in which the penultimate residue of the SSB_{Eco} protein (Pro 176) is replaced by serine (Chase, L'Italien et al. 1984). This mutation results in a UV- and temperature-sensitive phenotype. In addition, the mutant SSB lacking the last 10 amino acids is not able to substitute wild-type SSB_{Eco} *in vivo*. While the *in vitro* nucleic acid binding properties are slightly affected by deletion of C-terminus, this region is essential for *in vivo* function (Curth, Genschel et al. 1996). Interestingly, while lack of SSB_{Eco} C-terminus has fatal impact on *E. coli*, deletion of C-terminal end of SSB is not lethal to *Bacillus subtilis* and has moderate impact on its growth (Costes, Lecointe et al. 2010). However, it affects the efficiency of repair of damaged genomic DNA and of accidentally arrested replication forks. This was demonstrated by the lack of localisation of eGFP labelled replisome proteins (DnaE, SbscC, RecJ, RecO, RarA) in the cells containing truncated SSB (*ssb*ΔCter). Impact on DNA repair was more pronounced under growth conditions that are stressful for the genome. Mutant cells lacking C-terminus of SSB protein are sensitive to UV radiation nearly as *recA*$^-$ mutants. It was proposed that *ssbΔC*ter mutant cells of *B. subtilis* display defficiency in RecA loading on ssDNA, thus explaining bacterial inefficiency in triggering the SOS response upon exposure to genotoxic agents (Costes, Lecointe et al. 2010).

The crystal structure of SSB_{Eco} was solved first. As reported (Raghunathan, Ricard et al. 1997), each monomer in the tetramer is topologically similar to an OB-fold. In the SSB_{Eco} dimer, two antiparallel β-sheets from two SSB subunits form an extended six-stranded antiparallel β-sheet. Two dimer-dimer interfaces are observed within the structure. One of these stabilizes the tetramer in the solution (Raghunathan, Ricard et al. 1997). In addition, the crystal structures of bacterial SSBs from the following species have been determined: SSB_{Eco} with ssDNA (Raghunathan, Kozlov et al. 2000), *Mycobacterium tuberculosis (Saikrishnan, Jeyakanthan et al. 2003)*, the archeon *Sulfolobus solfataricus* (Kerr, Wadsworth et al. 2003), *Deinococcus radiodurans* (Bernstein, Eggington et al. 2004), *Mycobacterium smegmatis* (Saikrishnan, Manjunath et al. 2005), *Thermus aquaticus* (Fedorov, Witte et al. 2006; Jedrzejczak, Dauter et al. 2006), *Mycoplasma pneumoniae (Das, Hyun et al. 2007)*, *Helicobacter pylori* complexed with ssDNA (Chan, Lee et al. 2009), *Streptomyces coelicolor (Stefanic, Vujaklija et al. 2009)*, *B. subtilis* SsbB with ssDNA (Yadav, Carrasco et al. 2012), *Deinococcus radiodurans* with ssDNA (George, Ngo et al. 2012). Interesting variation in the quaternary structure with respect to quaternary structure of SSB_{Eco} and HmtSSB (Yang, Curth et al. 1997) were reported for two genus of Actinobacteria. Owing to the orientation between *AC* and *BD* subunits the DNA-binding surface in the case of *E. coli* SSB is an approximate spheroid, whereas mycobacterial or streptomycetes SSB is an ellipsoid (Saikrishnan, Manjunath et al. 2005; Stefanic, Vujaklija et al. 2007). The latest two belong to distantly related genera of the phylum Actinobacteria. Mycobacteria are

slow growing widespread bacteria with some pathogenic properties, while the streptmycetes are soil-inhibiting filamentous bacteria best known as antibiotic producers. What they have in common is high GC content in their genome. In addition to the observed differences in quaternary structure, SSBs from both bacteria also contain β-strand nine (Figure 1.5, Figure 1.15) as an insertion at the C-terminus of the OB-fold. These strands form clamp like mechanism that together with overall structure contributes to higher stability of the homotetramer (Saikrishnan, Manjunath et al. 2005; Stefanic, Vujaklija et al. 2009).

In the 2000, the crystal structure of SSB_{Eco} bound to two $(dC)_{35}$ oligonucelotides has been published (Raghunathan, Kozlov et al. 2000). Each $(dC)_{35}$ was bound to two OB-folds (Figure 1.7). The ssDNA is bound in a groove in which both nucleic acid backbone and base interact with the protein. Spectroscopic studies suggested that Trp 40 and Trp 54 form stacking interactions with the bases (Casas-Finet, Khamis et al. 1987) and mutagenesis of these residues reduced ssDNA binding affinity (Ferrari, Fang et al. 1997). Other studies have shown that Phe 60 is involved in ssDNA binding (Bayer, Fliess et al. 1989). Consistent with these observations, Trp 40, Trp 54 and Phe 60 make extensive interactions with the ssDNA in the SSB-ssDNA complex. In addition, several lysine residues have been proposed to participate in ssDNA binding. These Lys residues, as well as N-terminal amine, are within contact distance of the ssDNA backbone and selective acetylation of these residues is expected to have significant impact on ssDNA binding. Other basic residues make interactions with ssDNA, either with the ssDNA bases (Arg 3) or with the phosphate backbone (Arg 84).

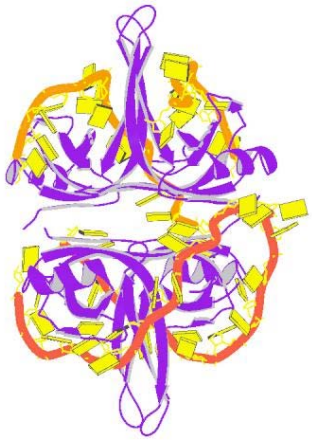


Figure 1.7. Crystal structure of chymotryptic fragment of *E. coli* SSB bound to two 35-mer single stranded DNAs, adapted from: http://www.rcsb.org/pdb/explore/explore.do?structureId=1EYG

Due to the presence of four OB-folds in each SSB tetramer, i.e. four ssDNA binding sites, these proteins can interact with ssDNA in multiple binding modes. Two major binding modes of the SSB tetramer have been proposed (Figure 1.8). In the $(SSB)_{35}$ binding mode two subunits of the *E. coli* tetramer bind to 35 nucleotides, while in the $(SSB)_{65}$ mode all four subunits participate in binding to

65 nucleotides. In $(SSB)_{35}$ mode eubacterial SSBs can bind ssDNA in highly cooperative manner, which leads to formation of SSB protein filaments on the long ssDNA (Lohman and Ferrari 1994). Although it has been proposed (Roy, Kozlov et al. 2007) that this mode has a role in DNA replication, the exact role of this cooperativity remains unclear. The relative stabilities of different binding modes are dependent on SSB/ssDNA ratio, monovalent salt concentrations, Mg^{2+} concentration, polyamines, spermine and spermidine (Shereda, Kozlov et al. 2008). As mentioned, it is not clear whether different binding modes have some role *in vivo*, although it has been proposed that they may be selectively used in different processes in the cell.

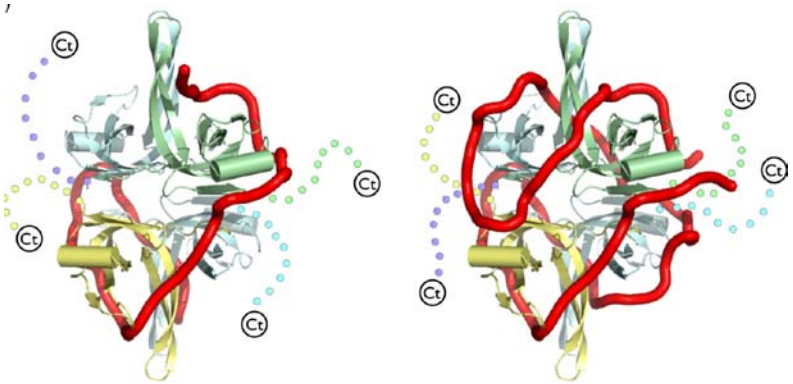


Figure 1.8. Proposed structures of the *E. coli* $(SSB)_{35}$ (left) and $(SSB)_{65}$ (right) ssDNA binding models 34. Each monomer in the tetramer is separately coloured and its Ct is shown as a dashed line. ssDNA is shown as a red tube. Adapted from (Shereda, Kozlov et al. 2008)

ssDNA binding studies of SSBs from two closely related species of Actinobacteria, *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*, disclosed both similarities and differences in comparison to SSB_{Eco}. The apparent equlibrium association and dissociation constants were similar for all three SSBs under comparable experimantal conditions. Nevertheless, the binding of SSB proteins from mycobacteria to ssDNA was substantially weaker in a high ionic strength buffers (Reddy, Guhan et al. 2001). The mode of DNA binding of mycobacterial SSBs is different from that of SSB_{Eco} partly on account of the difference in the shape of the tetramers. Owing to the variation in the quaternary structure, the DNA-binding surface in the case of SSB_{Eco} is an approximate spheroid, while that in the case of SSB from Mycobacteria is an ellipsoid (Saikrishnan, Manjunath et al. 2005).

More recently, ssDNA binding of SSB proteins from several bacteria was visualised by high resolution atomic force microscopy (AFM) (Grove, Willcox et al. 2005; Hamon, Pastre et al. 2007; Jain, Zweig et al. 2012). At a low ionic strength (20 mM NaCl) free ssDNA, partly formed complex and nearly fully saturated complex coexist, that is a typical feature of cooperative binding of *E. coli* SSB to ssDNA. Figure 1.9 represents partially saturated and fully saturated ssDNA. At a high inonic strength

buffer (300 mM NaCl), and *E. coli* SSB tetramer: nucleotide concentration ratio below the nucleoprotein filament saturation (1/120), different structures coexist in the AFM image, from naked ssDNA to more or less saturated nucleoprotein filaments. This represents the limited cooperativity which occurs at such high ionic strengths. At the higher *E. coli* SSB tetramer concentration (SSB_{Eco} tetramer: nucleotide concentration ratio of 1/40), the saturated filaments are observed, the contours of which were well defined. Compared to experiment in a low ionic strength, the M13 ssDNA–SSB_{Eco} complexes in high salt have a lower contour length which indicates a different binding mode of SSB_{Eco}. This is in agreement with the previously proposed model (Bujalowski and Lohman 1986).

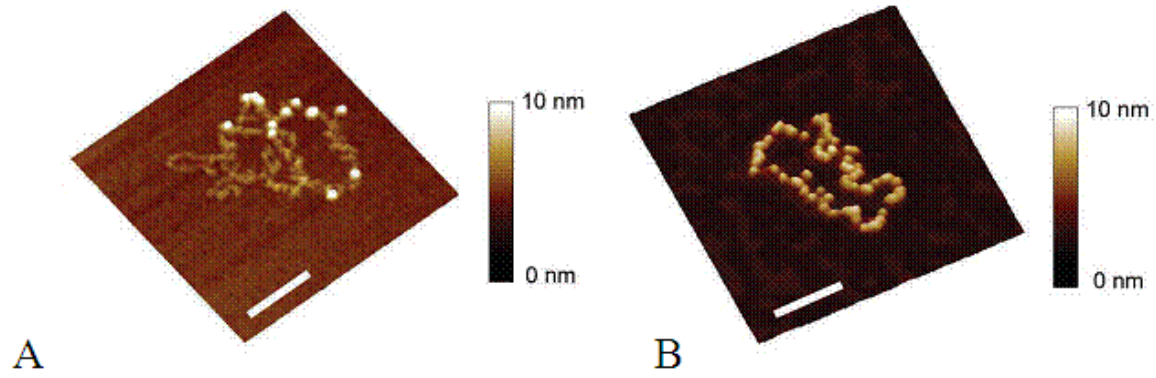


Figure 1.9. Atomic force microscopy of *E. coli* SSB bound to ssDNA. **A**. M13 ssDNA SSB complexes, above the level of saturation (R=1/320) few SSB proteins can be distinguished; **B**. M13 ssDNA SSB complexes, at R= 1/40 ssDNA is fully saturated. R=ssDNA/SSB ratio. Adapted from (Hamon, Pastre et al. 2007).

1.2.1. SSB proteins interactions

SSB proteins have often been described as inert, protective DNA coatings. Recent data demonstrate far more complex role of SSB proteins in all aspects of DNA metabolism. Eubacterial SSB proteins interact with more than a dozen cellular proteins, and most, if not all, of these interactions are mediated through Ct region of SSB protein (Shereda, Kozlov et al. 2008). The Ct tail of SSB_{Eco} has Asp-Phe-Asp-Asp-Asp-Ile-Pro-Phe (DFDDDIPF) sequence, and this motif is much conserved among SSBs from other bacteria (Figure 1.10). It is often referred to as “acidic tail”, although it should be considered as an amphypatic sequence element since it has three hydrophobic amino acids at the very end.

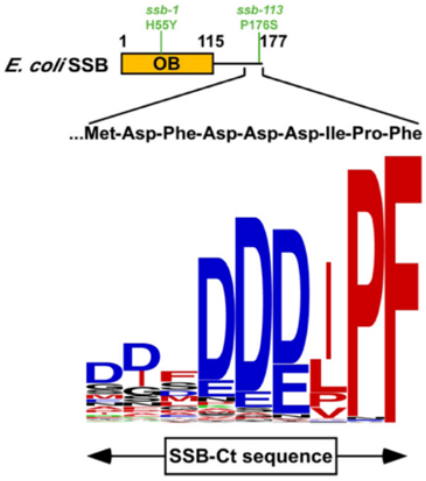


Figure 1.10. The sequence of the SSB_{Eco} -Ct element is displayed with its conservation across 280 eubacterial species represented as a logo in which the height of the residue relates to its frequency at the given position. Logo residues are coloured to indicate the hydrophobic (red), electronegative (blue), polar (black), or electropositive (green) nature of their side chains. Adapted from (Shereda, Kozlov et al. 2008).

C-terminus is structurally dynamic and unfolded, and thus cannot be seen in the crystal structures of SSB proteins (Savvides, Raghunathan et al. 2004; Lu and Keck 2008; Shereda, Kozlov et al. 2008). Ct region is more prone to proteolytic activity when SSB is bound to ssDNA. Moreover, it was shown that deletion of Ct has an effect on SSB binding mode (Kozlov, Cox et al. 2010). Removal of the acidic C-terminal end increases the intrinsic affinity for ssDNA and decreases cooperativity between ssDNA binding sites, indicating that the C-termini exert an inhibitory effect on ssDNA binding. This inhibitory effect decreases as the salt concentration increases. Binding of ssDNA to approximately half of the SSB subunits relieves the inhibitory effect for all of the subunits. Recent study on phage T7 gene 2.5 SSB protein showed that Ct tail competes with ssDNA for binding to OB-fold (Marintcheva, Marintchev et al. 2008). In the proposed model Ct tail is bound to OB fold in the absence of ssDNA, while the presence of ssDNA removes Ct tail, thus leaving it free for the interactions with other cellular proteins (Figure 1.11).

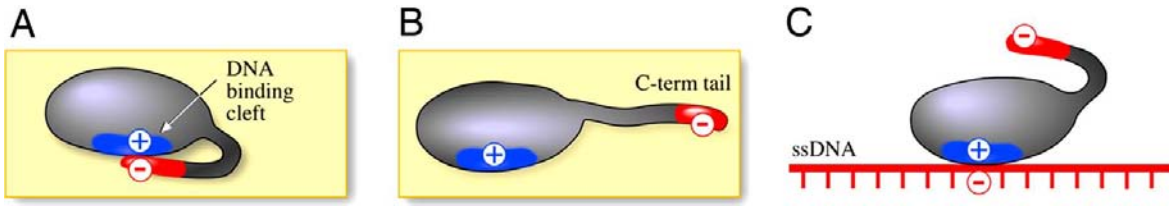
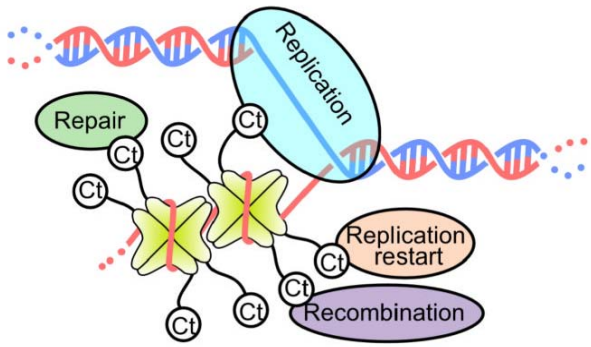


Figure 1.11. SSB binding to ssDNA is modulated through interaction with its Ct tail. **A.** Free protein. Negatively charged residues from Ct tail (red) are in contact with positively charged residues (blue) via electrostatic interactions. **B.** Displacement of the C-terminal tail results in the highly accessible DNA-binding cleft. **C.** Bound state. The C-terminal tail is displaced, and previously shielded basic residues contact the phosphate backbone (in red). Adapted from (Marintcheva, Marintchev et al. 2008).

E. coli SSB protein is known to interact with at least 14 other proteins (Figure 1.12). This interaction network involves proteins included in all aspects of DNA metabolism (Figure 1.12 B). For example, SSB interacts with primase during the process of DNA replication. Primase is a specialized form of RNA polymerase, which has a role in generating RNA primers in *oriC*-dependent replication and in replication restart processes (Benkovic, Valentine et al. 2001). Primase-SSB interaction is disrupted by binding of multi-subunit replicative DNA polymerase III (Pol III HE) to SSB (Yuzhakov, Kelman et al. 1999). This is achieved through direct binding of SSB to χ subunit of Pol III HE. Clamp loader is a complex within Pol III HE that loads the processivity subunit (β) onto DNA and helps to tie the holoenzyme together. Although χ subunit is not required for clamp loader onto ssDNA, it facilitates the assembly of clamp loader itself. In this replication step primase dissociates from the RNA-DNA duplex and allows the clamp-loader assembly to occur. Thus, the χ/SSB interaction plays a crucial role in the Pol III HE function by driving detachment of primase from RNA primers, which stimulates primer hand-off to Pol III HE.



SSB-interacting protein	Require SSB-Ct?
DNA polymerase III, chi	Yes
Primase	Unknown
Topoisomerase III	Yes
RecQ	Yes
RecO	Yes
RecJ	Unknown
RecG	Yes
PriA	Yes
PriB	Unknown
Exonuclease I	Yes
Uracil DNA glycosylase	Yes
DNA polymerase II	Yes
DNA polymerase V	Yes
Exonuclease IX	Unknown
vRNA polymerase	Unknown

Figure 1.12. A. Schematic representation of SSB interactions. SSB proteins (yellow) are depicted as tetramers with extruded Ct which interacts with the proteins involved in the major genome maintenance pathways of DNA replication (teal), recombination (purple), replication restart (orange), and repair (green). **B.** List of proteins that are known to physically interact with SSB. Highlighting colours indicate the major genome maintenance activities of the proteins (colour coding as in (A). Adapted from (Shereda, Kozlov et al. 2008).

SSB also interacts with many proteins known to play major role during DNA recombination, such as proteins from the RecF recombination pathway (Tseng, Hung et al. 1994). SSB associate with RecQ DNA helicase and increases its affinity for DNA (Shereda, Bernstein et al. 2007). It has been also shown that Topoisomerase III is involved in this interaction as well. The RecQ/Topoisomerase III pair, among other activities, has a role in the resolution of converging replication forks and this reaction is mediated through the interaction of both of these enzymes with the SSB (Suski and Marians 2008). The RecJ exonuclease from *E. coli* degrades single-stranded DNA (ssDNA) in the 5'-3' direction and participates in homologous recombination and mismatch repair. RecJ DNA binding and degradation are stimulated by SSB (Han, Cooper et al. 2006). RecG is another helicase which activity is influenced by SSB. This interaction was demonstrated *in vivo* in *B. subtilis* (Lecointe, Serena et al. 2007). SSB and RecG colocalise at the foci proposed to be stalled replication forks. This colocalisation is ablated in the cells where SSB lacks its 35 C-terminal amino acids.

RecO is important mediator protein in the RecF recombination pathway (Clark and Sandler 1994). It binds ssDNA and dsDNA and possesses a DNA annealing activity (Kantake, Madiraju et al. 2002). This annealing activity is stimulated by SSB (via Ct) and inhibited by RecR protein. Together with RecF and RecR, RecO functions as a modulator of RecA activity (Kowalczykowski, Dixon et al. 1994). RecO and RecR facilitate RecA loading onto SSB-coated ssDNA. The ability of RecOR to load RecA is greatly reduced in mutant cells with SSB variants lacking C-terminus (Hobbs, Sakai et al. 2007). Therefore, the RecA and SSB proteins are linked in their functions by three other proteins: RecF, RecO and RecR. RecA protein promotes DNA strand exchange which is the most important point of all recombination processes (Cox 2007). SSB plays a complex role in RecA mediated processes, however in *E. coli* there is no proof that they interact directly without the presence of DNA (Umezu and Kolodner 1994). Instead, RecFOR system serves for RecA loading onto SSB coated ssDNA (Umezu, Chi et al. 1993). Contrary to that, it was shown that SSB from *Mycobacterium smegmatis* interacts with RecA through its C-terminal region (Reddy, Guhan et al. 2001).

More recently study on *B. subitilis* SSB interactome revealed (Costes, Lecointe et al. 2010) 12 interacting proteins, among which there were some previously described for $SSB_{Eco,}$ such as RecQ and RecG, while others, such as XseA (large subunit of ExoVII) and YrrC, predicted protein of unknown function, and were not described before.

Interesting example of SSB activity is interaction of coliphage N4 virion RNA polymerase (vRNAP) with $SSB_{Eco.}$ SSB_{Eco} is required for early transcription of N4 virion on supercoild DNA containing vRNAP promoters. vRNAP initiates transcription on single-stranded ssDNA containing vRNAP promoter. However, RNA product is not displaced in this step, thus limiting template usage to one round. SSB_{Eco}

binds to the template and to the nascent transcript and prevents the formation of transcriptionally inert RNA-DNA hybrid. Thus, SSB activates vRNAP transcription through template recycling. On the contrary to T7 RNA that contains domain responsible for RNA displacement, vRNAP lacks this region. It was proposed that SSB_{Eco} functionally substitutes this N-terminal domain of T7 RNAP responsible for RNA binding (Davydova and Rothman-Denes 2003). Another example of SSB interaction with RNA polymerase comes from archea *Sulfolobus solfataricus*. SSB stimulates promoter melting and thus stumilates the transcription. This interaction requires both ssDNA binding domain and C-terminal region of SSB protein (Richard, Bell et al. 2004).

1.2.2. Phosphorylation of SSB proteins

Phosphorylation of single-stranded DNA-binding protein was firstly discovered in eukaryotic cell as phosphorylation of serine and threonine residues of replication protein A (RPA), which occurs during normal mitotic cell cycle progression and also in response to genotoxic stress (Din, Brill et al. 1990; Vassin, Anantha et al. 2009).

In comparison to this, eubacterial SSB proteins are modulated through phosphorylation of tyrosine residues. Experiments reported by the authors (Mijakovic, Petranovic et al. 2006) suggest that this modification is conserved among phylogentically distant bacteria, *B. subtilis*, *S. coelicolor* and *E. coli* (Mijakovic, Petranovic et al. 2006). During the overexpression in *E. coli* SSBs from all listed bacteria were posttranslationally modified by phosphorylation. MS-analysis identified the phosphorylation site of *B. subtilis* SSB to be tyrosine residue 82. This tyrosine occupies the same molecular space as tyrosine 98 in SSB_{Eco} and tyrosine 88 in SSB from *S. coelicolor* (Figure 1.13). Tyrosine phosphorylation of eubacterial protein is very interesting regarding evolution of posttranslational protein modifications, since it has been considered for a very long time that this type of modification is exclusively eukaryotic. The possible biological importance of this modification was further supported by *in vitro* phosphorylation of SSB protein from *B. subtilis* with the homologous tyrosine kinase. SSB phosphorylation increased its ssDNA binding affinity by 200 folds. In addition, it was shown by *in vivo* experiment that the induction of DNA damage by Mitomycin C, decreased phosphorylation of SSB from *B. subtilis*. These observations are in agreement with previously published data (Kuzminov 1999), i.e. SSB pre-bound to ssDNA inhibits the nucleation stage of RecA and subsequent RecA-dependent repair. Although SSB is crucial for maximal activity of RecA-mediated reactions, the interplay of affinity for DNA binding between SSB and RecA must be precisely tuned in concert with cell requirements. Thus, change in tyrosine phosphorylation during DNA damage response suggested that binding affinity of SSB must be tuned down to allow binding of

RecA protein at the damaged site. This example showed one biological aspect of this posttranslational modification in the regulation of DNA metabolism (Mijakovic, Petranovic et al. 2006). In addition, *B. subtilis* mutant with disrupted gene for tyrosine kinase (PtkA) showed inefficient DNA replication and arrest of replication complex (Petranovic, Michelsen et al. 2007).

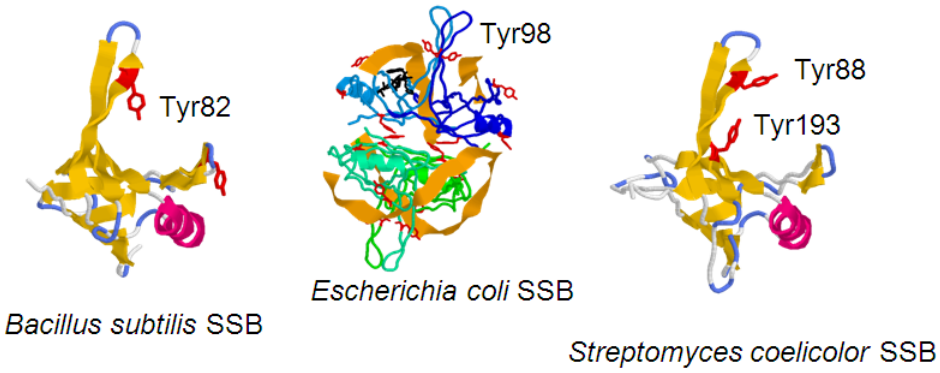


Figure 1.13. Three-dimensional modelling of *B. subtilis* and *S. coelicolor* SSBs strongly implies (A. Krisko and D. Vujaklija, unpublished data) that tyrosine residue 98 in SSB_{Eco} occupies nearly the same molecular space as tyrosine residue 82 in *B.* subtilis SSB and therefore is the most likely phosphorylation site in SSB_{Eco}. This is in agreement with the crystal structure data of the SSB_{Eco} tetramer which reveals that tyrosine residue 78 is placed inside the tetramer while tyrosine residue 98 is exposed on the surface.

1.2.3. SSB paralogues in Eubacteria

The processes of gene duplication generate gene copies that are identical to each other - such genes are called paralogues. Gene duplication and their subsequent divergence may play an important role in the evolution of novel gene functions (Innan and Kondrashov 2010). Few studies have shown that various bacterial orders possess paralogous *ssb* genes (Lindner, Nijland et al. 2004). The biological role of paralogous SSB proteins in bacteria is poorly studied. In *B. subtilis* this additional SSB has role in natural competence (Lindner, Nijland et al. 2004). This naturally transformable bacterium employs its second SSB (SsbB) for the DNA uptake, and its biological role correlates with *ssbB* expression, i.e. it has been shown that *ssbB* gene is expressed only in the late stationary phase in minimal media, during competence stage. Contrary to that, *ssbA* is mostly expressed during early stages of growth (Lindner, Nijland et al. 2004), when is most needed for DNA replication processes. In spite of their differential transcription profiles, recently the division of labour during *in vitro* genetic recombination between SsbA and SsbB has been shown in *B. subtilis.* These results indicate that these two proteins are co-players in some processes in the cell. SsbA from *B. subtilis* binds ssDNA with higher affinity than SsbB and co-assembles onto SsbB coated ssDNA, while both proteins inhibit ssDNA binding of RecA. RecO interaction with ssDNA-bound SsbA helps to

remove both SsbA and SsbB from the DNA more effectively than if the DNA is coated only with SsbA. Once RecA is nucleated onto ssDNA, RecA filament elongation displaces both SSBs and enables RecA - mediated DNA strand exchange (Yadav, Carrasco et al. 2012). *B. subtilis* SsbB-ssDNA complex structure has been solved (Yadav, Carrasco et al. 2012) and strongly resembles that of SSB_{Eco} and *Helicobacter pylori* SSB (Chan, Lee et al. 2009). For *Streptococcus pneumoniae*, bacterium from the same phylum (Firmicutes), it has been proposed that SsbB serves as a reservoir of ssDNA taken from the environment, allowing successive recombination cycles, and contributing to the genetic plasticity (Attaiech, Olivier et al. 2011). SsbB from *S. pneumoniae* directly protects internalized ssDNA. This SsbB is highly abundant in the cell during competence, potentially allowing the binding of ~1.15 Mb ssDNA (half a genome equivalent). The SsbBs from S*. pneumoniae* and *B. subtilis* participates in the same processes in the cell, eventhough these two proteins differ greatly in their lengths and in composition of the C-terminal regions. SsbB from *B. subtilis* lacks conserved acidic C-tail, and is much shorter. However, it has been shown truncated version (delta Ct) from S*. pneumoniae* still excerts the same proposed reservoir function (Attaiech, Olivier et al. 2011).

It has been reported that SsbA and SsbB from S. *pneumoniae* have different binding properties (Grove, Willcox et al. 2005). SsbB protein from *S. pneumoniae* differs from both SsbA and SSB_{Eco} both in its binding to short and long ssDNA. Contrary to *B. subtilis*, which SsbA binds ssDNA with a much greater affinity (>5-fold) over that of SsbB (Yadav, Carrasco et al. 2012), competition experiments indicate that the affinity of SsbB protein for dT_{50} is higher than that of either SsbA or SSB_{Eco} protein. However, the results of gel shifts experiments indicate that the two SSB_{Eco} or SsbA tetramers are able to bind to intermediate length oligomer, dT_{75}, while only one SsbB is able to bind to the same length of such ssDNA (Grove, Willcox et al. 2005). *S. pneumoniae* SsbA and SsbB binding to ssDNA has been compared using atomic force microscopy at saturated protein concentrations. ssDNA was completely covered by extended tracts of SsbA protein, similar to that of *E. coli*. In contrast, the complexes that were formed by SsbB proteins were highly condensed in appearance and contained numerous stem like projections that were not prominent in SsbA complexes. An analysis of the electron microscopy images revealed that the projected surface area of the SsbB-ssDNA complexes were 30 % smaller than that of the SsbA-ssDNA, indicating that there was less protein bound in the SsbB than in SsbA complexes. Therefore, the manner in which multiple SsbB proteins assemble onto ssDNA molecules differs from that observed for SsbA and SSB_{Eco} proteins.

It has been reported recently (Jain, Zweig et al. 2012) that SsbB from phylogenetically distant bacterium *Neisseria gonorrhoeae* (β division of Proteobacteria), although highly competemet, is not involved in DNA uptake from the environment. It was found that its SsbB strongly stimulates

topoisomerase I activity. This SsbB has a conserved N-terminal OB fold, as most SSBs. Both OB fold and the C-terminal regions share just a relatively low sequence similarity with other SSBs. This SsbB was expressed only at low levels under normal growth conditions. It has been suggested that SsbB under these growth conditions is either distributed evenly over exposed ssDNA stretches, or is specifically targeted to specific DNA regions by other proteins. This statment was partially confirmed by the atomic force microscopy showing that SsbB from *N. gonor*rhoeae binds ssDNA randomly at low SsbB/DNA ratios, as individual blobs. This might suggest that under such conditions SsbB binds preferentially to DNA regions without secondary structures, and is initially excluded from condensed regions. At higher ratios, DNA molecules are saturated, thereby resolving the condensed ssDNA structures. As observed for SSB_{Eco} the co-existence of different types of structures (semi-naked ssDNA, more or less saturated complexes) was observed in the same deposition. This indicates that there might be some limited cooperativity in their SsbB-ssDNA interaction, upon binding to longer ssDNA molecules. Binding of this SSB to ssDNA is not sensitive to salt or Mg^{2+} concentration. In a similar way as SSB_{Eco} it binds in highly cooperative manner to 35 nucleotides with two OB folds, while binding to 65 nucleotides at higher salt or Mg^{2+} concentrations uses four OB folds. This SSB cannot complement *E. coli ssb* mutation, as shown for many other SSB proteins encoded from chromosomes or plasmids (Jain, Zweig et al. 2012) .

1.2.4. Bioinformatic analyses of ssb genes and SSB proteins

The first bioinformatic analysis on bacterial genomes available at the time was reported in 2004 (Lindner, Nijland et al. 2004). 87 bacterial genomes were screened for *ssbA* homologues and their position. This analysis has shown (Lindner, Nijland et al. 2004) that *ssbA* gene in *B. subtilis* is positioned between genes for ribosomal proteins, *rpsF* and *rprR,* contrary to the position of *ssb* gene in *E. coli*, which is placed near *uvrA* gene. 15 genomes were found to have the same organisation as *ssb* gene from *E. coli*, while 35 genomes had positioning like *ssbA* from *B. subtilis.* Most bacteria within this group were Gram positive, and all of them had *ssb* paralogues in their genomes. The position of *ssbA* gene between ribosomal protein genes was also preserved in the representatives of Thermotogales and *Thermus/Deinococcus* group, and Gram-negative species from the phyla Spirochaetae, Aquificae and Chlorobi, and ε subdivision of Proteobacteria. Other bacteria such as Chlamydia, Cyanobacteria, and Fusobacteria did not display this type of *ssb* gene organisation (Lindner, Nijland et al. 2004). Untily now the phylogenetic studies on SSB proteins were limited. Few analyses showed that bacterial and mitochondrial SSBs are separated from the archeal SSBs/eukaryotic RPAs; and phage and bacterial SSBs do not form monopyletic groups since they are

intermixed with each other (Moreira 2000; Theobald and Wuttke 2005). Recently well defined clade of SSB proteins from virulent lactococcal phages was reported (Szczepankowska, Prestel et al. 2011). According to the phylogenetic analyses this group of SSB protein sequences has been transferred from Archea, although they are found in Eubacteria. Thus, present phylogenetic analyses support the hypothesis that ssDNA binding protein superfamily derived from the common ancestral domain present in the genome at the very base of the evolutionary tree. First phylogenetic analysis on paralogue SSB protein was published in 2012 (Jain, Zweig et al. 2012). A phylogenetic analysis of paralogous SSB protein was determined for SsbB from *N. gonorheae* and it was shown that this protein have close homologues in other Proteobacteria within conserved genetic cluster called genetic islands. This cluster contains genes encoding DNA-processing enzymes: ParA, ParB, and TopB topoisomerase and other conserved hypothetical proteins. The SSBs found in these clusters formed a family separated from other ssDNA binding proteins.

1.2.5. SSB proteins from Streptomyces coelicolor

Two paralogous genes encoding single-stranded DNA binding proteins (SSBs) were found in the core region of *S. coelicolor* chromosome. The structures of both SSB proteins from *S. coelicolor* have been determined (Stefanic, Vujaklija et al. 2009; Paradzik, Ivic et al. 2013). SsbA structure from *S. coelicolor* has been solved first and it has revealed variations previously described for SSB from pathogenic bacteria *Mycobacterium tuberculosis (Saikrishnan, Jeyakanthan et al. 2003).* Two subunits of tetramer are connected with additional β clamps which contribute to protein stability (Figure 1.14), while 3D structure has ellipsoid shape contrary to spheroid -like shape, described for SSB_{Eco} (Raghunathan, Ricard et al. 1997). These structure variations are probably a result of adaptation to high GC content of these bacteria (Saikrishnan, Manjunath et al. 2005; Stefanic, Vujaklija et al. 2009). Recently 3D structure of paralogous SSB protein (SsbB) from *S. coelicolor* has been also solved by our group (Paradzik, Ivic et al. 2013). SsbB from *S. coelicolor* lacks β clamps present in SsbA structure, but it has disulphide bridges connecting subunits (2 S-S bonds at dimer interface, A/B and C/D subunits). Calculations of free energies of tetramer dissociation (ΔG^{diss}) indicate that this structure could be the most stable among so far described SSB proteins (Paradzik, Ivic et al. 2013). This is the first example of solved crystal structures of two paralogous SSB proteins from the same organism with the reported disupphide bonds that connects dimmers of SsbB. Interesintgly, the crystal structure of dimeric PriB protein from *E. coli* (Liu, Chang et al. 2004) revealed presence of disulphide bridges between subunits. According to the literature (Ponomarev, Makarova et al. 2003) PriB protein also evolved from SSB in Proteobacteria.

In conclusion, until this work the studies on SSB proteins from model bacterium *S. coelicolor* were limited, and the biological roles of paralogous SSBs in streptomycetes have not been reported.

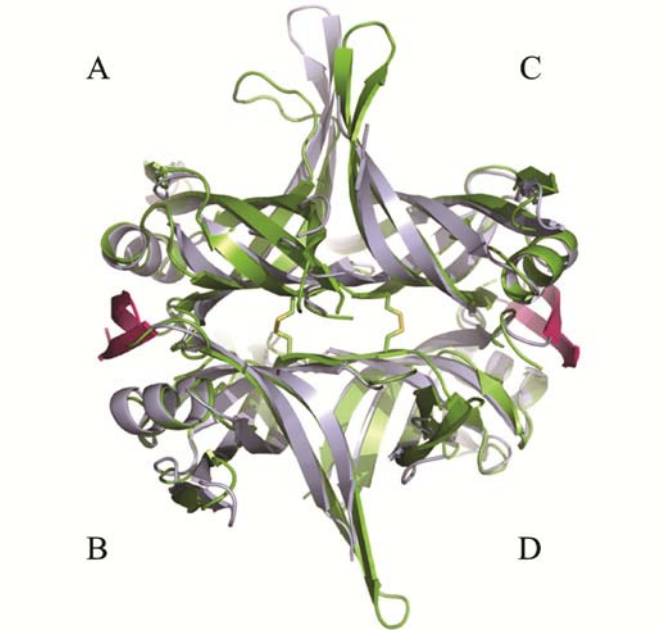


Figure 1.14. A superposition of SsbB (green) and SsbA (grey) structures. The monomers are designated A, B, C and D as shown. 3D structure of SSB proteins displayed similar structure with some unique variations. Disulphide bridges of SsbB are shown as yellow sticks, while the clamp structure of SsbA is in pink colour.

1.3. Aims

- The biological roles of paralogous SSBs in streptomycetes have not been reported until this study. Therefore, the main goal of this thesis is to investigate the cellular roles of paralogous SSB proteins from model bacterium *S. coelicolor*, starting with hypothesis that only one of them is essential for bacterial growth.
- Characterization of both SSB proteins will be performed using different molecular, biochemical and biophysical methods. This includes determination of transcriptional profiles during growth, localisation in *S. coelicolor* mycelia and analysis of *in vitro* activity, i.e. ssDNA binding affinities.
- Finaly, extensive bioinformatics and phylogenetic analyses on all available bacterial SSB protein sequences will be performed since the presence of paralogous SSB proteins in Eubacteria has been poorly studied until now.

2. Materials and methods

2.1. Materials

Laboratory chemicals used in this study were purchased from Kemika and Sigma-Aldrich companies. Commonly used buffers are listed in Table 2.1, molecular markers in Table 2.2, commercial kits in Table 2.3 and enzymes in Table 2.4.

Table 2.1. Buffers and solutions

Buffer	**Contents**
TAE 1x	TRIS–acetate, 1 mM EDTA, 20 mM Na–acetate (pH 8.3)
6 × SB (agarose gel electrophoresis)	0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 30 % glycerol
Protein lysis buffer	10 mM TRIS-HCl. pH 7.5, 1 mM EDTA, 1 mM DTT, 10 % glycerol
Buffer A	100 mM TRIS-HCl, pH 7.5; 300 mM NaCl; 10% glycerol
Buffer B	Buffer A+ 30 mM imdazole
Buffer C	Buffer A+300 mM imidazole
SDS-PAGE buffer 10x	30.3 g L^{-1} TRIS base, 144 g L^{-1} glicine, 10 g L^{-1} SDS
SDS-PAGE sample buffer 4 x	8 % SDS, 400 mM DTT, 200 mM TRIS-HCl pH 6.8, 40 % glycerol, 0.4 % bromophenol blue
Transfer buffer	25 mM TRIS–HCl, 192 mM glycine, 20 % methanol
TBS T 1 x	50 mM TRIS (pH 7.5), 150 mM NaCl, 0.1 % Tween–20
PBS 1 x	NaCl 8 g L^{-1}, KCl 0.2 g L^{-1}, Na_2HPO_4 1.44 g L^{-1}, KH_2PO_4, 0.24 g L^{-1}, pH 7.4
T + PEG buffer	25 mL sucrose (10.3 %), 75 mL MQ H_2O, 0.2 mL trace element solution, 1 mL K_2SO_4 (2,5 %), with the addition 0.2 mL 5M $CaCl_2$ and 0.5 mL Tris–maleic acid pH 8.0; 1 g PEG 1500 dissolved in 3 mL of sterile T buffer

Table 2.2. Molecular markers

	Trade name	**Manufacturer**
DNA	MassRuler DNA Ladder, Mix	Fermentas
protein	Precision Plus Protein Standard Unstained	BioRad
protein	Precision Plus Protein Standard: Prestained Dual Colour	BioRad

Table 2.3. Enzymes

Enzyme	Manufacturer
Restriction enzymes: BamHI, EcoRI, HindIII, SpeI, EcoRV, PstI	Fermentas
Klenow enzyme	Roche
T4 DNA ligase	Fermentas
Lyzozime	Sigma
RNase A	Sigma
Dnase I	Promega

Table 2.4. Commercial kits

Kit	Manufacturer
PCRquick-spin™ PCR Product Purification Kit	iNtRON Biotehnology
DNA-spin™ Plasmid DNA Extraction Kit	iNtRON Biotehnology
MEGA-spin™ Agarose Gel Extraction Kit	iNtRON Biotehnology
GEM®-T Easy Vector Systems	Promega
FastRNA®Problue Kit	Qbiogene, Inc, CA
High capacity cDNA Reverse transcription kit	Applied biosystems
5´ RACE System for Rapid Amplification of cDNA Ends	Invitrogen
Quick Start Bradford Protein Assay Kit	Biorad
SlowFade ® Antifade Kit	Initrogen

Cloning vectors used in this study are listed in Table 2.5, whereas primers are listed in Table 2.6.

Table 2.5. Plasmids and cosmids

Vector	Characteristics	Source
pQE-30	*Amp*^{R}	Qiagen
pGEM-T	*Amp*^{R}	Promega
pRSET-B mCherry	*Amp*^{R}	(Shaner, Campbell et al. 2004)
C607.1.c02.EZR1.seq.	Cosmid with transposon Tn*5062* in *ssbB* gene at the position 109 from the start codon, *Neo*^{R}, *Amp*^{R}, *Apr*^{R}	(Fernandez-Martinez, Del Sol et al. 2011)
SCH24	*Streptomyces coelicolor* cosmid, *Neo*^{R}, *Amp*^{R}	John Innes Centre
h24.EO1_05061309TW.seq	Cosmid with transposon Tn*5062* in *ssbA* gene at the position 372 from the start codon *Neo*^{R}, *Amp*^{R}, *Apra*^{R}	(Fernandez-Martinez, Del Sol et al. 2011)
cSA01	SCH24Δ*ssbA*, *Neo*^{R}, *Amp*^{R}, *Apra*^{R}	This study
cSA02	SCH24Δ*ssbA*::scar, *Neo*^{R}, *Amp*^{R},	This study
pIJ773	P1-FRT-oriT-aac(3)IV-FRT-P2	(Gust, Challis et al. 2003)
pMS82b	Integrative vector for Streptomyces, ΦBT1 *attP*-int, Hyg^R	(Gregory, Till et al. 2003)
pSET152	Integrative vector for Streptomyces, Φ C31 *attP*-int, Apra^R	(Bierman, Logan et al. 1992)
pANT849	Expression vector for Streptomyces, *Thio*^{R}	(DeSanti and Strohl 2003)
pMS-pssbA	pMS82 (containing *ssbA* gene and its promoter, *Hyg*^{R})	This study
pMS-pGFPssbA	pMS82 with promoter-eGFP-SsbA fusion *Hyg*^{R}	This study
pMS-pssbB	pMS82 (containing *ssbB* gene and its promoter, *Hyg*^{R})	This study
pSET-pmCherryssbB	pSET152 (containing mCherryssbB construct and *ssbB* promoter, Apra^R)	This study
pANTssbB	Isolation of His-tagged *SsbB* protein from *S. coelicolor, Thio*^{R}	This study
pANTmCherryssbB	*Thio*^{R}	This study
pMS-pssbBΔC	pMS82(containing *ssbB* ΔC gene and *ssbB* gene promoter, *Hyg*^{R})	This study

Table 2.6. Primers

Primer	Sequence 5'-3'	Purpose
SC-AF SC-AR	CGGGATCCATGGCAGGCGAGACCGTCATCACGGTC TTTTCTGCAGTCAGAAGGGGGGCTCGTCCGAGTAGCCGCC	*ssbA* cloning into pQE vector
F-L3pANT R-L3pANT	CGGAATTCGAGAGGTTCCCCATGAGAGGATCGCATCACCATCACCATCAC AATTAAGCTTGGCTGCAGTCAGAAGGGGGG	*ssbA* subcloning into pGEM vector
SSBAko F SSBAkoR	AGCAGCCAGAGCAGCAAACCCGCCGAGAGGTTCCCCATGATTCCGGGGA TCCGTCGACC CCTGTGTGATCAAGAAGTGTGGGTAGAGCCCGTCCCTTATGTAGGCTGG AGCTGCTTC	Amplification of pIJ773 cassette for *ssbA* gene disruption
GFPpromF GFPpromR	ATGGTGAGCAAGGGCGAGGAG CATGCCCGCCCTTGTACAGCTCG	eGFP cloning in front of *ssbA* gene
SSBApromF SSBApromR	CGGAATTCCTCGGGGCCGGATTTCAAC CTGGATCCCCTCTCTTTGGACTCAACGG	*ssbA* promoter cloning in front of *ssbA*
SC-BF SC-BR	CGGGATCCATGAACGAGACGATGATCTGCGCGGTGGGG TTTTCTGCAGTCATCCCACCGGGACGGGGTCCGGCCG	*ssbB* cloning into pQE vector
S6pANTF S6pANTR	CGGAATTCGAGAGGTTCCCCATGAGAGGATCGCATCACCATCACCATCAC AATTAAGCTTGGCTGCAGTCATCCCACCGGGACGG	*ssbB* subcloning into pGEM vector
SSB119 R	GGTCTGCAGTTACCTGGCGGTGCGCCGGAACGC	*sssB*Δ30 cloning into pQE vector
mcherryF mcherryR	GAGAATTCGAGAGGTTCCCCATGGTGAGCAAGGGCGAGG GGCATGGACGAGCTGTACAAGGGCGGGGCTGGATCCAG	mCherry cloning in front of *ssbB* gene in pGEM vector
pEcoRIssbbF pBamHIssbbR	CCTGAATTCGCTGCCACCCGTGCGGGTCG CTGGGATCCTCTCCTCCCTCCCGCACGGGTC	*ssb*B promoter cloning in front of ssbb gene
promssbbF promssbbR	GATATCGCTGCCACCCGTGCGGGTCG CTCCTCCCTCCCGCACGGGTC	*ssbB* promoter cloning in front of mcherrySsbB construct
SSBAGSP1 SSBAGSP2 RPR RPR2 RPR3 SSBAF SSBAR SSBAFRT SSBAR GSP1 izaSSBA	GGTCAGGAAGAGGCTTTCG CGACCGTGATGACGGTCTCG GCATCTCGGGACGGAGGACC CGACCTTCTCGACCTTTCC GGATGACCATCACCTCGTAG CGG TCGTCGGCAATCTTGTC, CGAGGTCTTGGTGACCTTGG GCGTCACTACGAGGTGATGG, CACCGCCGCCGTAACC TCTCCTGTGTGATCAAGAAG	Expression analyses of *ssbA*/RACE
SSBBGSP1 SSBBGSP2 SSBBGSP3 SSBBFRT	CGAGATCGTGGCCGATCG CCGAGGTCCGGCTCTGCC CGCACCTTCAGCCTGCC GGACTCGCGATGAACGAGACG	Expression analyses of *ssbB*/RACE
pMS82seqF T7f, pUCM13R EZR2, EZL1 SCO4848F SCO4848R SCO4848int AAP	TGGCCTTGAAATCGTTAGT, TAATACGACTCACTATAGGG, CAGGAAACAGCTATGAC, TCCAGCTCGACCAGGATG, ATGCGCTCCATCAAGAAGAG GTCTCCTGGTTCCTGCTCG, CGGCGCAGTGCGCGCACGC GCGACATGCCTAGATCTTGG GGCCACGCGTCGACTACGGGIIGGGIIGGGIIG	Sequencing primers Tn*5062* specific primers Integration into att site RACE-primer Invitrogen

EcoRI, BamHI, HindIII, EcoRV, PstI

2.1.1. Microorganisms

Bacterial strains used during the course of this study are listed in Table 2.7.

Table 2.7. Bacterial strains

Strain	Genotype	Source
E. coli XL1	endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB$^+$ lacI^q Δ(lacZ)M15] hsdR17(r_K^- m_K^+)	New England Biolabs
E. coli NM522	*sup*E, *thi*, Δ(*lac-pro*AB), *hsd*5 (r^-, m^-), [F′, *pro*AB, *lac*I^{q}ZΔM15]	New England Biolabs
E. coli ET12567(pUZ8002)	*dam13*::Tn*9(Chl^R) dcm-6 hsdM hsdR recF143 zjj-201*::Tn*10 galK2 galT22 ara14 lacY1 xyl-5 leuB6 thi-1 tonA31 rpsL136 hisG4 tsx-78 mtlI glnV44*, pUZ8002(*Kan*^{R})	(MacNeil, Gewain et al. 1992)
E. coli BW25113(pIJ790)	ΔaraBAD, ΔrhaBAD, pIJ790 -λ*RED (*gam, bet, exo*),* cat, araC*, rep101^{ts}*	(Gust, Challis et al. 2003)
E. coli BT340	DH5α, pCP20	(Cherepanov and Wackernagel 1995)
S. coelicolor M145	SCP1$^-$, SCP2$^-$	(Kieser 2000)
TSA01	M145ΔssbA/ssbA+ (ΦBT1 *attP*) , *Hyg^R*	This study
TSA02	M145 peGFPssbA (ΦBT1 *attP)*, *Hyg^R*	This study
TSB01	M145 *ssbB*::Tn*5062*, *Apra^R*	This study
TSB02	M145 *ssbB*::Tn*5062/ ssbB+*, *Apra^R*, pMS82 *ssbB* at ΦBT1 *attP* site, *Hyg^R*	This study
TSB03	M145 *ssbB*::Tn*5062/ ssbB+*, *Apra^R*, pMS82 *ssbBΔC* at ΦBT1 *attP*, *Hyg^R*	This study
TSB04	M145 pmCherryssbB (Φ C31 attP), *Apra^R*	This study
TSB05	M145, pANT849 ssbB, *Thio^R*	This study
TSAB1	M145 pmCherryssbA, peGFPssbA, *Apra^R, Hyg^R*	This study
TSAB02	M145 peGFPssbA, pANT849 mCherryssbB, *Hyg^R*, *Thio^R*	This study

2.1.2. Bacterial growth conditions

E. coli strains: NM522, ET12567 and XL 1-blue, were grown in the standard Luria-Bertani (LB) medium at 250 rpm and 37 °C, strains BW25113 (pIJ790) and BT340 at 30 °C until induction of temperature sensitive genes. Liquid LB was supplemented with 100 μg mL^{-1} ampicillin or 50 μg mL^{-1} kanamycin, and 25 μg mL^{-1} chloramphenicol, to maintain selection pressure. *Streptomyces coelicolor* M145 was grown at 250 rpm and 30 °C in media described in Table 2.8. The antibiotics were added at following concentrations: 50 μg mL^{-1} apramycin, 50 μg mL^{-1} kanamycin, 25 μg mL^{-1} nalidixic acid or 50 μg mL^{-1} hygromycin, when necessary.

Table 2.8. Growth media used for culturing *S. coelicolor*

Medium	Complete name/state	Reference	Purpose
CRM	Complete regeneration medium/Liquid	(Bielen, Cetkovic et al. 2009)	RNA isolation
MM	Minimal medium/solid or liquid	(Kieser 2000)	RNA isolation, microscopy
R5	Regeneration medium/solid	(Kieser 2000)	Protoplast regeneration, microscopy
TSB	Triptic soy broth/solid	Difco™	Fast non-sporulating growth
MS	Mannitol soya flour/solid	(Kieser 2000)	Sporulation, microscopy

2.2. Methods

2.2.1. DNA purification

Plasmid DNA was isolated from *E. coli* using QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions. DNA from solution or agarose gels was purified with appropriate kits (Table 2.3). Cosmid DNA was purified by alkaline lysis followed by phenol/chloroform extraction following described method (http://streptomyces.org.uk/redirect/RecombineeringFEMSMP-2006-5.pdf), (Gust, Chandra et al. 2004). Genomic DNA from *S. coelicolor* was isolated applying procedure as described (Kieser 2000).

2.2.2. Cloning

Unless otherwise stated, genes were cloned according to the standard procedures (Sambrook 1989). All constructs were generated following a multi-step cloning procedure:

1. PCR amplification of the desired fragments → **2.** Purification of the PCR products using a commercial PCR purification kit → **3**. Restriction enzyme digestion of the vector and DNA fragments → **4.** Separation of DNA fragments by agarose gel electrophoresis and purification of selected DNA bands by gel extraction kit → **5**. Ligation of purified DNA fragments → **6**. Transformation of competent bacterial cells → **7.** Screening of transformants with desired DNA insert using colony PCR method→ **8.** Purification of recombinant vector DNA from selected clone(s) → **9**. Verification of insert by DNA sequencing. Details of each step are described further in the text.

2.2.3. Polimerase Chain Reaction (PCR)

The setup for all PCR reactions were performed on ice in a volume of 50 µL (for cloning) or 25 µL (for colony PCR) and with different amount of DNA, in a volume from 0.5 to 1 µL. All components of different PCR reaction mixtures were listed in Table 2.9. PCR programmes were set up according to manufacturer instructions; extension time was adjusted according to the expected length of PCR product and to the annealing temperature of the each pair of primers (in dependence of their length and GC composition). Number of cycles was typically 30. PCR reactions were used for several purposes:

- Cloning; target genes or promoter regions were amplified from the genomic DNA (Table 2.9, mixture C)
- Screening of the transformants on LB agar plates to confirm the presence of specific inserts within the recombinant DNA constructs and within single bacterial colonies. PCR reactions were prepared using colony biomass (Table 2.9, Mixture B).
- Amplification of resistance cassette (REDIRECT PCR-targeting technology, section 2.2.8). Mixture A (Table 2.9) was used for this reaction. Cycling parameters were adjusted according to original protocol (Gust, Chandra et al. 2004):

1. Denaturation	94 °C, 2 min	
2. Denaturation	94 °C, 45 sec	10 cycles
3. Primer annealing	50 °C, 45 sec	
4. Extension	72 °C, 90 sec	
5. Denaturation	94 °C, 45 sec	15 cycles
6. Primer annealing	55 °C, 45 sec	
7. Extension	72 °C, 90 sec	
8. Final extension:	72 °C, 5 min	

PCR mixture for REDIRECT technology contained 100 ng of template DNA.

- Specific amplification of complementary DNA (cDNA) using mixture B (Table 2.9). Template for this reaction was total cDNA obtained with reverse transcriptase from total RNA (see section 2.2.10, RT-PCR).

PCR products were analyzed using 1 % agarose gel electrophoresis. When needed, PCR product was purified using PCR purification kit.

Table 2.9. PCR reaction (A, B or C) components

Mixture	Polymerase	Ammount of enzyme	dNTP	Primers	Buffer	DMSO
A	LA Taq (TAKARA)	1.25 U	400 µM	0.5 µM	GC II	2%
B	ReadyMix™ Taq (Sigma-Aldrich)	-	-	0.5 µM	-	-
C	*PfuUltra* (Stratagene)	0.6 U	200 µM	0.2 µM	1x Pfu buffer	4 %

2.2.4. Restriction digestion of DNA

The digestion of DNA was performed in the final volume of 15 - 30 µl in recommended reaction buffer and temperature for up to 3 h according to manufacturer's instructions. Double digestions were carried out simultaneously in a buffer compatible for both enzymes. If only specific bands or restriction products were required, DNA was separated on agarose gel and extraction of the desired bands was performed by the excision and subsequent purification using *DNA-spin™ Plasmid DNA Extraction Kit* (iNtRON Biotehnology).

2.2.5. Ligations

Vectors and DNA fragments selected for cloning were digested with appropriate restriction enzymes (section 2.2.4). DNA fragments were obtained either by PCR amplification or in subcloning experiments, by digestion of existing recombinant plasmid. Ligation reactions were performed with T4 DNA ligase (Fermentas). Reaction mixture and conditions for either protruding or blunt end ligations are described in Table 2.10. Typically, ligation was performed in 15 µL. Concentrations of DNA fragments used in ligations were determined by agarose gel electrophoresis and by comparison to DNA markers of known size and concentration (MassRuler DNA Ladder, Fermentas).

Table 2.10. Ligation mixtures

Type of ligation	Vector:insert ratio	Amount of vector	T4 DNA ligase/ Weiss units	Incubation	PEG4000
Sticky end	1:3	50 ng	5	16 °C, ON	/
Blunt end	1:6	100-150 ng	10	8 °C, ON	10 %

TA cloning of PCR products was performed using commercial GEM®-T Easy Vector Systems kit (Promega). Commercially obtained linearized pGEM-T vector has 3'- thymidine overhang timidins,

and it can easily be ligated to PCR products because *Taq* polymerases add 5' –adenine overgang to the ends of amplified product. Thermostable DNA polymerases with proofreading (3′-->5′ exonuclease) activity, such as *Pfu* DNA polymerase, generate blunt-ended fragments during PCR amplification. For efficient TA cloning PCR products need to be 3' adenylated and purified. Therefore PCR fragments generated using these polymerases had to be modified with dATP using TaqDNA polymerase before ligation to the vectors. A-tailing reaction mixture typically contained 100 - 400 ng of PCR product, 1 U *Taq* DNA polymerase (Fermentas), 1 x reaction buffer, 4 mM $MgSO_4$ and 0.2 mM dATP. Mixture was incubated 30 min at 72 °C.

2.2.6. Plasmid construction

Plasmid constructs were prepared according to general scheme in the section 2.2.1.
Cloning of almost all constructs was performed in a similar way using three vectors as follows:

1. Genes amplified from the genomic DNA were cloned into the expression vector pQE.

2. Genes were subcloned into sequencing vector pGEM-T. pGEM-T constructs were treated to add regulatory regions, or fluorescent labels to the 5' end of the genes.

3. Prepared constructs were ligated to pMS82b integrative vector for *Streptomyces*. Only pmCherrySsbB construct was ligated to pSET152 integrative vector, as described below.

Preparation of *ssbA* constructs is shown in Figure 2.1. The *ssbA* gene from *Streptomyces coelicolor* (NCBI, Gene ID 1099343) was previously cloned as described (Mijakovic, Petranovic et al. 2006). This construct was used for over-expression of *ssbA* gene, and also as a starting vector for preparation of all other recombinant *ssbA* genes used in this study. In order to obtain *ssb*A gene with its upstream regulatory elements, this gene was amplified by PCR (reaction C, Table 2.9) with F-L3pANT and R-L3pANT primers (Table 2.6) and subcloned into the pGEM-T following TA cloning procedure (Section 2.2.5). Further, *ssbA* promoter region was PCR amplified (reaction C, Table 2.9) from the *S. coelicolor* genomic DNA. pGEM-*ssbA* construct and *ssbA* promoter region were digested with EcoRI and BamHI restriction enzymes and ligated. Reconstructed DNA sequence containing *ssbA* gene and its upstream regulatory elements (promoter and ribosomal binding site; pssbA) was named pGEM-pssbA, as shown in Figure 2.1. This plasmid was digested with HindIII, treated with Klenow fragment to obtain blunt ends and digested with SpeI enzyme. DNA fragment carrying pssbA sequence was further subcloned into EcoRV and SpeI sites of pMS82b, integrative vector for streptomycetes. This construct, pMS-pssbA was used for the complementation experiment. For the preparation of fluorescently labelled *ssbA,* eGFP was PCR amplified (reaction C, Table 2.9) from pIJ786 (Table 2.5)

and ligated into BamHI site of pGEM-pSSBA construct. Promoter-eGFP*ssbA* fusion was subloned to pMS82 vector (Figure 2.1) using the same approach as described for pssbA construct.

Scheme showing the preparation of *ssbB* constructs is presented in Figure 2.2. The *ssbB* gene from *S. coelicolor* (NCBI, Gene ID: 1098117) was cloned into pQE-30 and designated pQE-ssbB, using the same approach as described for *ssbA* gene (Mijakovic, Petranovic et al. 2006). To obtain genetic construct with DNA sequence comprising of promoter region and *ssbB* gene, cloning was performed in the same way as for the *ssbA* gene. Construct, pMS-pssbB was used for complementation experiment. The *ssbBΔC* gene, lacking the last 37 amino acids at the C-terminus was obtained by PCR using SC-BF and SSB119 R primers (Table 2.6) and *S. coelicolor* genomic DNA as a template. The amplified DNA fragment was cloned into the vector pQE as described for the *ssbA* and *ssbB* genes. In the next step, *ssbB* from the pGEM-pssbB construct was replaced by *ssbBΔC* at the BamHI and HindIII restriction sites, following the same steps as described for pssbA and pssbB. pssbBΔC was further subcloned into pMS82, giving pMS-pssbBΔC.

mCherry labelling of the SsbB protein was performed as follows: (**i**) DNA sequence encoding mCherry was amplified by PCR from pRSET-B mCherry (reaction C, Table 2.9), (**ii**) DNA sequence encoding His-tag at the 5' end of *ssbB* gene was replaced with mCherry gene, and (**iii**) DNA fragment carrying promoter and SD sequence of *ssbB* gene was ligated in front of mCherry-*ssbB* sequence. Recombinant gene (mCherry*ssbB*) with promoter was further subcloned into pSET152, giving pSET-pmCherryssbB (Figure 2.2).

For the over-expression of the His-tagged *ssbB* gene in *S. coelicolor,* this gene was subcloned into pGEM-T using S6pANTF and S6pANTR primers (Table 2.6). This construct, pGEM-ssbB, that contained the original *ssbB* ribosomal binding site, was digested with EcoRI and HindIII and ligated into pANT849 generating pANT-ssbB that was used for the transformation of *S. coelicolor* protoplasts. For over-expression of mCherryssbB fusion protein in *S. coelicolor*, His-tag sequence in pGEM-ssbB was replaced with mCherry sequence using EcoRI and BamHI restriction sites. Further cloning into pANT849 was performed as described for pANT-ssbB, giving pANT-mCherryssbB. All constructs were verified by sequencing in-house with automatic sequence analyzer "ABI PRISM® 3100-Avant Genetic Analyser" (Applied Biosystem) using "ABI PRISM Big Dye Terminator v3.1 Ready Reaction Cycle Sequencing Kit.

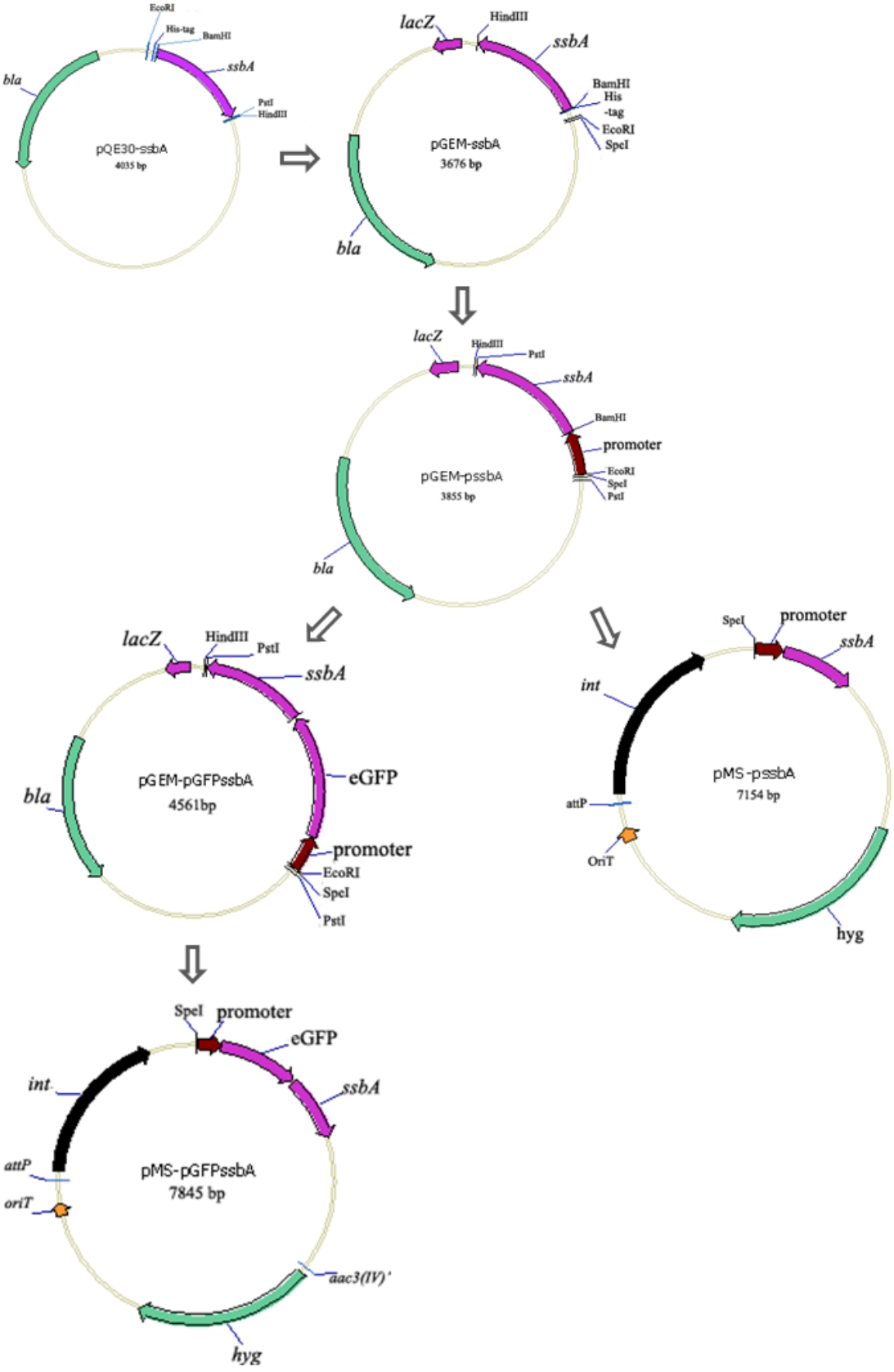


Figure 2.1. Flow chart showing steps in making different *ssbA gene* constructs

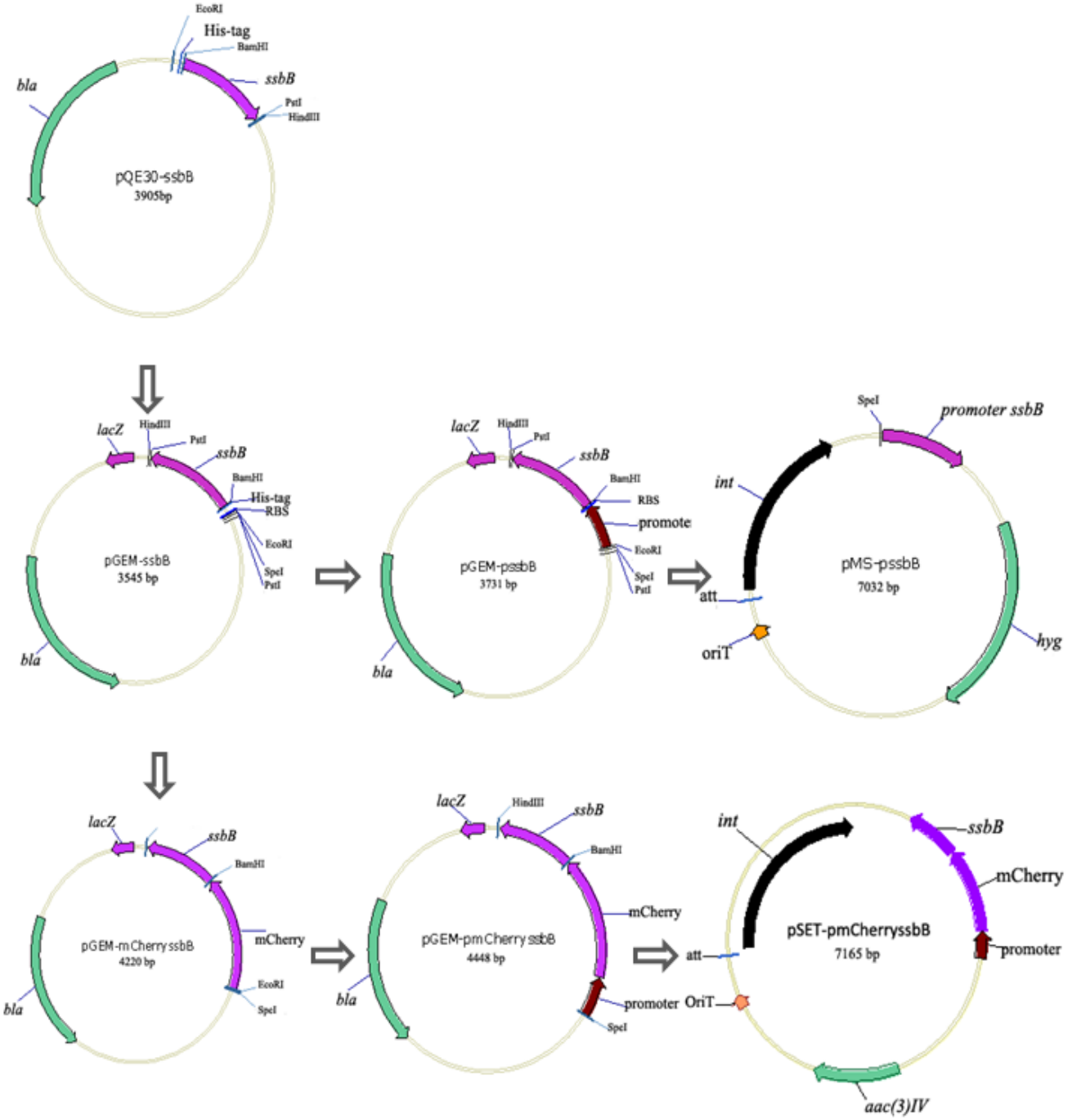


Figure 2.2. Flow chart showing steps in making different *ssbB* gene constructs

2.2.7. Transformation of bacterial cells

E. coli electrocompetent cells were transformed using "*Electroporator 2510*" (Eppendorf) and 2 mm electroporation cuvettets (BioRad). High voltage pulse (2.5 kV, 400 Ω and 25 µF) was applied to 40 µL of bacterial cell suspension mixed with DNA (plasmid, cosmid or ligation mixture) resuspended in a buffer of low ionic strength. Preparation of competent cells was performed as described (Sambrook 1989). Transformed cells were regenerated for 1 h at 37 °C in 1 mL of LB medium and then plated on LB agar with appropriate antibiotics and incubated over-night (ON) at 37 °C. Transformation of *S. coelicolor* was prepared applying two methods:

1. Protoplast transformation was performed as described by (Bielen and Vujaklija 2007). Ligation mixture was added to 50 µL of protoplasts and mixed immediately by tapping the tube. T + PEG buffer (200 µL) was added to a tube and gently mixed by pippeting up and down four times. Such suspension was spread on R5 regeneration plates and incubated ON at 30 °C. Next day, the plates were overlaid with appropriate antibiotic and incubated for three days.
2. Conjugation with *E. coli*, was performed as described in section 2.2.8.

2.2.8. Gene disruptions and complementation

Gene disruption experiments were performed using recently developed transposon insertion single-gene knockout cosmid library (Fernandez-Martinez, Del Sol et al. 2011). This procedure included three steps:

1. Transformation of electrocompetent *E. coli* ET12567 strain by cosmid containing transposon inserted into a gene of interest
2. Transfer of cosmid to recipient strain (*S. coelicolor* M145) by conjugation
3. Selection of exoconjugants in which double crossing-over occurred (recombination between cosmid and genome) with appropriate antibiotics

Cosmids containing inserted transposon inside *ssbA* (h24-1.E01_05061309TW.seq; at 124 aa) gene were introduced by transformation into ET12567 non-methylating *E. coli* strain containing the nontransmissible *oriT* mobilizing plasmid, pUZ8002. Following previously described method (Gust, Chandra et al. 2004), *E. coli* cells from mid-log phase were mixed with 10^8 *S. coelicolor* heat-shocked spores and plated on MS agar. After 20 hours nalidixic acid was poured onto plates to selectively inhibit DNA synthesis of *E. coli.* Exconjugants were selected for apramycin resistance and kanamycin sensitivity after one generation of growth without kanamycin.

REDIRECT PCR-targeting technology was used to replace the entire coding region of *ssbA* gene with an apramycin (aac(3)IV gene) resistance cassette, following the generation of the in-frame deletion of the *ssbA* gene (Figure 2.3).

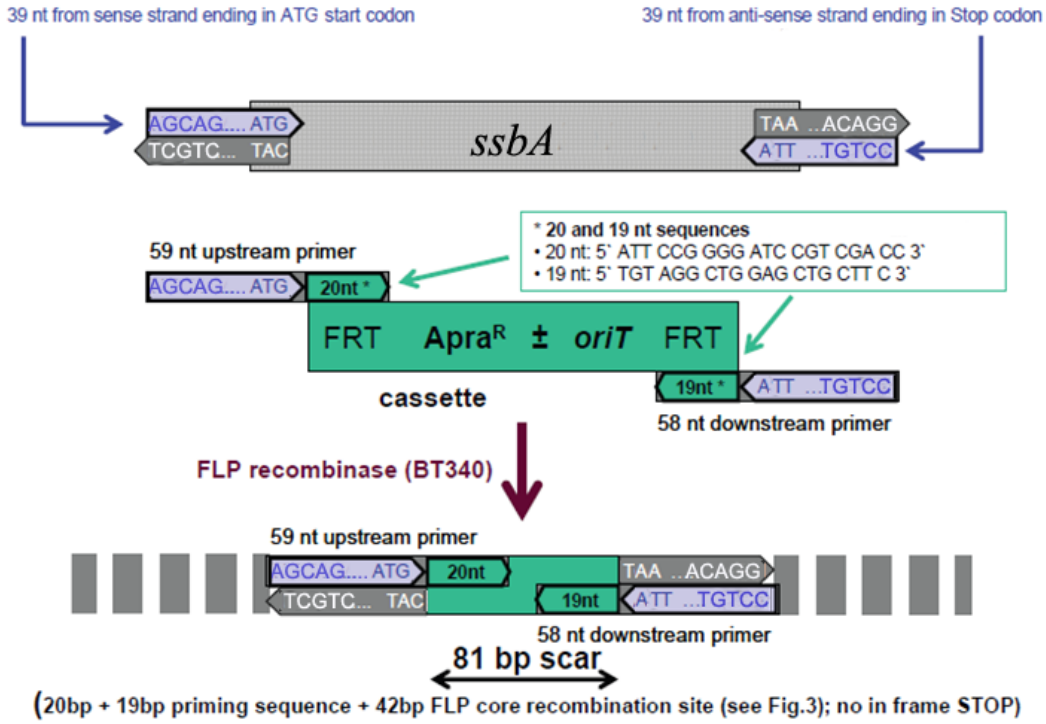


Figure 2.3. Scheme showing steps in generating an in-frame deletion of *ssbA* gene. Selectable marker is amplified by PCR using primers with 39 nt homology extensions (primers SSBAkoF and SSBAkoR, Table 2.6) These extensions allowed recombination with homologous regions in SCH24 cosmid. The inclusion of *oriT* in the disruption cassette allows conjugation to be used to introduce the modified cosmid DNA into *Streptomyces coelicolor (*adapted from http://streptomyces.org.uk/redirect/RecombineeringFEMSMP-2006-5.pdf *)*

Cosmid SCH24 that contains original *ssbA* gene was mutagenised in *E. coli* following the REDIRECT procedure schematically depicted in the Figures 2.3 and 2.4. Mutagenic cassette (FRT-oriTaac(3)IV-FRT) flanked by the recognition sequence of *E. coli* Flipase (FRT), contained the origin of transfer, *ori*T was PCR amplified using primers SSBAkoF and SSBAkoR (Table 2.6) and pIJ773 as a template. The PCR product was transformed to *E. coli* BW25113 (pIJ790) carryng recombinase gene and cosmid SCH24 (kanamycine resistance). Recombination between amplified resistance cassette and homologous regions on cosmid was induced with L-arabinose. The transformants with cosmid carrying cassette (i.e. disrupted *ssbA* gene) were selected with appropriate antibiotics as described (apramycin and kanamycin) (Gust, Chandra et al. 2004). After PCR and restriction verification, cosmid (cSA01, i.e. SCH24ΔssbA) was transformed into ET12567 *E. coli* for conjugation with *Streptomyces*. cSA01 was also used to generate in frame "scar" mutation at the position of the resistance cassette. This was

achieved by introducing cSA01 in *E. coli* BT340 strain (harbouring the Flipase gene, FLP) in which the temperature induced recombination between both FRT cassette-flanking regions takes place. Thus after recombination, cosmid labeled cSA02 had only 81 base pairs ("scar") left in the frame with the adjacent ORF and kanamycine resistance. This construct was prepared to avoid the effect of polar mutation to the downstream gene(s). Prepared cosmid was used to transform protoplasts of *S. coelicolor* carrying additional copy of *ssbA* gene at the ΦBT1 *attP* site and an unimpaired *ssbA* locus. The transformants were overlaid with kanamycin to select for the cosmid integration into the chromosome by single crossing-over. Kan^R recombinants were restreaked without kanamycin. Obtained colonies were cultured again for one generation with and without antibiotics to identify loss of kanamycin resistance i.e. presence of the double crossing-over event. Selected colonies were checked for the presence of "scar" sequence instead of *ssbA* gene by PCR reaction using SSBAFRT and izaSSBA primers (Table 2.6). In the original protocol cosmid with the unmarked "scar" sequence was used for the replacement of the already mutagenised target gene on the chromosome (Gust, Chandra et al. 2004). In this study, cosmid with scar mutation was used to replace the undisrupted *ssbA* because the disruption cassette in place of *ssbA* would not give any viable colonies (Gust, Chandra et al. 2004). This procedure is schematized in Figure 2.5.

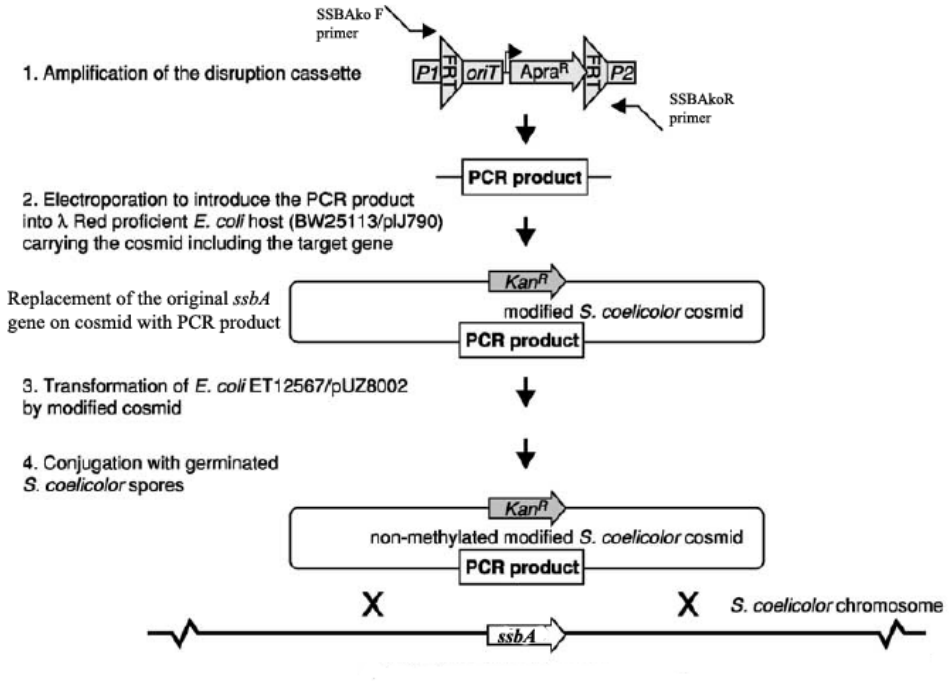


Figure 2.4. PCR targeting of *ssbA* gene. P1 and P2 represent the priming sites of the disruption cassette pIJ773. FRT sites (FLP recognition targets) flanking the disruption cassette allowed the elimination of the central part of the cassette to generate in-frame deletion. Adapted from (Gust, Chandra et al. 2004)

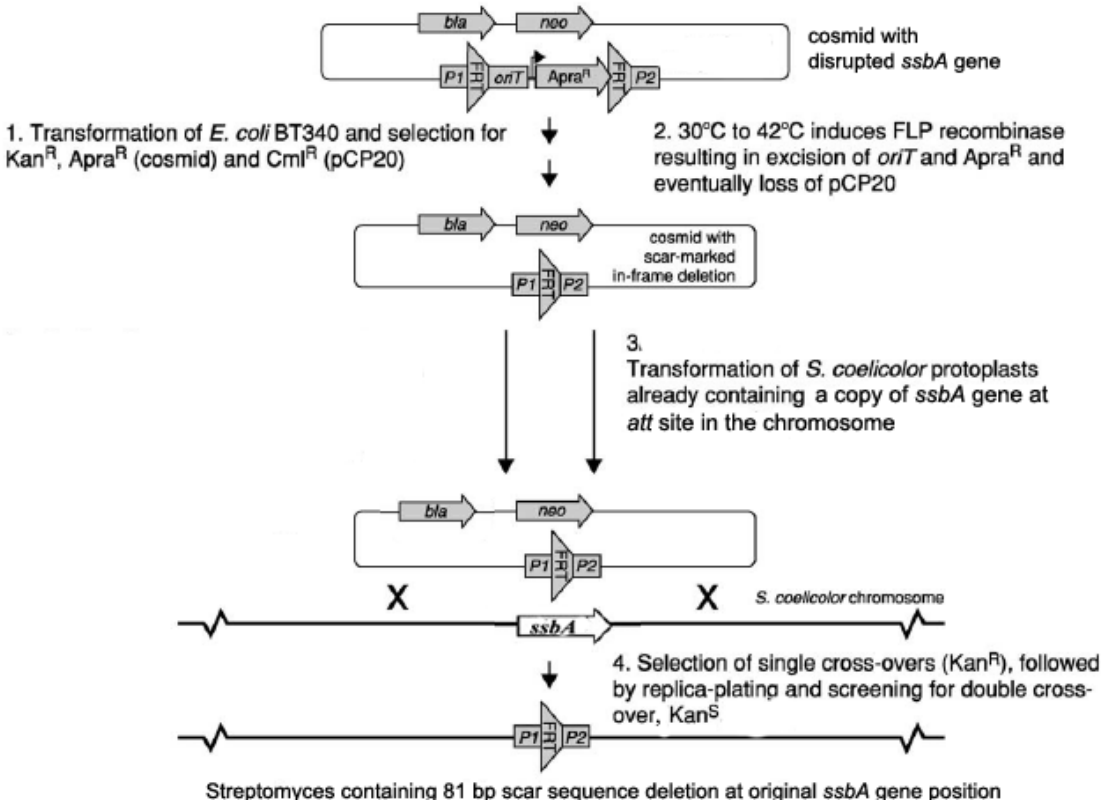


Figure 2.5. Steps in generating unmarked, in-frame deletion of *ssbA* gene in *Streptomyces* by site-specific recombination with the FLP recombinase. Adapted from (Gust, Chandra et al. 2004).

Cosmids containing transposon Tn*5062* in *ssbB* (6C07.1.c02.EZR1.seq, at 53rd aa) were conjugated into *S. coelicolor* as described for *ssbA*. The apramycin resistant and kanamycin sensitive colonies of *S. coelicolor* exconjugants were selected and double crossing over recombinants were confirmed by PCR using primers specific for the transposon cassette (EZR2 and EZL1) and *ssbB* (Table 2.6). Disrupted *ssbB* gene was complemented by introducing integrative vector, pMS82 carrying pssbB construct (pMS-pssbB), or pssbBΔC. The plasmid was firstly transformed into ET12567 *E.coli* to overcome restriction system and transferred by conjugation into *S. coelicolor.* After 20 hours hygromycin was poured onto the plates to select for exoconjugants with the sequence pssbB integrated into the genome. PCR with specific primers was used to confirm the plasmid integration into *S. coelicolor* chromosome at the ΦBT1 *attP* site.

2.2.9. RNA isolation

Total RNA was isolated from *Streptomyces coelicolor* M145 grown in 50 ml of CRM or minimal medium at 30°C and 250 rpm. Sampling (3 mL) was performed at 18, 24, 48 and 96 hours of bacterial

growth followed by immediate RNA extraction with FastRNA® Pro Blue Kit and FastPrep® Instrument (Qbiogene, Inc, CA), applying 40 seconds at a speed setting 6.0. Quality of RNA was checked spectrophotometrically measuring triplicate at A_{280} on the NanoDrop® Spectrophotometer and by agarose gel electrophoresis. Dnase I (Promega) was used according to the manufactures protocol for degradation of remaining DNA in the samples. The efficiency of these reactions was confirmed by PCR using *ssb* specific primers (Table 2.6).

2.2.10. Reverse transcription-PCR (RT-PCR)

Reverse transcription was performed using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Reaction mixture contained 2 µL of 10x reaction buffer, 2.5 mM dNTPs, 2 µL 10x random primers (hexanucleotides), 1 µL Rnase Inhibitor, 50U of Multiscribe™ Reverse Transcriptase, and 3 µg of RNA in total volume of 20 µl. Mixture was incubated in thermal cycler in conditions recommended by manufacturer: 10 min at 25 °C, then 120 min at 37 °C followed by 5 s at 85 °C. One-tenth of the reaction volume, typically 2 µL with cDNA, was used for PCR amplification (reaction B, Table 2.9) with *ssb* genes specific primers (Table 2.6). PCR products were analyzed by agarose gel electrophoresis to compare expression of *ssb* genes at different growth conditions and stages of growth.

2.2.11. Rapid amplification of cDNA ends-RACE

Rapid amplification of cDNA ends (RACE) was performed using Invitrogen™ RACE kit to determine transcriptional start sites of both *ssb* genes. Reactions were performed as follows:

1. Reverse transcription was performed using alternative protocol for the first strand cDNA synthesis of transcripts for high GC content as suggested by manufacturer. The reaction mixture containing 2.5 µl 10 x buffer, 10 mM DTT, 2.5 mM $MgCl_2$ and 400 µM dNTP, in total volume of 10 µl were pre-incubated at 50 °C. 3 µg RNA from the exponential phase of growth and 100 nM *ssb* primer (GSP1 or RPR) were mixed and denturated for 10 min at 70 °C, and then immediately transferred to prewarmed reaction mixture. 200 U (1 µl) of SuperScript™ II reverse transcriptase and 1 µl of Rnase inhibitor were added at the end, and the final reaction was incubated for 50 min at 50 °C. In some cases MultiScribe™ Reverse Transcriptase (Applied Biosystems) was used according to the manufacturer's instructions (described in Section 2.2.14). Reactions were terminated by incubating tubes 15 min at 70 °C.

2. Obtained cDNA was purified using S.N.A.P. columns from the kit according to manufacturers' protocol.

3. Purified cDNA was C-tailed using terminal deoxiribonucleoside transferase following instructions from the kit.

4. PCR of dC-tailed cDNA. 2 µl of tailed cDNA, *ssb* nested primers (Table 2.6) and Abridged Amplification Primer (AAP) from the kit were used to perform second round of PCR amplification. When required the third PCR was performed using second PCR product as a template as suggested by manufacturer. Non-tailed cDNA was used in PCR reactions as negative control.

5. PCR products were separated, extracted from the agarose gel, cloned into vector pGEM using TA cloning as described in section 2.2.5. Transformation and colony screening were performed as described in section 2.2.3. DNA from ten clones was sequenced to determine transcriptional start site.

2.2.12. Cell free extract preparation

The mycelia were harvested from 50 mL culture at different stages of growth by centrifugation at 5000 x g and resuspend in Protein lysis buffer (Table 2.1). Cells were disrupted by sonication (4 x 30 s). Amplitude of the ultrasound was set to 40 % of maximal amplitude that can be produced by the Ultrasonic processor (Cole–Parmer). The obtained suspension was centrifuged at 10 000 x g for 10 min to remove cell debris. Protein concentration was determined by Bradford reagent (section 2.2.14).

2.2.13. Protein purification

Over-expression of the *ssb* genes using pQE expression system was achieved in *E. coli* NM522 and in *S. coelicolor* carrying pANT-ssbB. Both SsbA and SsbB proteins were labeled with six hisitidines at the N-terminus to allow fast and easy protein purification from the cell lysate. Histidine tag enables binding of protein to agarose resin with immobilized Ni^{2+} ions. *E. coli* NM522 competent cells were transformed with pQE-ssb and plated on LB agar ON. Selected ampicilin resistant colonies were further incubated ON and next day diluted in fresh LB medium (1/100), supplemented with 100 µg mL^{-1} of ampicilin. The culture reached optical density (A_{600}) of 0.4 after 3-4 h of growth when the expression of recombinant proteins was induced by adding IPTG (isopropyl-b-D-thiogalactopyranoside) to a final concentration of 1 mM (Sambrook 1989). After 3 hours cells were harvested at 4000 g for 15 min and resuspended in Buffer A (Table 2.1). This suspension was supplemented with DNase (1 mg mL^{-1}), cells were disrupted by sonication (4 x 30 s) and cell lysate was centrifuged at 10 000 g for 10 min to remove cell debris. *S. coelicolor* transformed with pANTssbB plasmid was grown in MM (50 mL) for 48 h, mycelia were harvested by centrifugation and

disrupted by sonication (6 x 30 s). The cell lysate was centrifuged at 10 000 g for 15 min to remove cell debris. Either *E. coli* or *S. coelicolor* soluble fraction obtained as described was loaded onto Ni^{2+}-nitrilotriacetic acid agarose column (Qiagen) equilibrated with 10 column volumes of Buffer A (Table 2.1). The column was washed with Buffer B (Table 2.1) containing 30 mM imidazole. This step was monitored by Bradford reagent as described in section 2.2.14, and stopped when no proteins were eluted in the collected fractions. The His-tagged proteins bound to Ni-agarose were eluted with elution Buffer C (Table 2.1) containing higher imidazole concentration (300 mM) and the presence of proteins in the eluted fractions were analysed by Bradford reagent as descirbed. Fractions with purified proteins were pooled and desalted using PD10 columns (GE Healthcare) to 20 mM Tris buffer pH 7.4 for EMSA experiments or Western blot analyses.

2.2.14. Methods for determination of protein concentration

Concentration of proteins in cell extracts was determined by Bradford reagent and Quick Start Bradford Protein Assay Kit (Bio-Rad), following manufacturer's instructions. Three measurements were performed for each sample and protein concentrations were determined by comparison to the standard curve. This curve was created using Bovine Serum Albumin (BSA) Standard Set from Quick Start Bradford Protein Assay Kit (Biorad).

Bradford reagent was also used for monitoring the presence of the proteins in collected fractions during affinity chromatography. This was performed by mixing 5 µL of sample fraction eluted from the column with 20 µL of Bradford reagent in a drop. The protein presence was exhibited by a color change of the sample, from red to blue.

Concentrations of purified proteins were determined by measuring absorbance at 280 nm using NanoDrop® Spectrophotometer. Measurements were repeated three times for each sample. ProtParam tool (ExPASy) was used to determine extinction coefficient and molar mass of each protein. The Lambert-Beer equation (A = E x b x c) was used for all protein calculations to correlate absorbance with concentration, where A is the absorbance value (A), E is the wavelength-dependent molar absorptivity coefficient (or extinction coefficient) with units of L mol^{-1} cm^{-1}, b is the path length in centimeters, c is the protein concentration in mol L^{-1} or molarity (M).

2.2.15. Electrophoretic methods for protein analyses

Proteins were separated by denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE, 12%) usinga Mini-Protean IV vertical cell unit (Bio-Rad) at RT. SDS-PAGE were composed of two different gels: stacking gel (5 % akrilamide:bisakrilamide 29:1, 125 mM Tris-HCl, pH 6.8, 0.1 %

SDS, 0.05 % APS, 0.1 % TEMED) and separating gel (12 % akrilamid:bisakrilamid 29:1, 375 mM Tris-HCl, pH 8.8, 0.1 % SDS, 0.05 % APS, 0.1 % TEMED). Samples were mixed with protein loading dye (Table 2.1) and heated for 5 min at 95 °C before loading on the gel. Run through the stacking gel was performed at 150 V, followed by separation at 180 V until the dye front reached the bottom of the gel. The gels were stained with PhastGel® Blue R (Sigma) after the gel runs were finished. Destaining was performed using following solution: 10 % acetic acid and 10 % 2–isopropanol.

2.2.16. Western blotting

Proteins were electroblotted (300 mA for 1 h) from SDS–PAGE gels onto PVDF membranes using Transfer buffer (Table 2.1.). Membranes were blocked for 1 h at RT in 5 % powder milk (Roth) in TBST buffer (Table 2.1). Incubation with Anti-SSBA antibodies (1:10 000) was performed for 2 h at RT. Membranes were incubated with Anti-SSBB (obtained from Pineda, Anticörper-Service, Germany) antibodies (1:10 000) ON at 4 °C in the blocking solution. Membranes were further incubated with secondary antibodies (Polyclonal Anti-Rabbit Antibodies, 1:10000, GeHealthcare) for 1 h at RT. The signals were visualized with a peroxidase ECL™ Primer Western Blotting Detection Reagent kit (GeHealthcare).

2.2.17. DNA binding shift assay

For DNA band–shift assay experiments increasing concentrations of SSB proteins (see section Results, Figure 3.21) were mixed with 7.7 nM ΦX174 DNA. The reaction solutions containing 40 mM Tris acetate (pH 7.5), 10 mM $MgCl_2$ and 1 mM EDTA in total volume of 20 µl were incubated for 15 min at 37 °C and then run at 0.5 % agarose gel. To test the inhibitory effect of DDT, increasing amounts (10-100 mM) were added to the binding buffer.

2.2.18. Fluorescence microscopy

For the phase contrast and fluorescence microscopy *S. coelicolor* M145 strains were grown in the acute angle of a sterile coverslips inserted obliquely in MS, R5 or MM plates. The coverslips were removed at various time periods and mycelia were treated using Fixative (Glutaraldehyde: Formaldehyde, 0.0045%:2.8% in PBS) as decribed in (Schwedock, McCormick et al. 1997). Propidium iodide stain was used after fixation for visualization of DNA and FITC-WGA for visualization of peptidoglycan using SlowFade ® Antifade Kit. For comparative purposes *S. coelicolor* mycelium was also stained immediately after coveslips removal with DAPI (4',6-diamino-2-phenylindole) at concentration of 1 µg mL^{-1}. For localisation experiments mycelia were stained with Syto 42 live cell

dye(10 μM) and images were taken during the first 30 min. Samples were studied using a Nikon TE2000S inverted microscope and observed with a CFI Plan Fluor DLL-100X oil N.A. 1.3 objective lens, and images were captured using a Hamamatsu Orca-285 Firewire digital charge-coupled device camera. Captured images were processed using IPlabs 3.7 image processing software (BD Biosciences Bioimaging, Rockville, MD). Statistical analyses of spores were performed using Statistica 8 (StatSoft Inc.).

2.3. Computational methods

2.3.1 Collecting sequences

Ref Seq completed genomes database from NCBI (dated August 2013) was used for inquiring/searching for SSB protein sequences. Species containing unusual number of *ssb* genes were rechecked manually by usage of NCBI protein database and BLASTP similarity search restricted to that species. If *ssb* genes other than those of chromosomal origin were annotated (i.e. plasmid), they were excluded from further analyses. Collected sequences were filtered to final dataset consisting of only one representative species from every genus. This was performed in order to reduce large number of very similar sequences. Genomic position of every *ssb* gene from final dataset was checked manually. Five genes upstream and downstream of *ssb* gene were taken into account, not considering the gene order. Conserved genomic neighbourhood was predicted by STRING 9.0 (Szklarczyk, Franceschini et al. 2011).

2.3.2. Annotation of paralogous SSB proteins

Taking into account that until now all studied paralogous SSB proteins, mostly from Firmicutes, have been annotated as SsbA and SsbB, we apply the same style. Therefore all paralogous proteins found in specific species were annotated as SsbA, SsbB, SsbC etc. In addition to ascribe particular protein to SsbA group the following criteria were taken into account: (**i**) existence of the C-terminus within the primary structure, (**ii**) conserved genomic neighbourhood, (**iii**) overall homology to the group of SSB proteins previously confirmed to be essential and (**iv**) the rate of mutation i.e. genes that are changing more slowly are presumed to be essential.

2.3.3. Clustering of SSB protein sequences from Eubacteria

Cluster Analysis of Sequences (CLANS) was used to identify subfamilies of eubacterial SSB sequences and elucidate the relationships among and within the ssDNA-binding protein subfamilies

(Frickey and Lupas 2004). CLANS is a Java utility based on the Fruchterman-Reingold graph layout algorithm. It runs BLAST on given sequences, all-against-all, and clusters them in a 3D space according to their similarity. A 2D-representation was obtained by seeding sequences randomly in the arbitrary distance space. The entire final dataset (see Section 2.3.1) was used for CLANS.

2.3.4. Computing phylogenetic trees

Multiple sequence alignments (MSA) of protein sequences were obtained with Promals (Pei and Grishin 2007). Substitution modeling was completed using AIC in ProtTest (Abascal, Zardoya et al. 2005). Multiple alignment of 16S rRNA gene sequences was performed with ClustalW (Larkin, Blackshields et al. 2007), while statistical selection of models of nucleotide substitution was performed under the AIC in jModeltest (Posada 2008). Phylogenetic trees were constructed using maximum likelihood method in PhyML (Guindon, Dufayard et al. 2010). aLRT values were used to infer branch support. The branches with aLRT values over 0.9 were considered as well supported. Programme Seaview (Gouy, Guindon et al. 2010) was used for statistic report and CorelDRAW® for graphic presentation of the results**.**

2.3.5. Analysis of recent duplications

Synonymous and non-synonymous substitution rate ratio was estimated by using codon models of sequence evolution implemented in CodeML (Goldman and Yang 1994). Values of d_S (substitutions per synonymous site) and d_N (substitutions per non-synonymous site) were calculated as a cumulative value for the pairs of paralogous sequences from the same species by using PAL2NAL (Suyama, Torrents et al. 2006). PAL2NAL converts multiple sequence alignment of proteins and the corresponding DNA (or mRNA) sequences into a codon alignment. Furthermore, it calculates synonymous (d_S) and non-synonymous (d_N) substitution rates in DNA sequence alignment. When the ratio of d_N/d_S (ω) is much lower than one ($\omega \ll 1$) a gene is considered to be under selection; when ω is close to one ($d_N/d_S \sim 1$) a gene is considered to evolve under neutral model (no selection). SSB sequences behaving as recent duplicates, i.e. those branching together on phylogenetic trees and showing high percent of homology were used for d_N/d_S analysis.

3. Results

Both *S. coelicolor ssb* genes are positioned in the central region of the chromosome (Figure 3.1) that is not exposed to frequent deletion events and mostly consists of essential genes (Bentley, Chater et al. 2002).

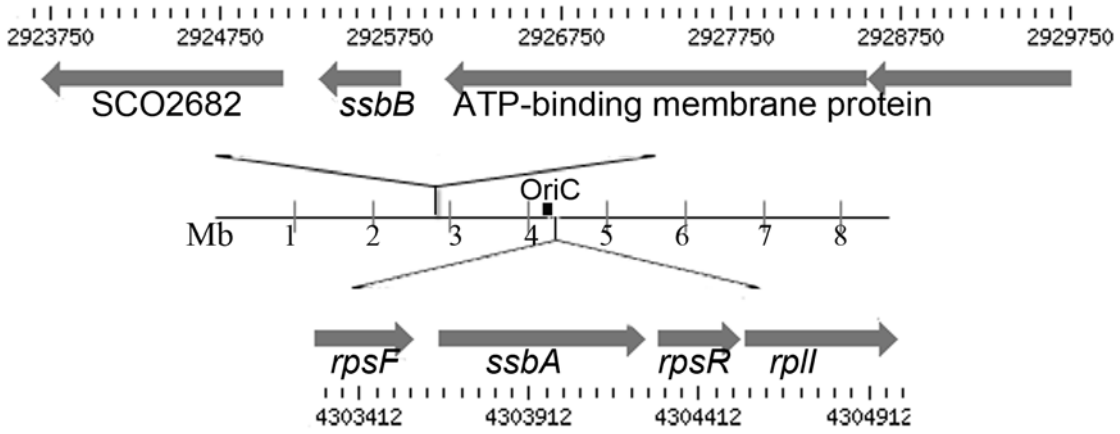


Figure 3.1. Physical map of *S. coelicolor* chromosomal region with two paralogous *ssb* genes. The gene SCO3907 (*ssbA*) encoding a protein of 19.9 kDa, is located 40 kb on the right side of the origin of replication (*oriC*), while the gene SCO2683 (*ssbB*) encoding protein of 16.8 kDa, is located 1.36 Mb on the left from the *oriC*. The both genes are positioned in a region that is not exposed to frequent deletion events and contains many essential genes.

SCO3907 (*ssbA*) encodes a protein of 199 amino acids, while SCO2683 (*ssbB*) encodes a protein of 156 aa (16.8 kDa). The two SSB proteins share an overall 35 % sequence identity mostly conserved in the first 110 aa that contains the OB-fold ssDNA-binding domain. The C-terminal region (89 aa long) of SsbA is enriched in glycine (over 50 %), making it highly flexible and disordered. An acidic ^{194}DEPPF sequence, at the very end of this protein, is essential for interacting with proteins (Shereda, Kozlov et al. 2008) or with the OB-fold (Marintcheva, Marintchev et al. 2008). In SsbB, the C-terminus is reduced to 50 aa and it lacks the acidic tail, which is highly conserved in all SsbA sequences (Shereda, Kozlov et al. 2008). Figure 3.2 represent schematic comparison of SsbA and SsbB primary structures.

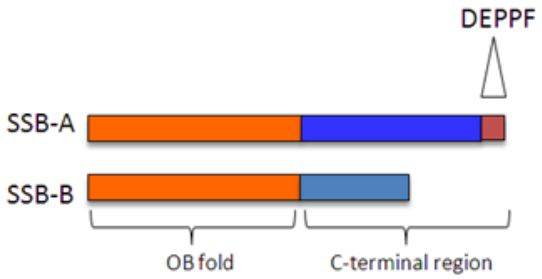


Figure 3.2. Comparison of SsbA and SsbB primary structures. The homology is conserved mostly within first 100 aa, whereas C-terminal region is much shorter in SsbB and lacks conserved acidic Ct

3.1. Determination of biological roles of SSB proteins from S. coelicolor

3.1.1. Disruption of ssbA gene

Retention of duplicate genes depends on their adoption of novel functions in the cell (Bratlie, Johansen et al. 2010). In order to elucidate the biological role of paralogous SSBs in *S. coelicolor*, knock-out experiments were performed as described in Materials and Methods. In order to disrupt *ssbA* by semi-targeted *in vitro* transposon mutagenesis cosmid h24.EO1 (apramycin resistance - Apr^R, kanamycin resistance - Kan^R) was introduced to *S. coelicolor* by intergeneric conjugation and selected for Apr^R exconjugants. After screening 500 colonies for the Apr^R Kan^S phenotype indicative of integration of h24.EO1 by a double cross over event, all colonies displayed the Apr^R Kan^R, phenotype that indicated a single recombination event. The failure to obtain Apr^R Kan^S colonies suggests that either allelic exchange could not occur easily, or that *ssbA* is an essential gene or the third, that this exchange would cause a block in transcription of downstream essential genes. Since the *ssb*A gene is clustered with genes encoding ribosomal proteins (*rpsF* and *rpsR*), as in *B. subtilis* (Lindner, Nijland et al. 2004) and owing to the reported co-regulation of *ssbA* and ribosomal genes, an in-frame deletion of *ssbA* to allow transcription of downstream gene(s) was created (Figure 2.3). Firstly, recombination between cosmid SCH24 and PCR amplified resistance cassette took place in BW25113 (pIJ790) recombinogenic *E. coli* strain. Primers for the resistance cassette had homologous regions for upstream and downstream sequence of *ssbA* gene which allowed recombination of resistance cassette with SCH24 cosmid. Obtained recombinant cosmid was designated cSA01 (SCH24ΔssbA). Purified recombinant cosmid was transformed into *E. coli* BT340 strain (harbouring the Flipase gene, FLP) in which next recombination step occurred. Resistance cassette was flipped out using FRT recombination sites in the cassette. New cosmid, cSA02 consisting of 81 base pairs (“scar“) remained in frame with the adjacent ORF. These events were monitored by the loss of apramycin resistance and PCR reactions (Figure 3.3). Primers SSBAFRT and izaSSBA (Table 2.6), complementary to flanking regions of *ssbA* gene were used for this purpose (Figure 3.1).

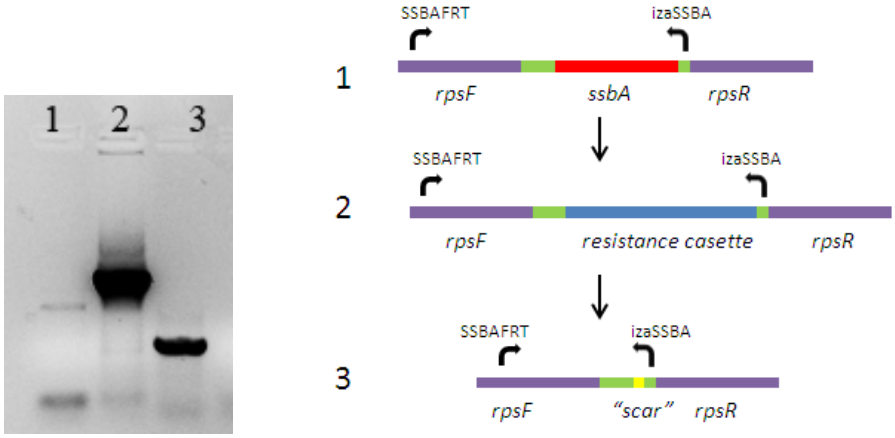


Figure 3.3. PCR reactions showing steps in "scar" preparation using flanking primers SSBAFRT and izaSSBA. Lane 1. SCH24 (wild type cosmid), lane 2. cSA01 (SCH24ΔSsbA), lane 3. cSA02 (SCH24scar). Scheme of *ssbA* gene disruption, flanking regions and used primer pairs are presented on the right; number of each step is equivalent to the PCR products obtained in the lanes 1, 2 and 3 on the left.

"Scar" cosmid was transformed into *S. coelicolor* protoplasts. Namely, transfer by conjugation was not possible since excision of resistance cassette led to loss of *oriT* as well. In spite of repeated attempts, no viable mutant was obtained. Deletion of *ssbA* was successful only when an additional copy of the gene with its original promoter sequence was integrated into the genome at the *att* site. Exoconjugants with an extra copy of *ssbA* gene were selected by hygromycin, and verified by PCR reactions using Sco4848 primers (Table 2.6). These primers were designed to detect disruption in the SCO4848 gene, which was published previously as a site of integration of ΦBT1 (Gregory, Till et al. 2003). *S. coelicolor* carrying additional copy of *ssbA* gene was transformed with a scar cosmid cSA02. Double crossing-over was indicated by the loss of kanamycin resistance. Desired event, i.e. presence of scar sequence in place of *ssbA* gene had to be confirmed confirmed by PCR reaction since the scar sequence does not contain any selection marker. Figure 3.4 represents colony PCR with flanking primers (as in Figure 3.3) on several kanamycin sensitive transformants of *S. coelicolor* after one generation without selection pressure. Most of the transformants revert to wild type state as upper band indicates, i.e. the whole cosmid has been excised and only the wild type gene copy of *ssbA* has been present. Lanes 1 and 2 represent transformants with *ssbA* gene replaced with the "scar" sequence. Lane 4 is showing existence of both wild type gene, and scar sequence. This clone was still resistant to kanamycin.

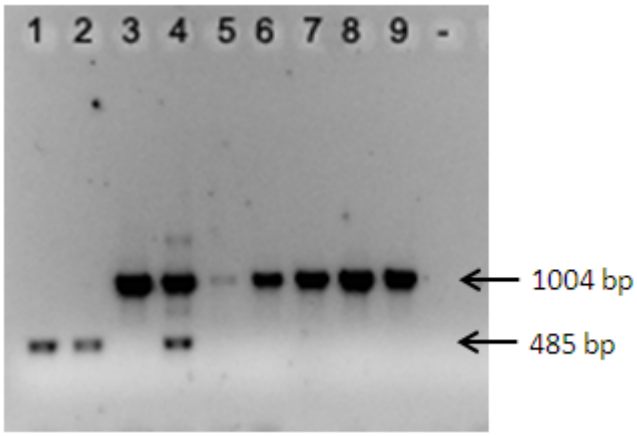


Figure 3.4. PCR products on *S. coelicolor* kanamycin sensitive colonies after one generation without selection pressure. The 485 bp product apperas in lanes where double cross-over occured leaving „scar" sequence instead of ssbA gene

Based on this result as concluded that *ssbA* is essential for *S. coelicolor* survival.

3.1.2. Disruption of ssbB gene

To examine the biological role of *ssbB*, this gene was disrupted with Tn*5062* transposon in the same manner as described for *ssbA* gene (section 3.1.1.). Seven out of one hundred exconjugants exhibited double cross-over, i.e. displayed $Apr^{R}Kan^{S}$ phenotype. It is expected that less than 10% percent of *Streptomyces* colonies will loose kanamycin resistance (Gust, Chandra et al. 2004). Correct positioning of the transposon in the obtained mutant strain (TSB01) was confirmed by three PCR reactions (Figure 3.5).

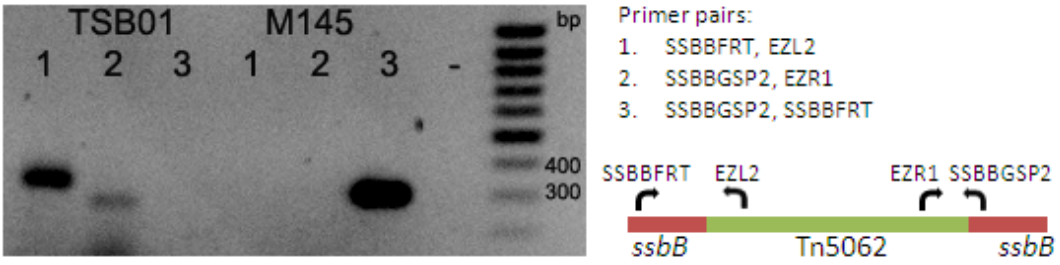


Figure 3.5. PCR products showing correct positioning of the transposon Tn*5062* within *ssbB* gene in TSB01 mutant. PCR reactions with the same primer pairs were performed on the wild type strain (M145). The lane numbers are equivalent to reactions with primer pair number on the scheme. Mutant strain had integrated transposon (Lanes 1 and 2, TSB01), and product corresponding to the wild type gene amplification was absent (Lane 3, TSB01), whereas for the wild type strain (M145) PCR product was obtained only in lane 3, as expected.

Interestingly, after four days of growth, the mutant strain (TSB01) displayed white aerial mycelium on MS medium (Figure 3.6.A). Formation of the white aerial hyphae (*whi* phenotype (Jakimowicz, Mouz et al. 2006)) is characteristic of *Streptomyces* mutants that fail to develop grey pigment associated with mature spores. Further, we observed that after prolonged growth (10 days) the surface of our mutant turned light grey (Figure 3.6.B) indicating that a small proportion of mutant

hyphae could complete the spore maturation process. To exclude the possibility of an unpredicted polar mutation, the strain TSB01 lacking *ssbB* was complemented with the native gene integrated at the $attB_{\Phi BT1}$ site under the control of its own promoter. Strain TSB02 carrying this complementation vector, fully restored the original phenotype (Figure 3.6). This was achieved by conjugation between TSB01 and *E. coli* ET12567, previously transformed with pMS-pssbB construct (*ssbB* gene with its native promoter). Hygromycin resistant colonies had intergrated pMS-pssbB construct precisely at the unique *att* site, inside of SCO4848 gene, which was confirmed by PCR reactions using SCO4848 primers (not shown).

Figure 3.6. Phenotypes of the *ssbB* mutant (TSB01), complemented strain (TSB02) and their parental strain *S. coelicolor* M145. Strains grown at 30 °C on solid MS medium. TSB01 displayed a white phenotype on MS medium after 4 days (**A**) and light gray after 10 days of growth (**B**), while the phenotypes of parent strain and TSB02 were similar.

To check if C-terminus of SsbB is important for observed phenotype, mutant carrying *ssbBΔC* has been constructed i.e. TSB03 strain (see Materials and Methods). The expression of truncated SsbB in TSB03 strain was confirmed by Western blot using antiSSBB antibodies (Figure 3.7).

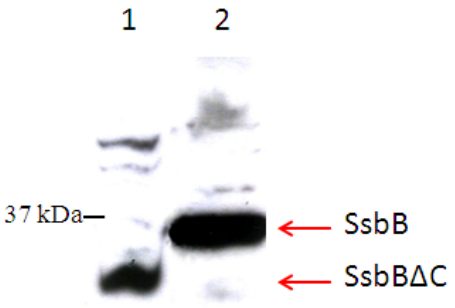


Figure 3.7. The comparison of expression of truncated SsbB (SsbBΔC) in TSB03 strain and SsbB in the wild type strain (wild type). Western blot analysis was performed on 30 µg of *S. coelicolor* cell lysates using antiSSBB antibodies. Signal corresponding to protein of lower molecular weight was obtained for strain TSB03 (Lane 1), while signal for whole length protein (34 kDa) was present in the wild type strain (Lane 2). SsbB in both strains was present in dimeric form.

This strain also exhibited white phenotype as shown for TSB01 (Figure 3.8.A). Thus this result confirmed that SsbB truncated at the C-terminus and with intact OB-fold could not restore spore maturation process. Both, TSB01 and TSB03 displayed a similar phenotype on R5 medium. A delay in formation of aerial mycelium was observed and after prolonged incubation mutant colonies displayed a more pronounced pink colour (Figure 3.8.B) (Ryding, Bibb et al. 1999).

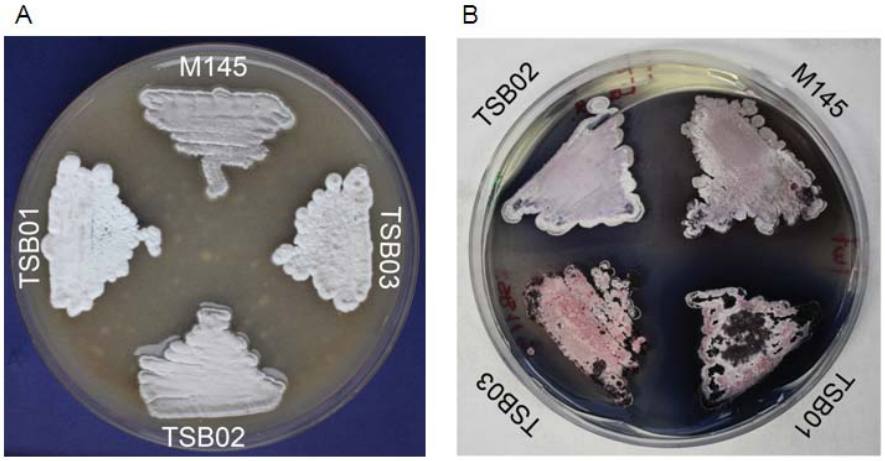


Figure 3.8. Comparison of parent *S. coelicolor* M145, mutant (TSB01), complemented strain (TSB02) and strain complemented with ssbBΔC (TSB03). Growth on MS (**A**) and R5 (**B**) plates after 4 days.

Microscopy analysis was performed to examine mutants displaying whi phenotype. It was observed that it is more was more difficult to find regions with abundant spore chains in TSB01. Mutant strain produced 32.8 % less spores than the wild type strain (Paradzik, Ivic et al. 2013). As shown in Figure 3.9 A1-3, closer inspection showed that spores in the mutant strain frequently lacked DNA (30 %; n=2200) or contained excessive amounts of DNA (23 %, n=900), suggesting a defect in chromosome segregation. In contrast, the complemented strain (TSB02) displayed restoration of the wild type phenotype (Figure 3.9 C1-2), with just 1 % of spores lacking DNA in TSB02 compared to the wild type in which 0.8 %–1.1 % spores lacked DNA (Dedrick, Wildschutte et al. 2009; Ditkowski, Troc et al. 2010).

To exclude the possibility that observed defect is an artefact because of the difficulty in stain penetration through the spore wall *in vivo* staining of the *S. coelicolor* wild type, TSB01, TSB02 and TSB03 (Figure 3.10) was performed. Results of DAPI staining showed similar results as observed previously, in TSB01 25 %, TSB03 26 % and wild-type 1.5 % of spores were deficient in DNA (n=800).

Statistical analysis showed that TSB01 had a slightly increased spore length and number of spores in spore chains (Table 3.1). The length of the spores in TSB01 was less uniform than in the wild type,

as could be seen from the standard deviation values. An F-test was also applied showing that TSB01 spores were significantly longer than those of the wild-type strain (P< 0.001).

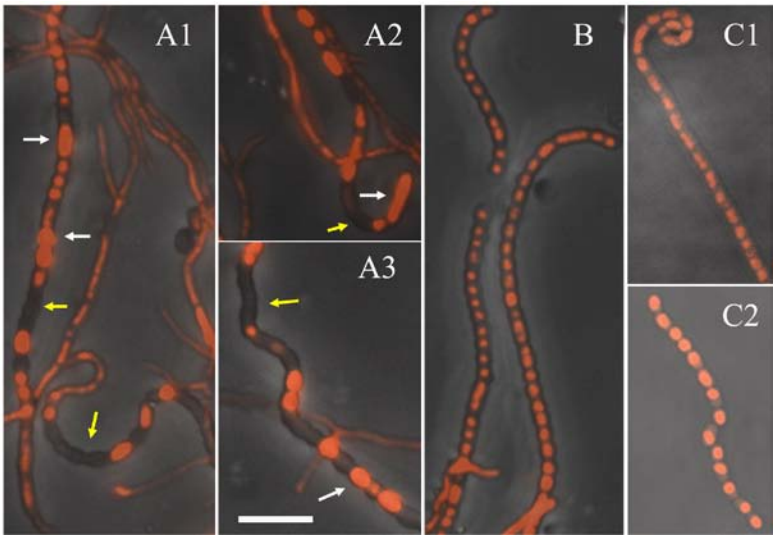


Figure 3.9. Phenotypic characterisation by fluorescence microscopy revealed segregation defect in the strain lacking *ssbB*. Strains were grown on glass cover slips for 3 days and imaged with the fluorescence microscopy. The aberrant distribution of DNA is shown in spore chains of TSB01 (**A1-3**) whereas proper segregation can be seen in the wild-type strain M145 (**B**), and TSB02 (**C1-2**). White arrows point to the spore with an excessive amount of DNA while yellow to the compartment lacking DNA. Bar 5 µm.

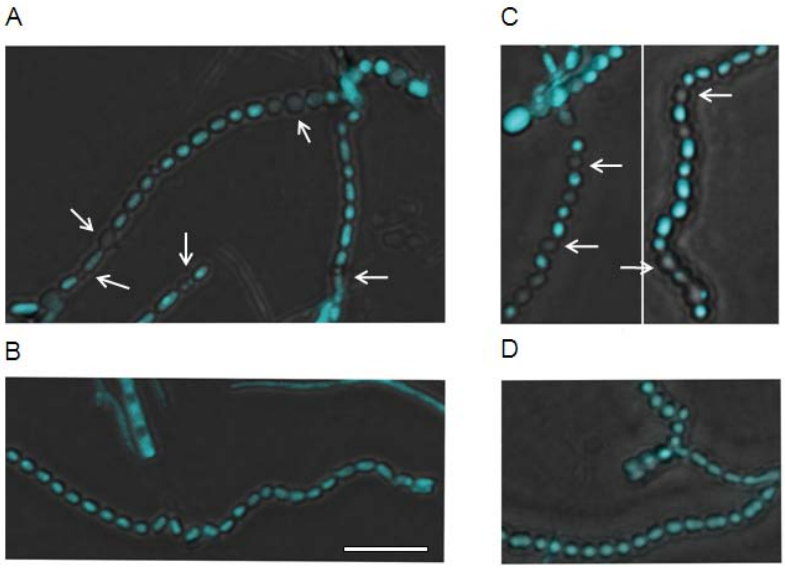


Figure 3.10. Phenotypic characterisation by fluorescence microscopy after *in vivo* staining by DAPI revealed a segregation defect in (**A**) the strain lacking *ssbB* (TSB01) or (**C**) only C-terminus of *ssbB* (*ssbBΔC* ,TSB03), whereas proper segregation can be seen in (**B**) M145, wild type strain, (**C**) and (**D**) TSB02, *ssbB* mutant complemented with *ssbB.* Strains were grown on cover slips for 3 days and imaged as described in Materials and Methods. White arrows point to the spore without DNA. Bar 5 µm.

Table 3.1. Effect of the absence of *ssbB* in *S. coelicolor*. The mutant strain, TSB01 has an increased number of spores in spore chains and slightly increased spore length.

Variable	$\bar{x} \pm s^{*}$ (95 % CI)†		d‡ (95 % CI)§	p||
	M145	TSB01		
N_o of spores in chain (n=100)	27.20 ± 5.99 (2.50 -28.90)	40.42±13.60 (6.54 -44.30)	13.22 (9.04 -17.40)	< 0.001
Spore length (n=200)	1.24 ± 0.19 (1.21 – 1.28)	1.36 ± 0.35 (1.30 – 1.43)	0.12 (0.04 – 0.20)	< 0.001

*Mean ± standard deviation
† Lower and upper limit of the interval estimate for the mean with 95% confidence
‡ d is the difference between the means of groups
§ Lower and upper limit of the interval estimate for d values in the population with 95% confidence
|| p value obtained with t-test

The aberrant distribution of DNA in the spores and irregularly sized spores suggest that SsbB has a distinctive function for chromosome segregation (Flardh and Buttner 2009). We also observed that the strain deficient in SsbB synthesis occasionally produced longer and coiled aerial hyphae, (Figure 3.11) as previously reported for the *whiA* mutant (Ainsa, Ryding et al. 2000).

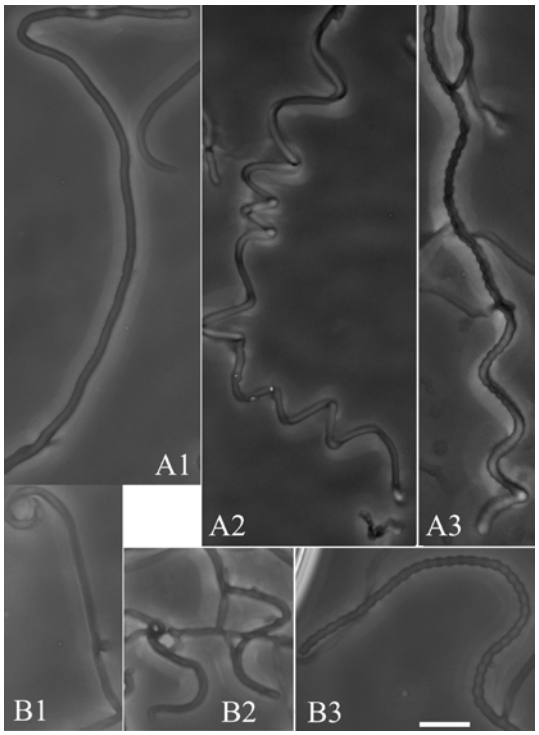


Figure 3.11. (**A1-3**) Phase-contrast micrographs of TSB01 (*ssbB* mutant) and (**B 1-3**) the wild-type strain M145. **A1** and **B1** represent young aerial hyphae, **A2** and **B2** more mature aerial hyphae and **A3** and **B3** spore chains. Bar = 5 µm.

The result clearly showed that *ssb* paralogues perform different biological roles; *ssbA* is essential for cell survival, while *ssbB* is important for production of viable spores.

3.2. Transcriptional analyses of ssb genes

3.2.1. Expression profiles of ssb genes

S. coelicolor was grown in MM (minimal medium) and CRM (rich medium) media for 4 days and total RNA was isolated at different time intervals that coincided with metabolic changes, monitored by accumulation of actinorhodin (Figure 3.12 C1-2), a red/blue acid/base indicator pigment with weak antibiotic properties (Kieser 2000). Total cDNA was obtained with random hexanucleotides using High capacity cDNA kit (Materials and methods). As a RNA quantity control, PCR with 16S rRNA primers was performed. As shown in Figure 3.12, the reaction yield was the same in every reaction i.e. the same amount of starting RNA was used to perform reverse transcription for every time point. Gene specific primers were used in different combinations (Figure 3.12 A) for *ssbA* transcript since it is predicted that this gene is transcribed in operon. RT-PCR analysis of *ssbA* confirmed that this gene is expressed as a long mRNA with *rpsF* and most likely *rpsR* from a promoter positioned upstream of *rpsF*. Expression is more pronounced in the exponential phase of growth in rich medium (Figure 3.12 A1), while in minimal medium (Figure 3.12 A2) expression is prolonged into stationary phase. Under the same conditions, the *ssbB* transcript was hardly detectable in MM, while in rich medium it was undetectable.

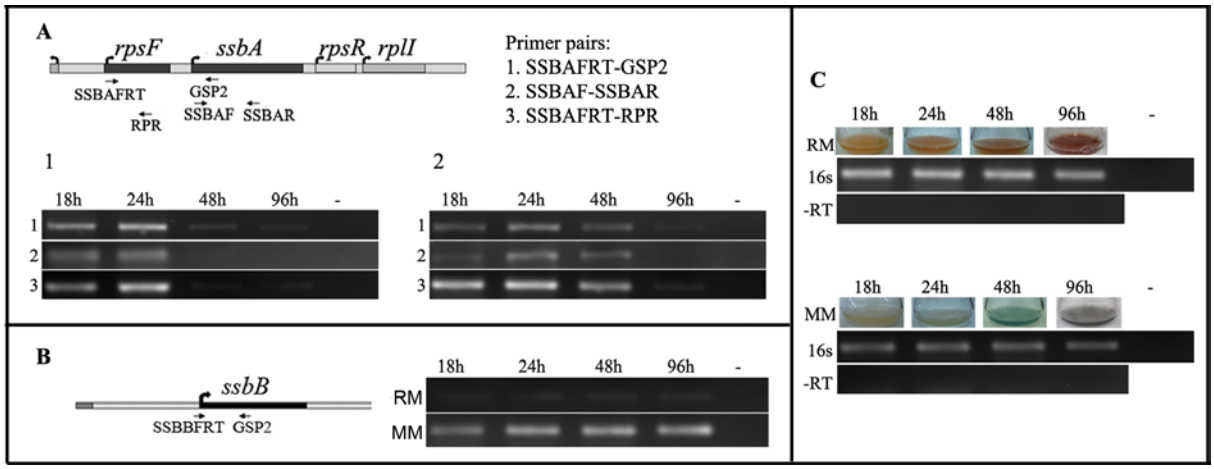


Figure 3.12. Expression profiles of *ssb* genes. Samples were taken at different stages of growth characterised by visible actinorhodin accumulation. Gene expression was analysed by RT-PCR as described in Materials and Methods. Total cDNA was obtained using random hexanucleotide primers and PCR was performed with gene specific primers positioned as indicated in **A** and **B**. (**A1**) Expression of *ssbA* in rich medium (RM), (**A2**) expression of *ssbA* in minimal medium (MM). (**B**) Expression of *ssbB* gene in RM and MM. (**C**) Control reactions and *S. coelicolor* culture flasks at given times, 16S - upper panel shows expression of 16S rRNA in RM, while 16S - lower panel in MM, -RT shows control reactions without reverse transcriptase.

Therefore RT-PCR was extended to 35 cycles; even under modified condition, *ssbB* mRNA was barely detectable in rich medium. On the contrary, in minimal medium gene is expressed more or less equally throughout 96 hours of growth (Figure 3.12 B).

3.2.2. Identification of the promoter regions of ssbA and ssbB genes

Since no obvious conserved promoter regions are present in the upstream regions of the two *ssb* genes, mRNA transcripts were analysed using 5'-RACE. As shown in Figure 3.12 A, RT-PCR revealed that *ssbA* was transcribed as a long mRNA with ribosomal gene(s). Two primers, GSP1 located at the beginning of *ssbA* and RPR located at the end of *rpsF* (Figure 3.12 A) were used to obtain cDNA in a RACE experiment. Two transcriptional start points (TS1 and TS2) of *rpsF* were identified. TS1 was determined to be a G or C 75 nucleotides upstream of the start codon of *ssbA*, while TS2 was located 163 nucleotides upstream of the translational start codon of *rpsF*. By inspecting upstream regions of TS1 and TS2 we could not find obvious promoter consensus sequences. However, detailed analysis of the upstream region showed a TTTACT sequence, 6 nucleotides apart from the TS1 and a GAC motif 16 bp upstream of this sequence (Figure 3.13). These two putative promoter elements most closely resembled previously reported subclass G of *Streptomyces* promoters, GAC ($N_{18/19}$) T(N_4)T (Bourn and Babb 1995). Comparative analysis of these regions from other *Streptomyces* genomes displayed genetic variation upstream of the TS1 owing to short insertion and deletion events (Figure 3.14). Therefore, in spite of the syntheny, the proposed promoter elements of the proximal promoter are located in the poorly conserved DNA region. The sequence homology blocks were found upstream of the TS2 (Figure 3.14). The sequence TACGCT, 6 nucleotides apart from the TS2, might act as a -10 consensus sequence (4 out of **TA**g(**Pu**)(Pu)**T,** (Strohl 1992)) while sequence TTTCAA spaced 18 bp from the proposed -10 sequence, might act as a -35 consensus sequence (4 out of **TT**G**A**C(**Pu**) (Strohl 1992)). To check the possibility that transcription of *ssbA* occurs from a proximal promoter located in the intergenic (81 bp) or coding region of *rpsF*, additional RACE experiments with a SSBAR primer complementary to the 3' end of the *ssbA* gene were performed. However, no other transcriptional signal(s) could be observed.

The transcriptional start of *ssbB* was determined to be a C, 73 nucleotides upstream of the AUG codon. Analysis of *ssbB* promoter region did not show typical ***S****treptomyces-****E****. coli* like **p**romoter sequence (**SEP**-like; TTGACA-18bp-TCTTAT). Detailed inspection showed complex DNA region with a palindromic sequence, DnaA box and two long imperfect direct repeats as indicated in Figure 3.13. Moreover, part of this long repeat (GGCTGCCACCCGTGC) was found at three additional positions in this promoter region (Smulczyk-Krawczyszyn, Jakimowicz et al. 2006).

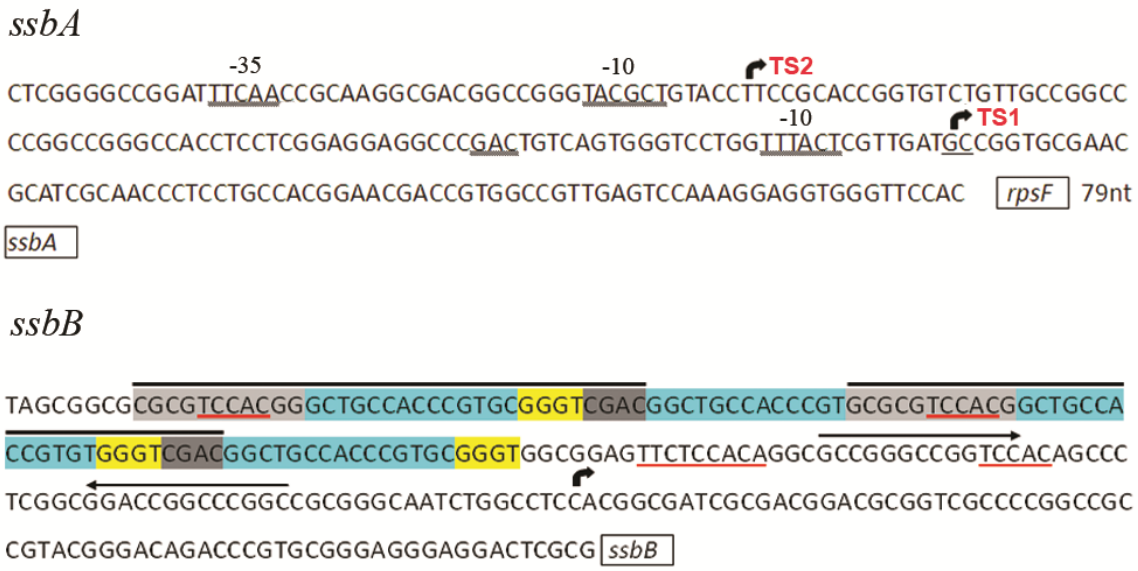


Figure 3.13. Transcriptional start sites of the two *ssb* genes. The results of RACE experiment revealed that the transcription of *ssbA* gene starts at 75 bp (TS1) and 163 bp (TS2) upstream of *rpsF* gene (indicated by rightwards arrows). -10 consensus sequence and -35 consensus sequence are underlined. Proposed promoter elements upstream of TS2 most closely resemble **SEP**-like sequences (TTGACPu-18- TAgPuPuT (Strohl 1992).The transcription start site for *ssbB* gene is 73 nucleotides upstream of start codon (indicated by rightwards arrow). Promoter region of *ssbB* contains a palindrome sequence (opposite directed arrows), DnaA box (red underline) and two long imperfect direct repeats (black overline), while identical shorter repeats are labelled in the same colour.

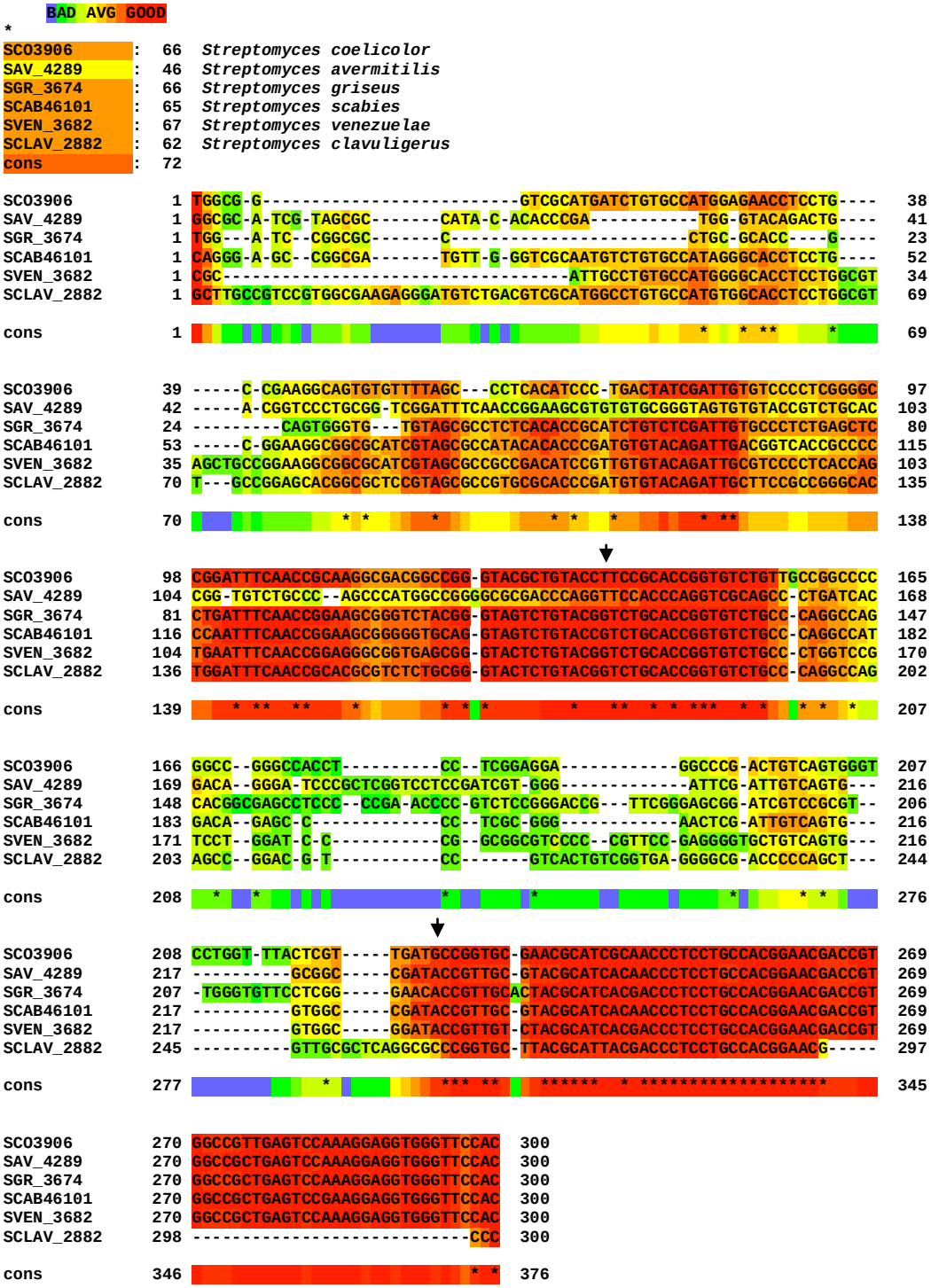


Figure 3.14. Multiple alignment of *rpsF* upstream regions from six *Streptomyces* species. While region upstream of TS2 (upstream black arrow) transcriptional site is conserved, significant genetic variation upstream of the TS1 (downstream black arrow)owing to the short insertion and deletion events is observed.

3.3. Expression of recombinant ssb genes in heterologous and homologous host

3.3.1. Gene cloning and expression in E. coli

Genetic manipulation in streptomycetes is often tedious and complex; therefore all genetic constructs were prepared in *E. coli* and then either transferred to *S. coelicolor*, or used for over-expression in *E. coli* for biochemical characterization. Constructs pGEMssbA and pGEMssbB were initial genetic constructs for all further cloning and subcloning. Existence of EcoRI/BamHI/HindIII restriction sites allowed easy replacement of sequence encoding His-tag (EcoRI/BamHI) with other suitable tags (Figures 2.1 and 2.2), such as eGFP, mCherry or regulatory regions when needed. Tags were always added at N-terminal regions of SSB proteins, since C-terminal regions could be involved in protein interactions (Shereda, Kozlov et al. 2008). All constructs for expression in *S. coeliolor* had corresponding streptomycete RBS, to allow proper translation in homologous system. pQE vector was used for heterologous over-expression of both *ssb* genes, since it has N-terminal in-frame sequence encoding for His-tag, that allowed easy purification from the cell lysate. During the over-expression of *ssbA* and *ssbB* genes inhibition of *E. coli* growth was observed. The inhibition was more pronounced when SsbB was over-expressed (Figure 3.15).

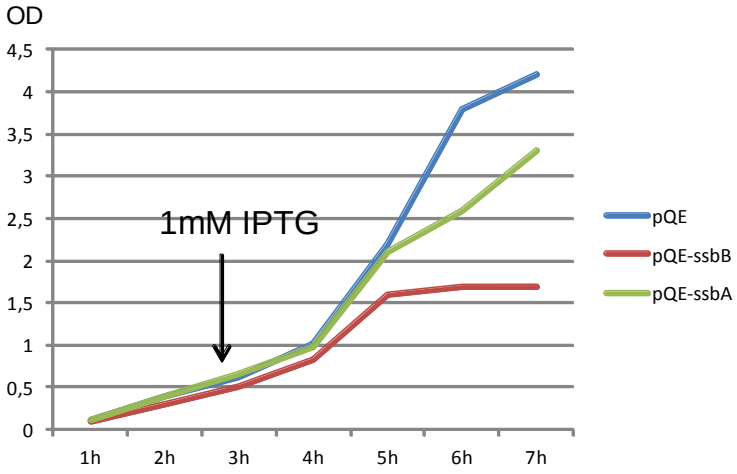


Figure 3.15. Induction of expression of *ssbA* (pQE-*ssbA*) and *ssbB* (pQE-*ssbB*) genes in *E. coli* NM522. IPTG was added at OD 0.5. *E .coli* cells expressing *ssbA* and *ssbB* genes showed delay in growth compared to strain carrying only pQE30.

To elucidate localisation of SSB proteins within *S. coelicolor* mycelia, they were labelled with fluorescent proteins. SsbA was labelled with enhanced green fluorescent protein (eGFP), and SsbB with red fluorescent protein (mCherry). Both constructs were integrated into *S. coelicolor* chromosome, using integrative vectors, which were delivered to *S. coelicolor* via conjugation with *E. coli* ET12567 strain. Natural *ssb* promoter regions were added in front of fluorescent tags. This was performed to assure near-physiological expression of both proteins. PromoterlessmCherrySsbB

construct was cloned into pANT849 plasmid vector, in case that localisation of integrated mCherryssbB would not be detectable due to low level of activity of ssbB natural promoter.

Interestingly, while constructing mCherry-*ssbB* construct with its own promoter, we observed activity of this promoter in *E. coli*. Although this activity was not comparable to that of the T7 promoter, accumulation of mCherry-SsbB caused filamentation (Figure 3.16). Moreover, SsbB showed alternating localisation pattern, as indicated in Figure 3.15 A.

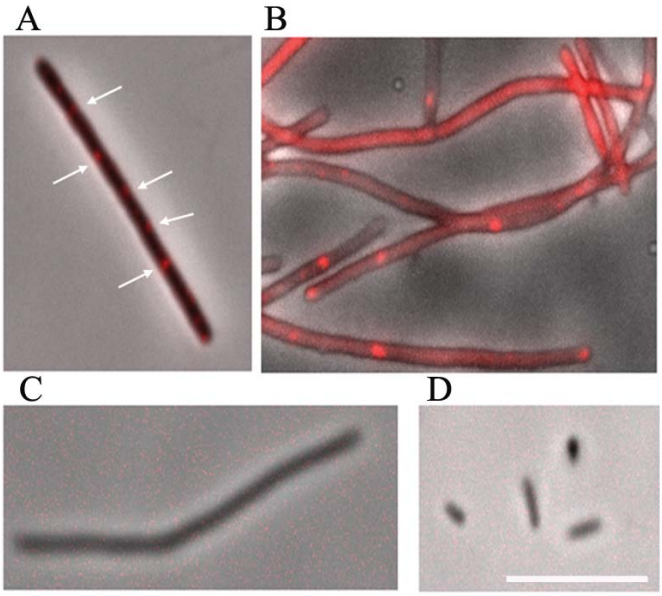


Figure 3.16. Expression of *S. coelicolor ssb* genes in *E. coli*. Fluorescence microscopy of *E. coli* carring different *ssbB* constructs: **A**. cell expressing mCherryssbB gene fusion under the *ssbB* gene native promoter (pSET-pmCherryssbB), **B.** cells expressing mCherry *ssbB* gene fusion under the T7 promoter (pGEM-mCherryssbB), **C**. cell expressing the *ssbB* gene under its native promoter (pMS-pssbB),**D**. wild type *E. coli* cells. Alternating pattern of SsbB localisation was indicated by white arrows. Bar 10 µm. Magnification 100x.

3.3.2. Localisation of SSB proteins in streptomycete mycelia

In order to determine localisation of SsbA and SsbB protein in *S. coelicolor* mycelia, SsbA was labelled with eGFP and SsbB with mCherry. Both genes were placed under the control of its native promoters, and were integrated into *att* sites within *S. coelicolor* chromosomes (see Material and methods). Expression of fusion SSB proteins was confirmed by Western blot (Figure 3.17). The detected amounts of eGFPSsbA and mCherrySsbB was approximately similar to that of the wild type proteins.

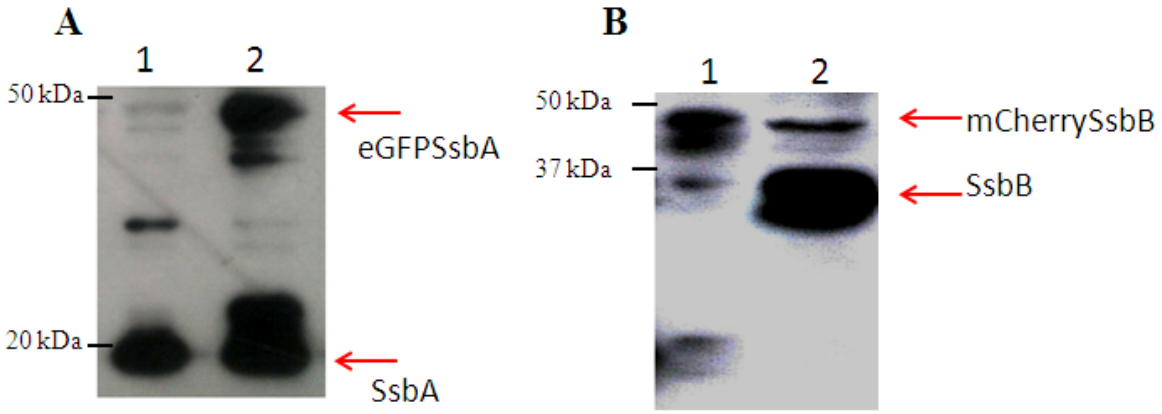


Figure 3.17. The expression of fluorescently labelled SsbA and SsbB in *S. coelicolor*. **A.** Western blot analysis was performed on 2 µg of cells lysates using antiSSBA antibodies. In the cell lysate obtained from the *S. coelicolor* M145 (lane 1), the signal belongs to SsbA, whereas in the lysate obtained from TSA02 (*S. coelicolor* M145 peGFPssbA) strain, there is additional signal belonging to eGFPSsbA fusion protein (lane 2). **B**. Western blot analysis was performed on 30 µg of cells lysates using antiSSBB antibodies. Signal corresponding to a fusion protein (mCherrySsbB) was obtained for strain TSB04 (Lane 1), while signal for dimeric SsbB protein (34 kDa) was present in wild type strain (Lane 2).

The *S. coelicolor* M145 strains used for fluorescent microscopy and general observations regarding SSB proteins localisation were described in Table 3.2. Strains TSA02 (eGFPSsbA) and TSB04 (pmCherrySsbB) were used as negative controls to confirm low level of fluorescence in the absence of fluorescently labelled protein (i.e. no fluorescence with TRITC filter (red) without the presence of mCherry label, and no fluorescence with FITC filter (green) without the presence of eGFP).

Table 3.2. General observation on SSB protein localisation in *S. coelicolor*

M145 strain carrying	**Localisation in rich medium**			**Localisation in minimal medium**		
		CL	**DNA**		**CL**	**DNA**
TSAB01 GFPSsbA- NP mCherrySsbB- NP	 -diffuse or dispersed -foci occasionally observed	- 	 + -	 -bright foci in young hyphae, or diffuse -bright foci, diffused	+/- 	 + -
TSAB02 GFPSsbA-NP pANTmCherrySsbB, PP	 -bright regular foci -overexpression	? 	 + -	 -bright regular foci -overexpression	? 	 +/- ?

NP-natural promoter, **PL**-plasmid promoter, **CL**-colocalisation of SsbA and SsbB, **DNA**- colocalisation with DNA

Different types of fluorescence were observed along hyphal compartments, from bright foci to diffuse foci (often in the same compartment) and from dispersed fluorescence to a lack of fluorescence. SsbA showed diffused localisation most of the time. When this signal was compared to the wild type strain at the same exposition parameters, this diffuse pattern was not the result of autofluorescence. eGFPssbA bright foci were found mostly within young vegetative hyphae and at the tips of young aerial mycelia (on average 6 foci, ranging from 1-14). At the same time, they are visible only in 10 % of spore chains, while in all other spores, signal was dispersed (Figure 3.17 A), or no fluorescence was observed. When present, bright foci were spaced 0.3- 5.5 µm, but sometimes this distance was larger.

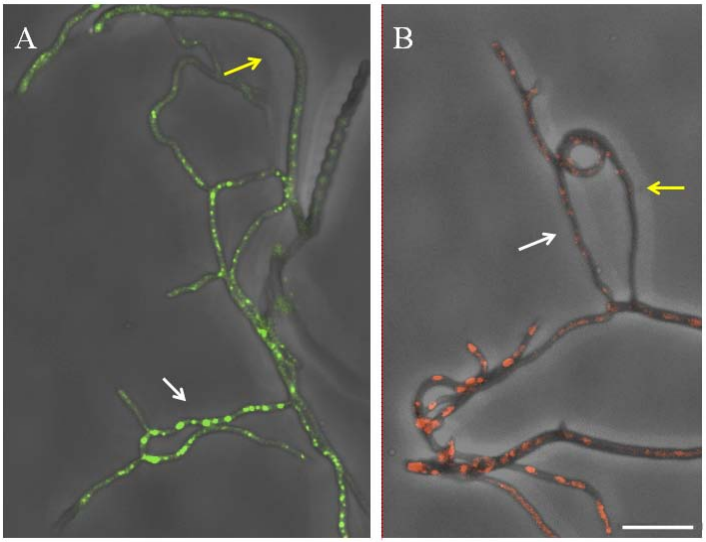


Figure 3.18. Localisation of SsbA (**A**) and SsbB (**B**) in *S. coelicolor* mycelia grown in minimal medium. Both proteins were found vegetative (white arrows) and in aerial mycelia (yellow arrows). Bar = 5 µm.

SsbB often had more intensive fluorescence than SsbA (Figure 3.18 B), with average of eight bright foci per lateral vegetative branch (up to 10 µm), and 11 bright foci in aerial mycelia. The fluorescently bright foci had variable spacing that ranged from 0.7 to 6 µm. This is statistically significant difference in comparison to distance between foci of SsbA protein ($P<0,001$). mCherrySsbB signals were weaker; both in intensity and in the number of foci in rich medium (average 5 foci per hyphae for both vegetative and aerial mycelia). Colocalisation of SsbA and SsbB was almost never observed during growth in rich media (Figure 3.19), while it was possible to detect occasional colocalisation in minimal medium (not shown). The SsbB foci were always more bright in such cases. While SsbA localise with DNA (Figure 3.19 D), SsbB was found between nucleoids (Figure 3.19 E).

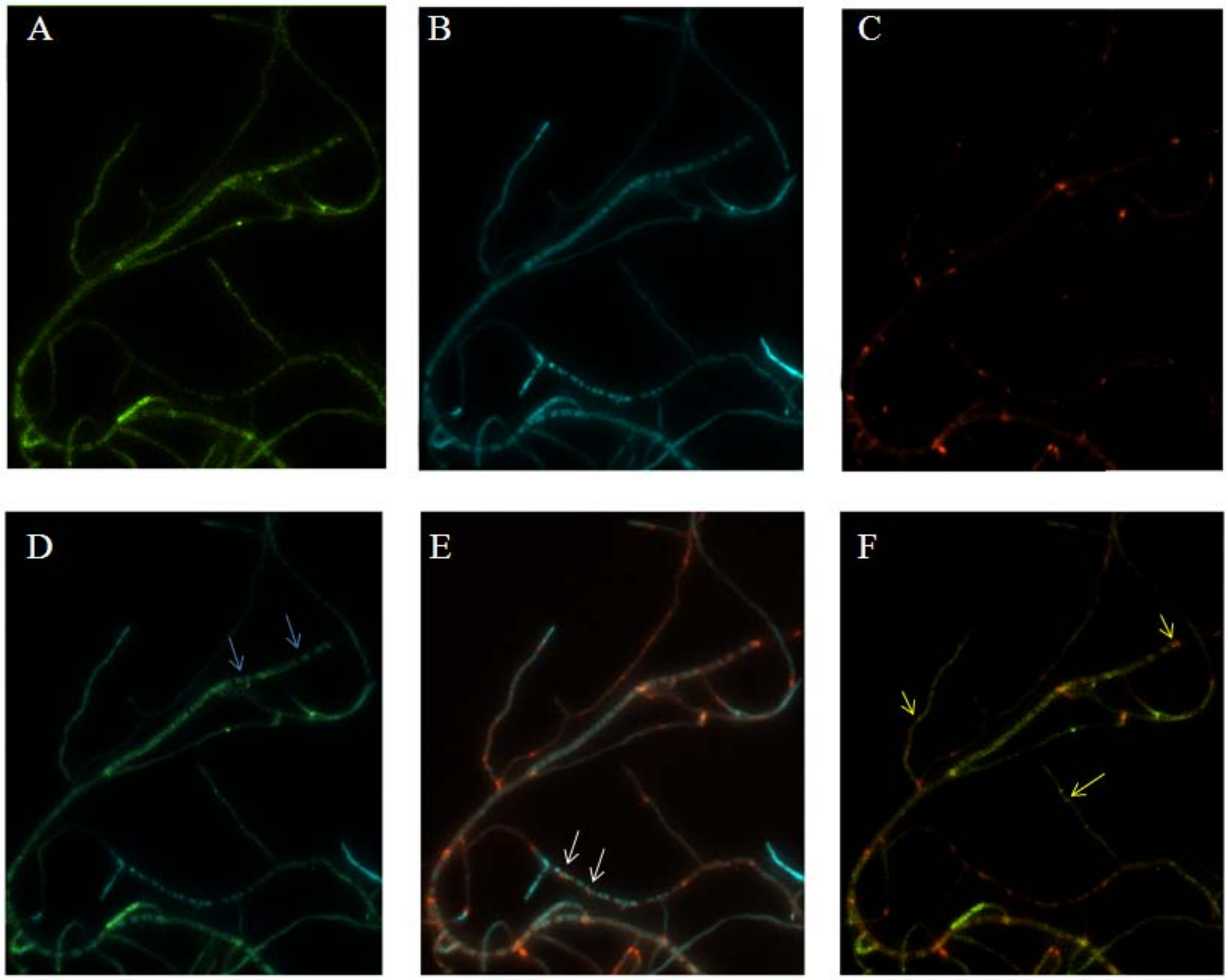


Figure 3.19. Localisation of SsbA and SsbB proteins in *Streptomyces coelicolor* mycelia grown in rich medium. Images A-C were taken with different fluorescent filters, whereas images D-E are superpositions of images A-C in different combinations. **A**. Diffused localisation of eGFPSsbA fusion protein. **B**. DNA stained with Syto 42 dye. **C**. Localisation of mCherry-SsbB. **D**. Overlap of A and B, eGFPSsbA and DNA, blue arrows point to areas where the same localisation pattern is observed. **E**. overlap of B and C, mCherrySsbB and DNA, white arrows indicate SsbB foci between DNA **F**. ovelap of A and C, eGFPSsbA and mCherrySsbB, nearby localisation (but not colocalisation) is indicated with yellow arrows.

When mCherrySsbB was expressed from a plasmid vector promoter (pANT849), obvious delay in growth and unusually short and thick branches of vegetative mycelia were observed. SsBA was present as regularly positioned bright foci in areas where SsbB was accumulated (Figure 3.20). Moreover, DNA was unusually distributed in vegetative hyphae. Between each nucleoid there was bright SsbA focus (Figure 3.20 E). The SsbA foci were averagely spaced 2.2 µm (in range from 1.2 to 4.0 µm). Even when over-expressed, SsbB was usually localised between chromosomes (Figure 3.20 F), although partially overlapping with DNA probably due to the high amount of SsbB protein.

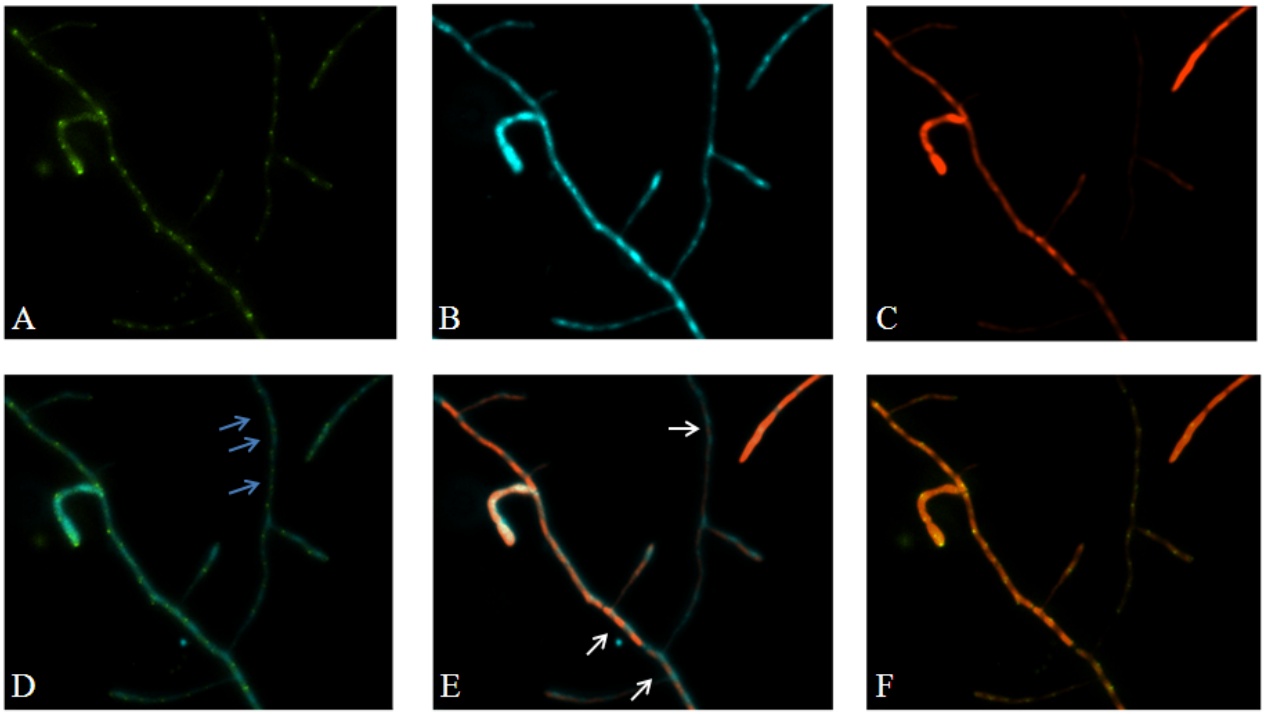


Figure 3.20. Localisation of SsbA and SsbB in *S. coelicolor* overexpressing mCherrySsbB. Images A-C were taken with different fluorescent filters, whereas images D-E are superpositions of images A-C in different combinations. **A**. Localisation of eGFPSsbA fusion protein. **B**. DNA stained with Cyto42. DNA is unusually distributed throughout vegetative hyphae. **C**. Localisation of mCherry-SsbB **D**. Overlap of A and B, eGFPSsbA and DNA, bright foci are present between segregated nucleoids **E**. overlap of B and C, mCherrySsbB and DNA, SsbB is found beween nucleoids, although they are partially overlapping **F**. ovelap of A and C, eGFPSsbA and mCherrySsbB. SsbA are possibly colocalizing with SsbB.

3.4. Binding properties of paralogous SSB proteins

3.4.1. Electrophoretic mobility shift assay

Cooperative binding mode of SSB proteins are important for DNA replication (Shereda, Kozlov et al. 2008). To reinforce finding that SsbA is a key player during DNA replication the binding efficiency of the SsbA and SsbB proteins to the long ssDNA fragment by electrophoretic mobility shift assay (EMSA) was evaluated. Binding was tested using the circular ssDNA of Φ1748 (5386 nt). When increasing concentrations of SSB proteins were incubated with a fixed concentration of Φ1748 a progressive decrease in the mobility of circular ssDNA could be seen for both proteins (Figure 3.20 A). However, a moderate concentration of salt (100 mM NaCl) significantly reduced cooperative binding of SsbB (Figure 3.21 B). ssDNA binding activity of truncated version of SsbB (SsbB$_{\Delta C}$) was determined as well (Figure 3.21 C). The results confirm that SsbB$_{\Delta C}$, in spite of lack of C -terminal

region, still possesses similar pattern of DNA binding efficiency as SsbB. Nevertheless, the binding was weaker at high SsbB/ssDNA ratio than those observed for SsbB.

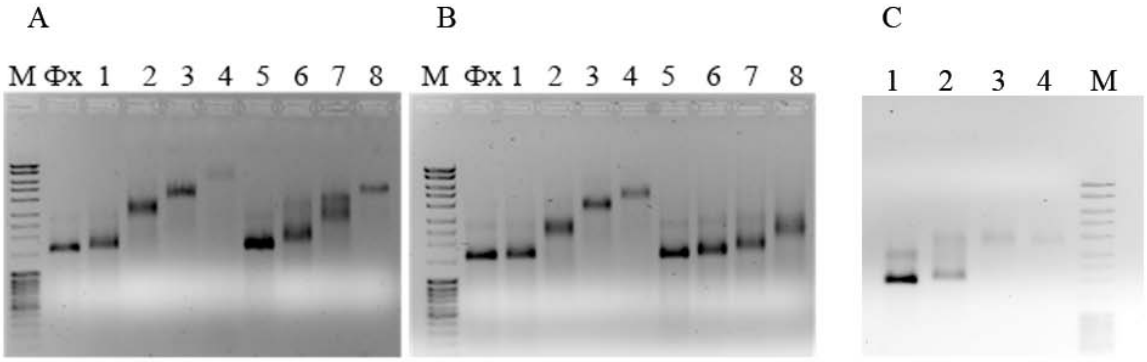


Figure 3.21. Electrophoretic mobility shift assays. Binding of SSB proteins to ΦX174 DNA was examined by EMSA. (**A***)* Binding affinity without salt or (**B**) with 100 mM NaCl added to the reaction mixture. Lanes 1-4 contained SsbA protein at the following concentrations: 0.25 µM, 1.25 µM, 2.5 µM, 5 µM, and lanes 5-8 contained SsbB with the following concentrations: 0.25 µM, 1.25 µM, 2.5 µM, 5 µM. (**C**) ssDNA binding activity of truncated version of SsbB (SsbBΔC). Lanes 1-4 contained SsbBΔC at the following concentrations 0.25 µM, 1.25µM, 2.5 µM, 5 µM. No salt was added to the reaction

3.4.2. Effect of disulphide bridges in SsbB on ssDNA binding

Recently the 3D structure of SsbB from *S. coelicolor* has been published (Paradzik, Ivic et al. 2013). This structure revealed the presence of S-S bridges between SsbB dimers. Since the presence of disulphide bonds was not previously described for SSB proteins, the effect of reducing agent on SsbB binding affinity was examined. The presence of S-S bridges was detected by SDS-PAGE or Western blot (Figure 3.22 A and B) using SsbB purified from *S. coelicolor* and *E. coli* cells. The specificity of protein band at the expected position of SsbB dimer was confirmed by Western blotting. The dimer structure of SsbB was disrupted upon additon of DTT (100 mM). In addition, it was shown that 10 mM DTT almost completely abolished SsbB binding to ssDNA, while affinity binding of SsbA was not changed even in the presence of 100 mM DTT (Figure 3.22 C). This confirmed that formation of disulphide bridges is not a purification artefact. Beside the fact that disulphide bridges exist both in cell lysate (not shown) and purified proteins (Figure 3.22 A and B), their formation has distinct effect on SsbB binding activity.

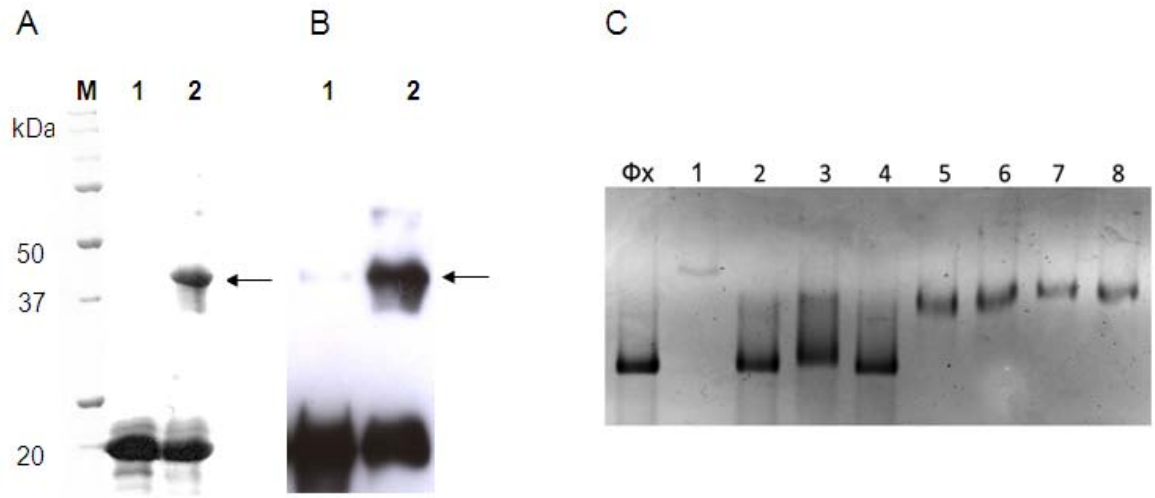


Figure 3.22. Detection of disulfide bonds in SsbB. (**A**) *S. coelicolor* SsbB isolated from *E. coli*. Lane 1. Reducing SDS sample buffer (SB) was mixed with SsbB and heated for 10' at 95 °C. Lane 2. Non-reducing SB was mixed with SsbB and heated 10' at 95 °C. 4 µg of SsbB protein was used in each lane. (**B**) Western blot analysis: SsbB isolated from *S. coelicolor* was detected with antiSSBB antibodies. Lane 1. 50 ng of SsbB in reducing buffer heated for 10' at 95 °C, Lane 2. 50 ng of SsbB in non-reducing buffer heated for 10' at 95 °C. DTT at 100 mM concentration was used as reducing agent and SsbB in dimer formation is indicated by arrows. (**C**) Binding of SSB proteins to ΦX174 DNA in the presence of DTT. SSB proteins were incubated for 10 min at 37 °C with increasing DTT concentrations prior to ssDNA binding. Lanes 1-4, 5 µM SsbB protein was mixed with the following DTT concentrations 0, 10 mM, 50 mM and 100 mM. Lanes 5-8. 1.25 µM SsbA protein mixed with the following DTT concentrations 0, 10 mM, 50 mM, and 100 mM.

3.5. Bioinformatic analyses of ssb genes and SSB proteins

3.5.1. Genomic analysis of the ssb genes in Eubacteria

All completed eubacterial genomes were surveyed for the presence and copy number of the *ssb* genes. The sequences were retrieved using genomic BLAST in NCBI database against every eubacterial phyla (August 2013 5100 completed genomes available). For the phylum Proteobacteria, BLAST search was performed against lower taxonomic units (classes/subdivisions). The datasets were examined in order to exclude extra chromosomal SSB sequences, as described in Materials and Methods and regularly updated. Finally, one representative species from each sequenced genus was used for further analyses. In such a way dataset consisting of many highly similar sequences originating from closely related species was reduced, and bias deriving from over-represented species in the dataset (e.g. human pathogens) was minimised. Summary of the presence of paralogous *ssb* genes and their location on the bacterial chromosomes across Eubacteria is presented in Table 3.3. All bacterial genomes examined in this study (558 in total) displayed 965 putative SSB protein coding genes, among which 326 genomes (58 %) had only one *ssb* copy. From the total number of analysed genomes, 84 species had three (56 species) or more (28 species) *ssb* genes. Genomes with large number of *ssb* genes (five or more) were relatively rare, only nine such genomes were found within our datasets and occasionally these copies were greatly reduced in size. Phytoplasma is a good example of this group; Aster yellows witches'-broom phytoplasma has 11 *ssb* copies and one of their *ssb* products is reduced to only 59 amino acids. All other copies are longer, but none of them exceed 115 aa.

Correlation between the number of *ssb* gene numbers on the chromosome and taxonomic affiliations could not be found (Figure 3.23). However, within specific phyla, most species often possess the same number of *ssb* genes (Table 3.3). For example, in phylum Actinobacteria, most species had two genes (39), although significant number of species (19) had one or three genes (Table 3.3). In contrast, Firmicutes had approximately same number of species with one and two *ssb* genes and most divisions of Proteobacteria have predominantly one *ssb* gene. Similarly to this, lower taxonomic units (orders and families) also displayed variability in copy number of *ssb* genes on the chromosome. However, members of the particular genus usually had the same number of *ssb* genes. Exception to this is β proteobacteria *Bordetella duttoni*, which has six *ssb* copies on the chromosome, while other sequenced members of the same genus (13 in total) have variable number of *ssb* genes ranging from one to four.

In summary, to demonstrate variations in *ssb* copy numbers graphically, the phylogenetic ML tree based on 16S rRNA gene sequences was constructed. The sequences for tree construction were selected from all orders using representative species with one, two or more *ssb* genes. As described above, the obtained tree (Figure 3.23) clearly shows random distribution of bacterial species with the additional number of *ssb* genes across phyla.

Table 3.3. Comparative analysis of the *ssb* genes located on the bacterial chromosomes from different phyla. Representative species were selected among sequenced bacterial genomes as described in Material and methods. Position of *ssbA* gene other than in ribosomal operon is highlighted in gray

Phylum and number of representative species		Number of species	Number of *ssb* genes	Chromosomal position of *ssb* genes-neighbour genes
Actinobacteria 64		6	1	*rpsF-ssbA-rpsR* (97 %) *ssbB*- ABC transporter (63 %)
		39	2	
		13	3	
		3	4	
		3	5	
Firmicutes 75		36	1	*rpsF-ssbA-rpsR* (95 %)
		25	2	
		8	3	
		5	4	
		1	5	
Thermus/Deinococcus 6		5	1	*rpsF-ssbA-rpsR* (100 %)
		1	2	
Chloroflexi 9		5	1	*rpsF-ssbA-rpsR* (88 %)
		2	3	
		2	4	
Cyanobacteria 26		21	2	*ssbA*- ABC transporter (38 %) *ssbB- MreB* (58 %)
		3	3	
		2	4	
Spirochaetes 7		3	2	*rpsF-ssbA-rpsR* (100 %)
		3	3	
		1	6	
Tenericutes 8		3	1	*rpsF-ssbA-rpsR* (100 %) *
		3	2	
		2	>10	
Chlorobi 5		5	2	*ssbA*DNApolIII(Δ) (80 %) *rpsF-ssbB-rpsR(* 100 %)
Bacterioides 57		16	1	*ssbA MutY* (68%)
		29	2	
		6	3	
		4	4	
		1	7	
Chlamydiae 6		4	1	*ssbA*- Leucine aminopeptidase (67 %)
		2	2	
Aquificae 8		8	1	*rpsF-ssbA-rpsR* (100 %)
Others (underrepresented) 49		39	1	*rpsF-ssbA-rpsR* (*rplI*) (79 %)**
		8	2	
		2	3	
Proteobacteria	δ 24	17	1	Not conserved
		4	2	
		3	3	
	ε 11	10	1	*rpsF-ssbA-rpsR* (100 %)
		1	3	
	α 65	46	1	*ssbA-UvrA* (53 %), Fe-transport (9 %), transglycosilase(14%)
		13	2	
		3	3	
		2	4	
		1	5	
	β 44	19	1	*ssbA-UvrA*(75%) *ssbB-TopB*(48%)
		19	2	
		5	3	
		1	5	
	γ 95	72	1	ssbA-UvrA/MFP(72 %)
		17	2	
		6	3	

* species with multiple *ssb* copies have great reductions in *ssb* size

**members of Verrucomicrobia/Planctomyces group lack *rpsR* gene from the *rpsF-ssbA-rpsR* operon

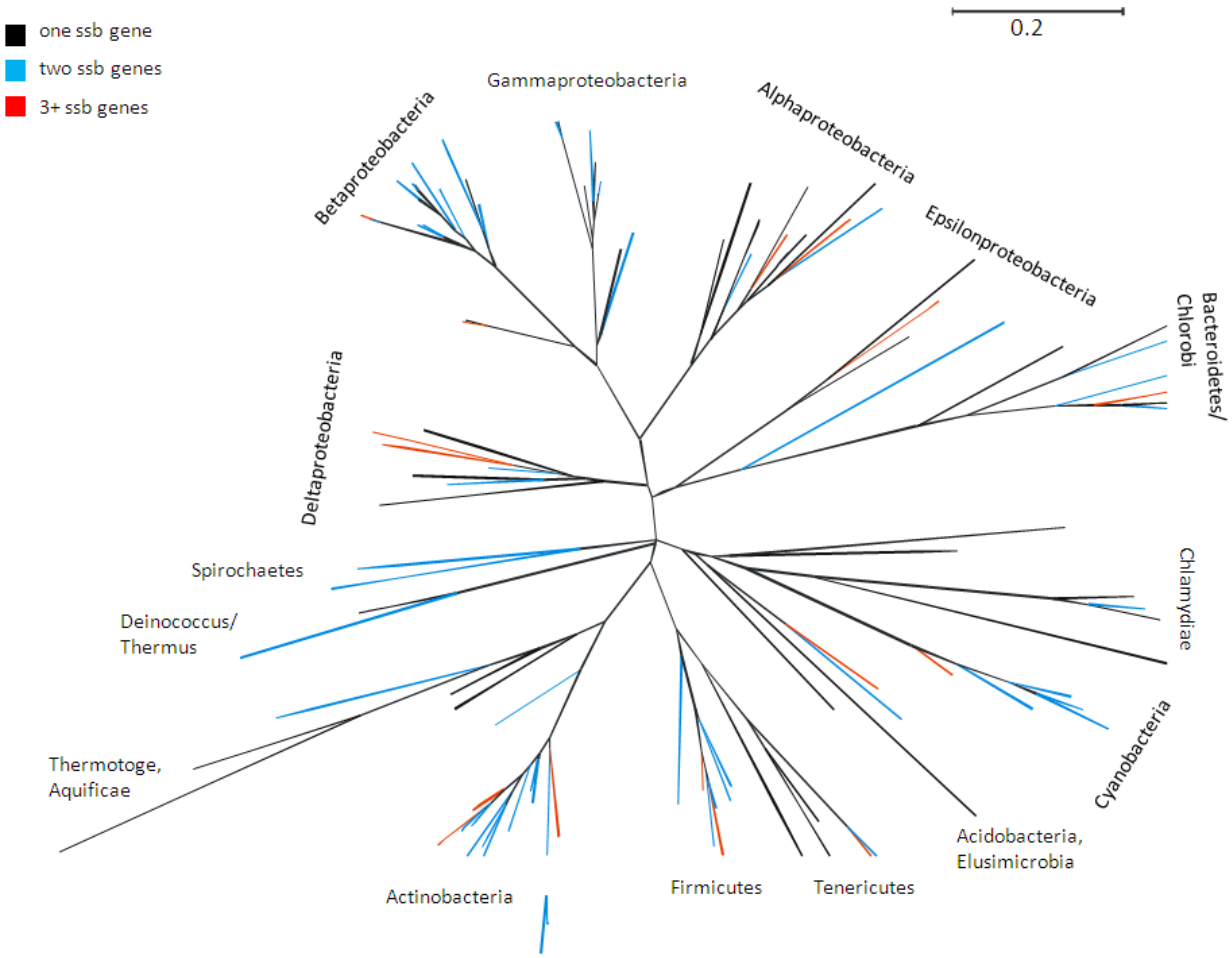


Figure 3.23. ML tree constructed with the selected 16S rRNA gene sequences from different Eubacterial phyla. Each terminal branch of the phylogentic tree represents one species that is coloured according to the number of *ssb* genes on its chromosome.

It was reported previously that many *ssb* genes are often encoded with a ribosomal protein operon (Lindner, Nijland et al. 2004). In this study position of *ssb* genes was examined on the large set of sequences. As shown in the Table 3.3, most of the bacterial species have *ssbA* gene (annotated as explained in Materials and Methods) positioned within ribosomal protein genes (*rpsF-ssbA-rpsR*). Analysis of 28 bacterial phyla showed that 23 phyla (83 %) possess conserved syntheny of *ssbA* gene, i.e. this gene was located within ribosomal operon (Figure 3.24). All exceptions found among the following phyla: Cyanobacteria, Bacteroides, Chlamydia and certain classes of Proteobacteria (with exceptions of ε division) are listed in Table 3.3 and shown schematically (Figure 3.24). Contrary to *ssbA*, the position of *ssbB* gene is generally not conserved, with the exception of phylum Actinobacteria, order Actinomycetales. In this order 56 out of 58 genomes have multiple copies of the *ssb* genes and most of them have *ssbB* placed near putative ABC transporter (54 %). Another example of the syntheny of *ssbB* gene is found in β proteobacteria in which this gene is predominantly located near *TopB* gene (Table 3.3). In only one case, gene labelled as *ssbB* from phyla Chlorobi, was located within ribosomal operon, since three out of four criteria were fulfilled to annotate this paralogous gene as *ssbB,* as described in Materials and methods.

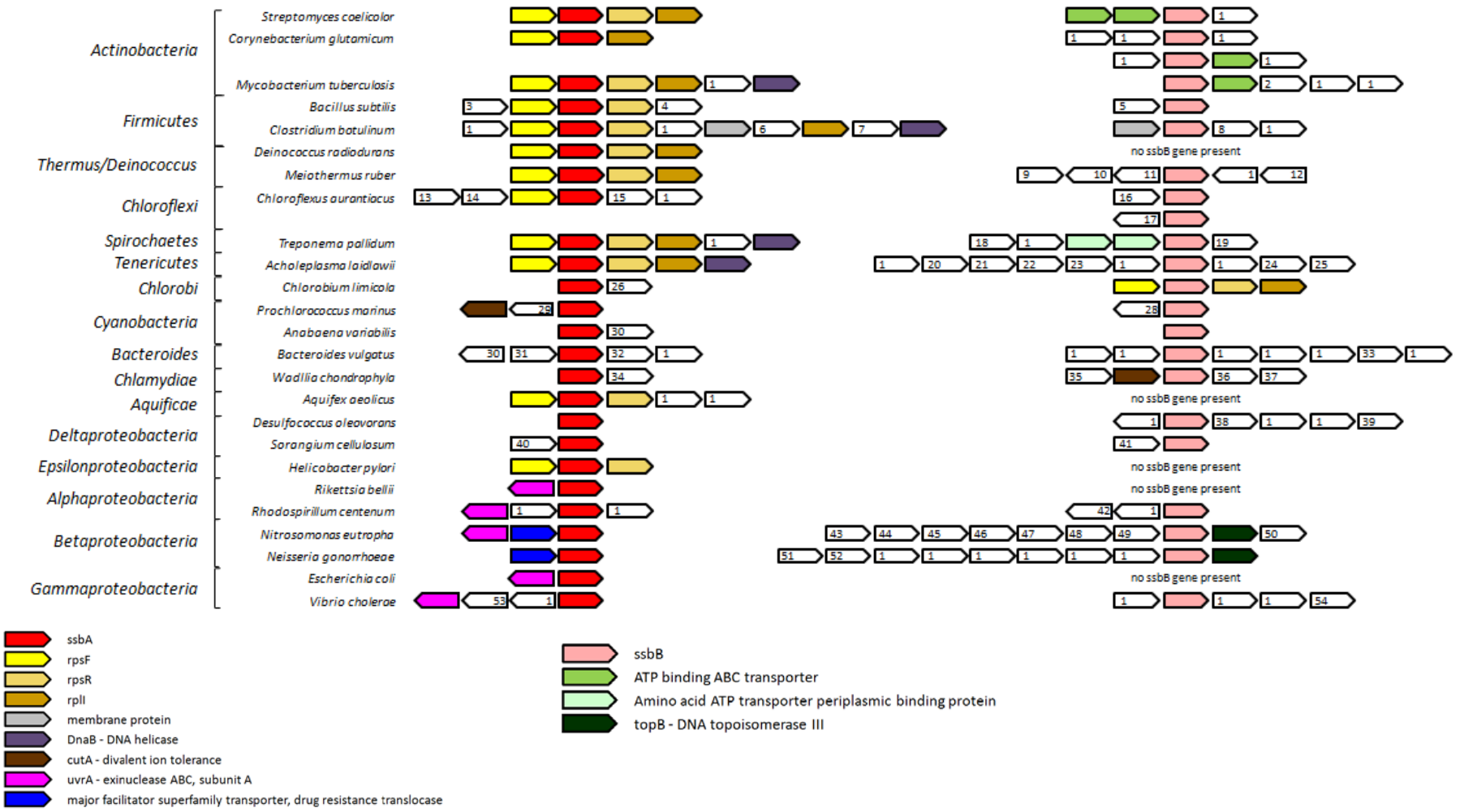


Figure 3.24. Genomic neighbourhood of *ssb* genes in different eubacterial phyla. The left side of the panel represents positions of the *ssbA* gene while the right side represents *ssbB* genes. Coloured boxes represent genes appearing more than once in the scheme. The white boxes represent genes encoding for hypothetical proteins and those appearing only once. These genes are listed in supplement.

3.5.2. CLANS clustering of bacterial SSB sequences

CLANS programme was applied for analysis of the relationship and clustering of bacterial SSBs. This programme uses blast-all approach, i.e. it compares every sequence with all other sequences in the input file and built 3D representation based on sequence similarities. All collected SSB sequences were used to run CLANS analyses. As shown in Figure 3.25, SsbA and SsbB protein sequences do not form two separate clusters. Most of the bacterial SsbA sequences cluster according to their taxonomic affiliation. Actinomycetales and Bifidobacteriales are displaced forming separate cluster while other SsbA sequences retrieved from the same phyla are clustered closer to the other bacteria. SSBs belonging to Thermus/Deinococcus group (most representatives with one SSB) clustered separately as well (visible in 3D graph). This was not unexpected since these proteins possess unique primary structure. They function as homodimers and show unusual domain arrangement (Bernstein, Eggington et al. 2004). Phyla Proteobacteria and Firmicutes form distinct groups of SsbAs, with many highly similar sequences, as can be seen by high concentration of red and blue dots. There are few distinctive groups of SsbB sequences: 1.SsbB from Cyanobacteria, 2.SsbB from Actinomycetales, 3. SsbB from α, β, and γ Proteobacteria, 4. SsbB from α Proteobacteria. 5. SsbB from Clostridia family of Firmicutes. Interesting clustering is observed for both SsbA and SsbB from Cyanobacteria. All analysed Cyanobacteria have at least two *ssb* genes clustering into two groups; however none of them fulfils all criteria for a clear annotation into SsbA or SsbB group. After CLANS analysis sequences annotated as SsbA are positioned in a separate cluster displaced from the centre of the graph, while sequences annotated as SsbB are positioned closer to SsbAs from the other bacteria.

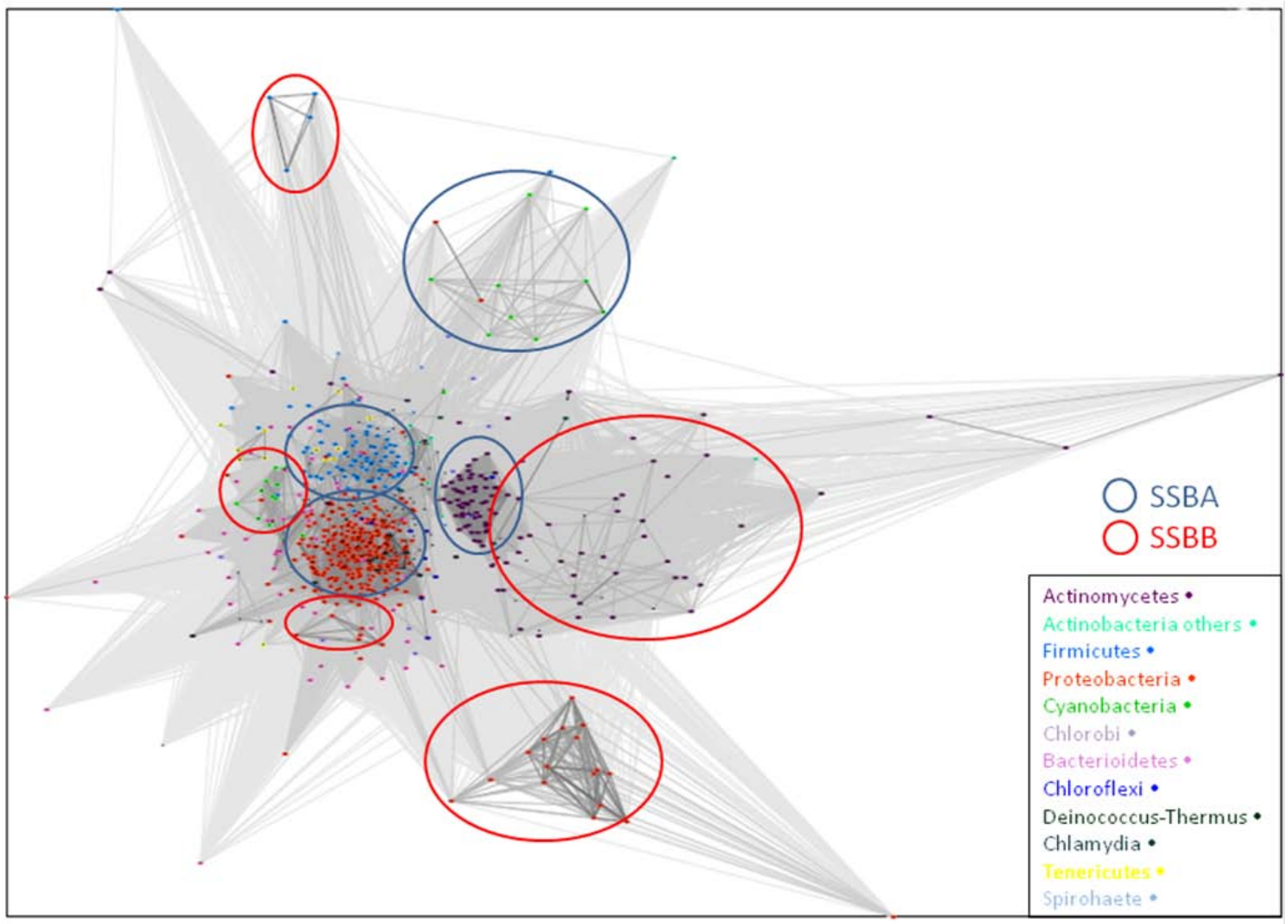


Figure 3.25. Clustering of SSB protein sequences from Eubacteria. Each dot represents one sequence. Sequences from different phyla are differently coloured, while groups of SsbA and SsbB sequences are marked with blue and red circles, respectively.

3.5.3. Phylogenetic analysis of SSB proteins from Eubacteria

Clustering of SSB sequences obtained by CLANS provided the first rough insight into the relationship of bacterial SSBs. In the next step phylogenetic trees were built using sequences from the individual phyla. Since our model bacterium belongs to Actinobacteria, the first SSB tree was constructed for this phylum (Figure 3.26). Representative paralogous genes from each suborder belonging to five orders of Actinobacteria, and two paralogous SSBs from *B. subtilis*, were used as an outgroup to construct maximum likelihood tree. The phylogenetic analysis (Figure 3.26) demonstrated that SsbB and SsbA are clustered separately within the orders Actinomycetales and Bifidobacteriales. This exeption is SsbB from *F. alni,* positioned outside the SsbB cluster. Branch lengths indicate that cluster contiaing *S. coelicolor* SsbB protein is more heterogeneous than cluster with SsbA; that might be indicative of the essentiality of the latter cluster of proteins. This is in agreement with CLANS clustering, where SsbB sequences are much more divergent (Figure 3.25). Two species with three SSBs (*Stackebrandtia nausensis* and *Corynebacterium glutamicum)* were also included in the tree. As shown (Figure 3.26), their SsbC proteins were clustered differently; *C. glutamicum* SsbC was clustered within conspecific SsbA protein group, while SsbC from *S. nausensis* was positioned outside the cluster of Actinomycetales SSB proteins. SSBs from the order Coriobacteriales make separated, well supported cluster (see aLRT values at nodes); however their SsbB proteins branch with SsbA as recent duplicates. The two SSBs from *B. subtilis* (Yadav, Carrasco et al. 2012) representing an outgroup were positioned together with high confidence (0.99).

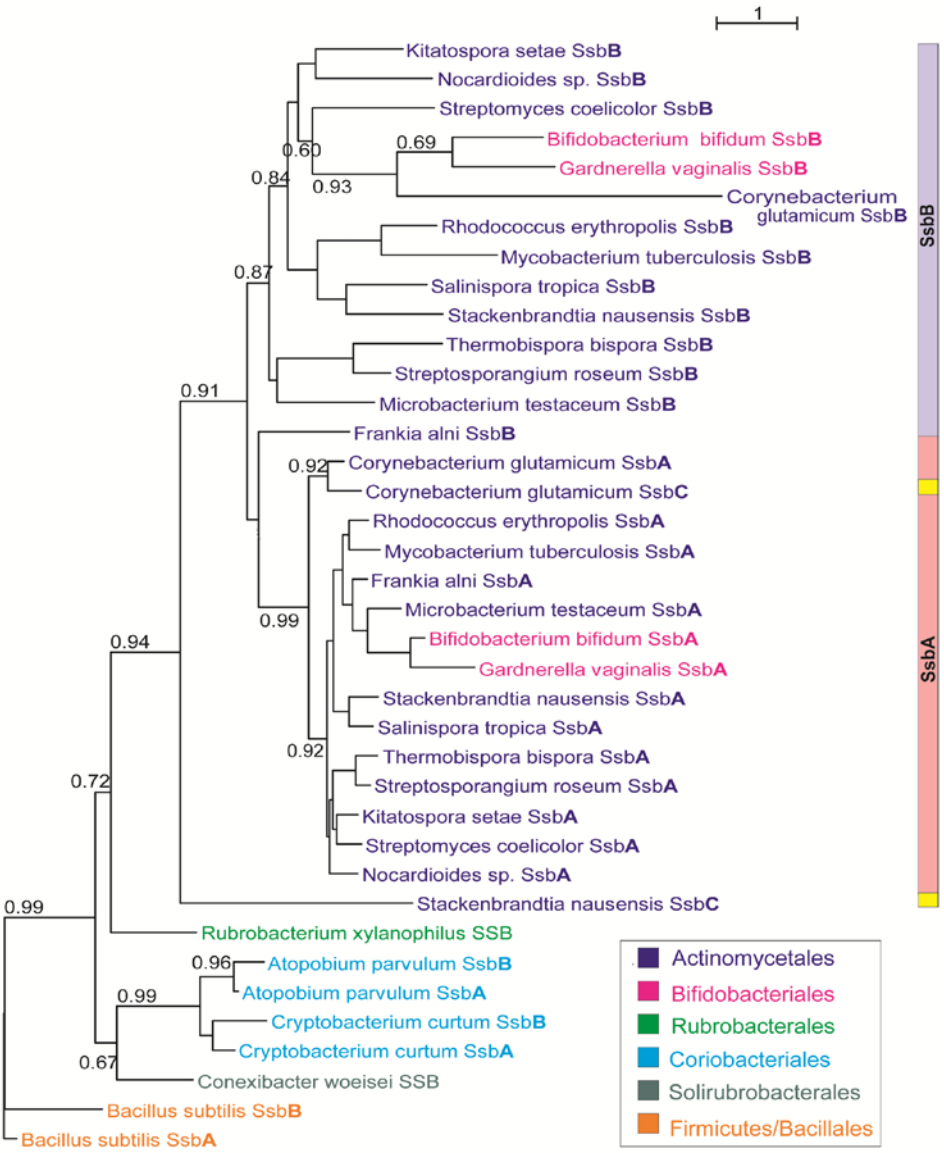


Figure 3.26. Phylogenetic tree of SSB proteins based on multiple alignment of full length protein sequences from phylum Actinobacteria. The representatives of different orders are coloured as shown in the legend and aLRT values are shown for the main branches. Additional Ssbs (SsbC) are marked by yellow on the bar on the right. *B.subtilis* SsbA and SsbB were used as an outgroup.

The tree obtained with SSB sequences from Cyanobacteria is represented in the Figure 3.27. This tree was built from SSB sequences belonging to five orders of this phylum. There is a clear grouping of SsbA and SsbB sequences; moreover, SsbB probably adopted important role in this phylum since it was found in every species sequenced so far. *Gloebacter violaceus*, species with four SSBs, was included in the tree as well. The third and fourth SSB are grouped with either SsbA or SsbB, indicating more recent duplication event. Two SSBs from *S. coelicolor* were included in this and subsequent phylogentic analysis.

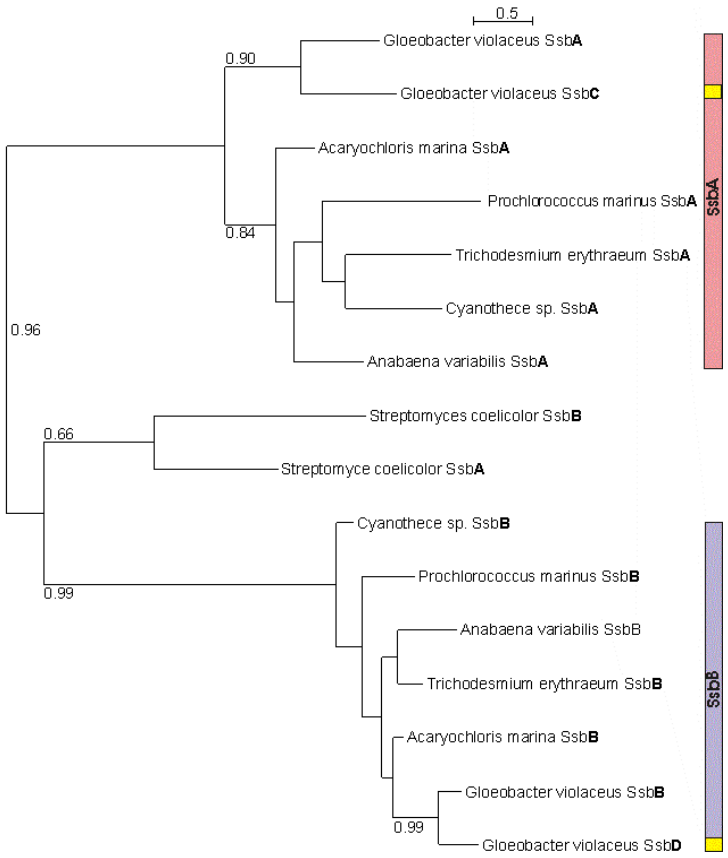


Figure 3.27. Phylogenetic tree of SSB proteins based on multiple alignment of full length protein sequences from phylum Cyanobacteria. Additional Ssbs are marked by yellow on the bar on the right. aLRT values are shown for main branches.

As seen from CLANS clustering (Figure 3.25), two closely related phyla (Gupta and Lorenzini 2007) Bacteroidetes (anaerobic pathogens) and Chlorobi (green sulphur bacteria) have very divergent SSB sequences, and do not form clusters of SsbAs and SsbBs. Representatives from Bacterioidetes, members of three orders and four sequenced Chlorobi genera were used for the construction of phylogenetic tree. Order Bacteroidales was excluded from this analysis, since its SsbBs were clustered with SsbAs in preliminary analysis and could not contribute to the clarification of complex phylogenetic relationships in this group. Results of this phylogenetic analysis are shown in Figure 3.28. There are many examples of multiple *ssb* copies in Bacteroidetes, including species *Zunonwangia profunda* with seven *ssb* genes. Although CLANS analysis did not show any significant clustering, three distinct groups of SSBs have well supported branching (Figure 3.28; groups SsbA, SsbB and SsbC). Position of the SsbA and SsbB group of this phylum indicates that duplication event occurred after separation from the last common ancestor with Chlorobi, while SsbC group formed separate cluster. Chlorobi formed strongly supported groups of SSBs, as indicated in Figure 3.28, but

separated from Bacteroidetes. In addition to this, tree indicates that duplications of *ssb* genes are very common event.

Figure 3.28. Unrooted phylogenetic tree of SSB proteins based on multiple alignment of SSB protein sequences from phyla Bacteroidetes and Chlorobi. The representatives of different orders are marked in different colour (see legend). aLRT values are shown for main branches.

Phylogenetic analysis of SSBs from phylum Firmicutes is presented in Figure 3.29. The biological role of SsbB protein is well described for two representative members of this phylum. It was shown that these proteins play important role in natural competence (Lindner, Nijland et al. 2004; Attaiech, Olivier et al. 2011). However, more than half of analysed Firmicutes possess only one *ssb* gene. The phylogenetic analyses of the SSBs belonging to Firmicutes are built using sequences from the two classes: Clostridiae and Bacilli. As seen from the branch lengths (Figure 3.29) SsbA from Bacilli are highly homogenous, whereas SsbA from Clostridia are more divergent. In addition, there is group of

SsbBs from Clostridia separated from all other SSBs. Blast analysis did not show any significant similarity of this protein with other protein sequences outside of this class. SsbBs from the orders Bacillales and Lactobacillales (both belong to class Bacilli) are clustered separately. However, there are some SsbB sequences which do not belong to these well defined groups and are positioned outside main branches. There are also multiple examples of SSB expansions - such as in *Ruminococcus albus* (Clostridia), *Macrococcus caseolyticus*, *Staphylococcus aureus* (Bacillales), or *Lactococcus lactis* (Lactobacillales).

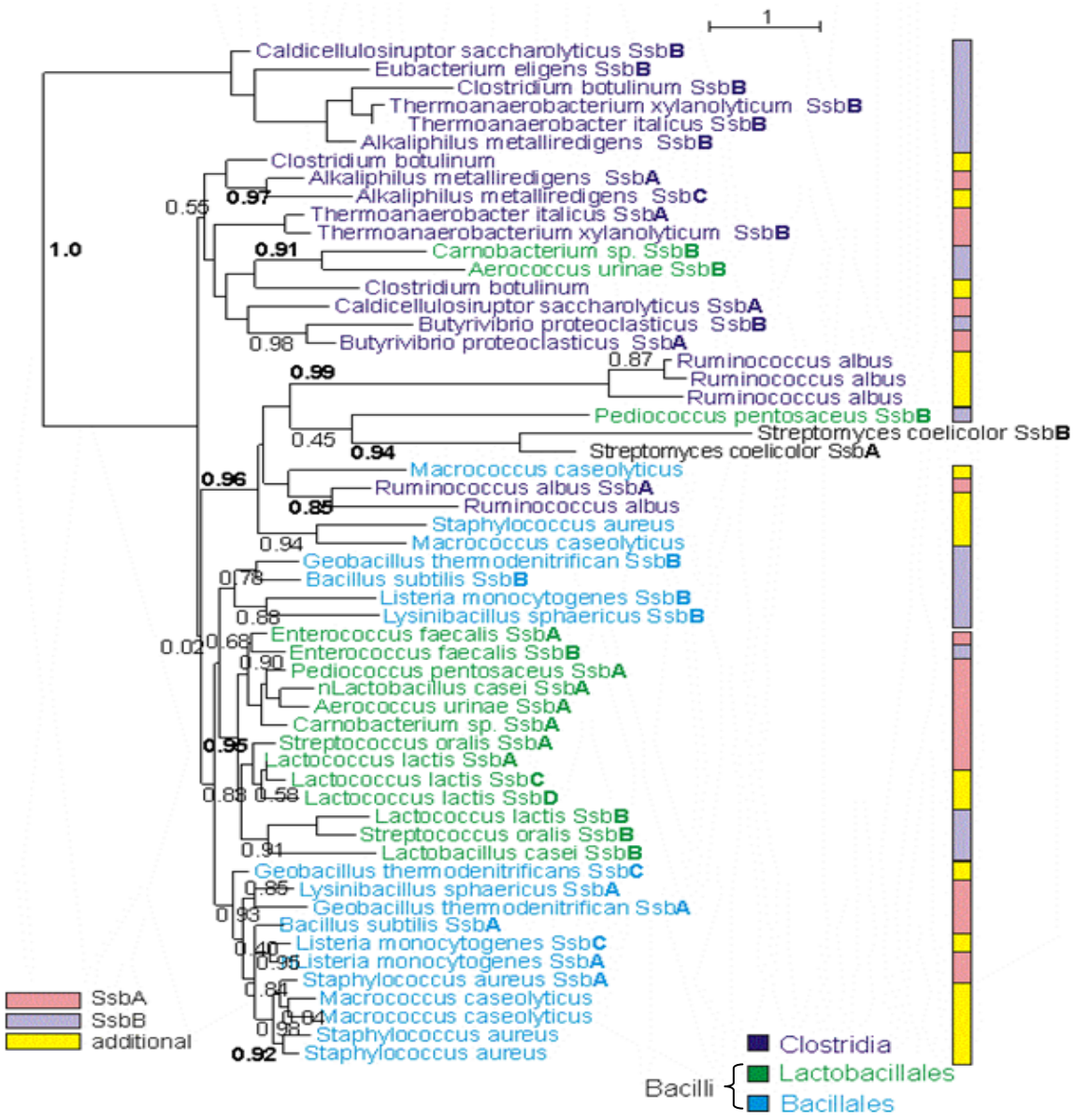


Figure 3.29. Phylogenetic ML tree of SSB proteins from phylum Firmicutes based on multiple alignment of SSB protein sequences. The representatives of different orders are coloured acording to the legend. aLRT values are shown for the main branches.

Proteobacteria the most studied Eubacteria phylum and is represented by extremely large number of sequenced genomes. For example, there are 61 completed genomes of various *Escherichia coli* strains, and many more in a process of finishing (NCBI database). The Proteobacteria are divided in five sections (classes): α, β, γ, δ and ε. In preliminary analyses, each division was analyzed separately. Most α proteobacteria have only one *ssb* gene in their chromosomes (for example order Rickettsiales, or genera *Rhizobium* and *Agrobacterium*). On the contrary, other orders (Caulobacterales, Parvularculales, Rhodospirillales, Sphingomonadales and other families of order Rhizobiales) have variable *ssb* numbers. In division β proteobacteria there is approximately same number of genera with one or two copies of *ssb* genes. The γ proteobacteria is a class of several medically, ecologically and scientifically important groups of bacteria. 72 out of 95 analysed genera have only one *ssb.* Occasionally when large number of sequenced strains from the same species is available (e.g. the model bacterium *E. coli*), certain variability could be observed among different strains. Even though most *E. coli* strains have a single *ssb* gene, there are substantial number of strains with two *ssb* genes, and even one example with five *ssb* genes. Furthermore, there are 31 sequenced strains of *Salmonella enterica*, and approximately 50 % have a single *ssb* gene, while others have two *ssb* genes. Therefore, in these pathogenic bacteria the number of *ssb* genes is not uniform even at the species level. The *ssb* genes from δ proteobacteria do not share any conserved syntheny, which is unusual for *ssbA* genes (Table 3.3). In the preliminary trees all SsbBs from δ proteobacteria branched as they are result of recent duplications. ε proteobacteria are only group of Proteobacteria in which all SSBs have conserved position within ribosomal genes. Only one genus out of 11 is found to have three SSBs, while all the others have one.

Phylogenetic tree of Proteobacteria SSB sequences was built to resolve relationships among SSBs from different divisions. Sequences included in this analysis were taken from the species with multiple *ssb* copies that do not branch together with SsbAs in the preliminary individual phylogenetic trees. All representatives from subdivisions α, β and γ which satisfy this criterium were used for analyses. Subdivisions ε and δ were excluded since they had only one SSB or their multiple SSB sequences branch together and therefore would not contribute to clarification of SSB relationships within Proteobacteria. Representatives of three classes are indicated by different colours (Figure 3.31). As depicted, most of the species showed clear grouping of SsbA (one cluster) and SsbB (two clusters) at the subdivision level with the exception of γ proteobacteria. As can be observed, many SsbBs from α proteobacteria are branched together which is supported by high aLRT values (this group is indicated by the brace). Interestingly, there are few multiple expansions of sequences within this cluster (for example there are four SsbBs from *Erythrobacter litoralis)*. SsbBs in a large cluster

marked as “TopB” group, are present mostly in β proteobacteria species. SsbB from three species belonging to γ proteobacteria and only one species from α proteobacteria are also found inside this group. Beside this, as shown in Figure 3.30 there are certain numbers of sequences from all three subdivisions which do not have clear grouping and whose branching is not supportive enough (below 0.9).

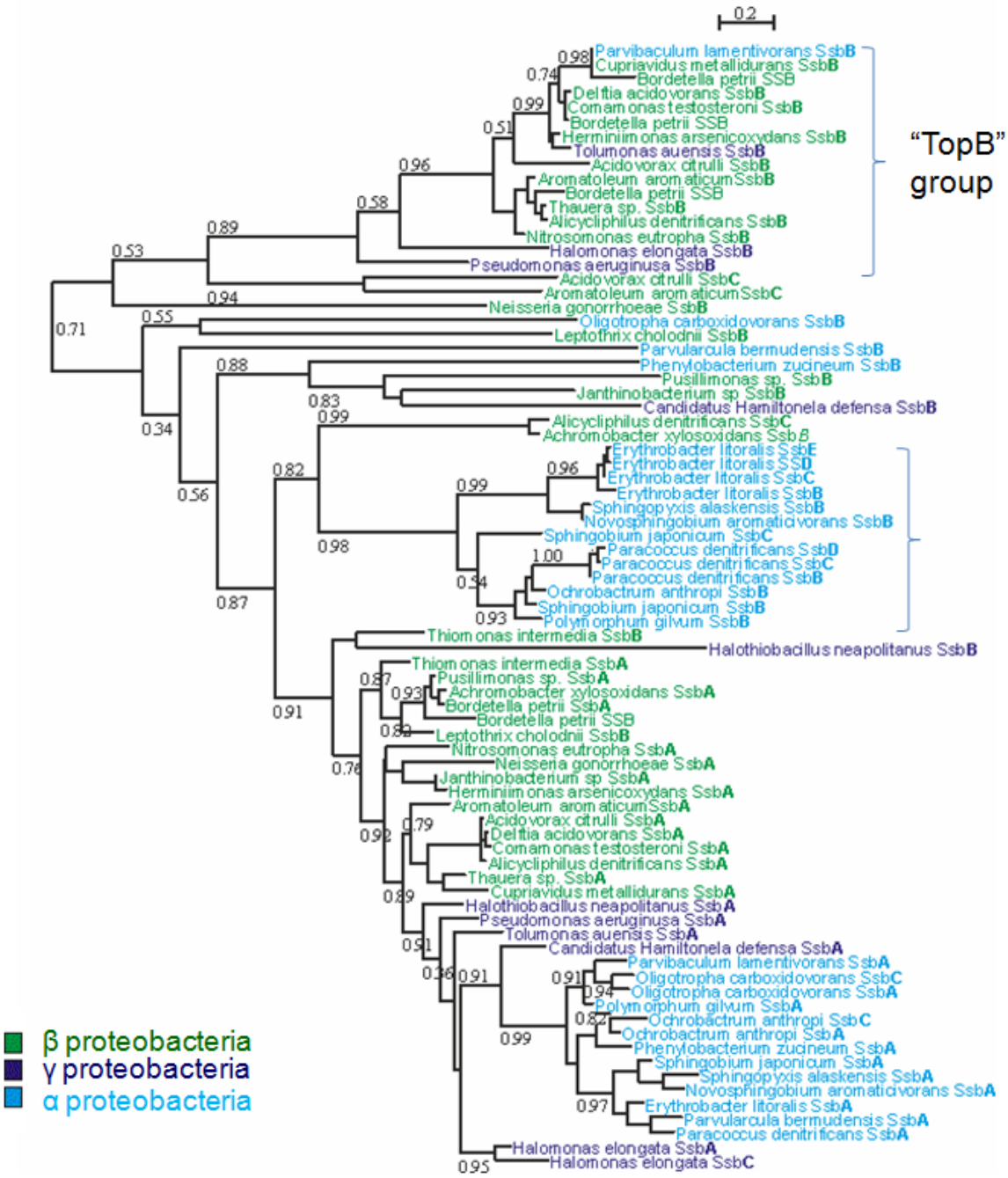


Figure 3.30. Unrooted proteobacteria ML tree. Subdivisions are marked in different colours. aLRT values are shown for the main branches. Two distinctive groups of SsbB proteins are indicated by the braces.

As shown in figures 3.26-3.30, phylogenetic analyses were firstly performed using representative members of separate phyla. In order to get better insight into the phylogenetic relationship of SSB proteins from different phyla, a single tree was built using selected members from previous analyses.

Selected species were those showing clear grouping of SsbA and SsbB at the phylum level. In addition, this dataset was expanded with the members of phyla Spirochaetes and Chloroflexi. Other phyla, such as Aquificae were not included, as those species mostly have only one *ssb* gene copy and this study was focused on analyses of paralogous SSBs. Furthermore, all SSBs that did not show specific grouping in preliminary trees were also excluded (i.e. all species with two or more SSBs that were clustered as recent duplicates). Distribution of the SSB groups in the tree based on sequences from seven eubacterial phyla is presented in Figure 3.30. Members of the SsbA group of the same phylum are in all cases clustered together. On the contrary, members of the SsbB groups are clustered either within phyla (see Actinobacteria or Spirochaetes) or form distinct clusters. SsbB groups did not include members of more than one phylum, for example SsbB group formed by the members from Bacteriodetes.

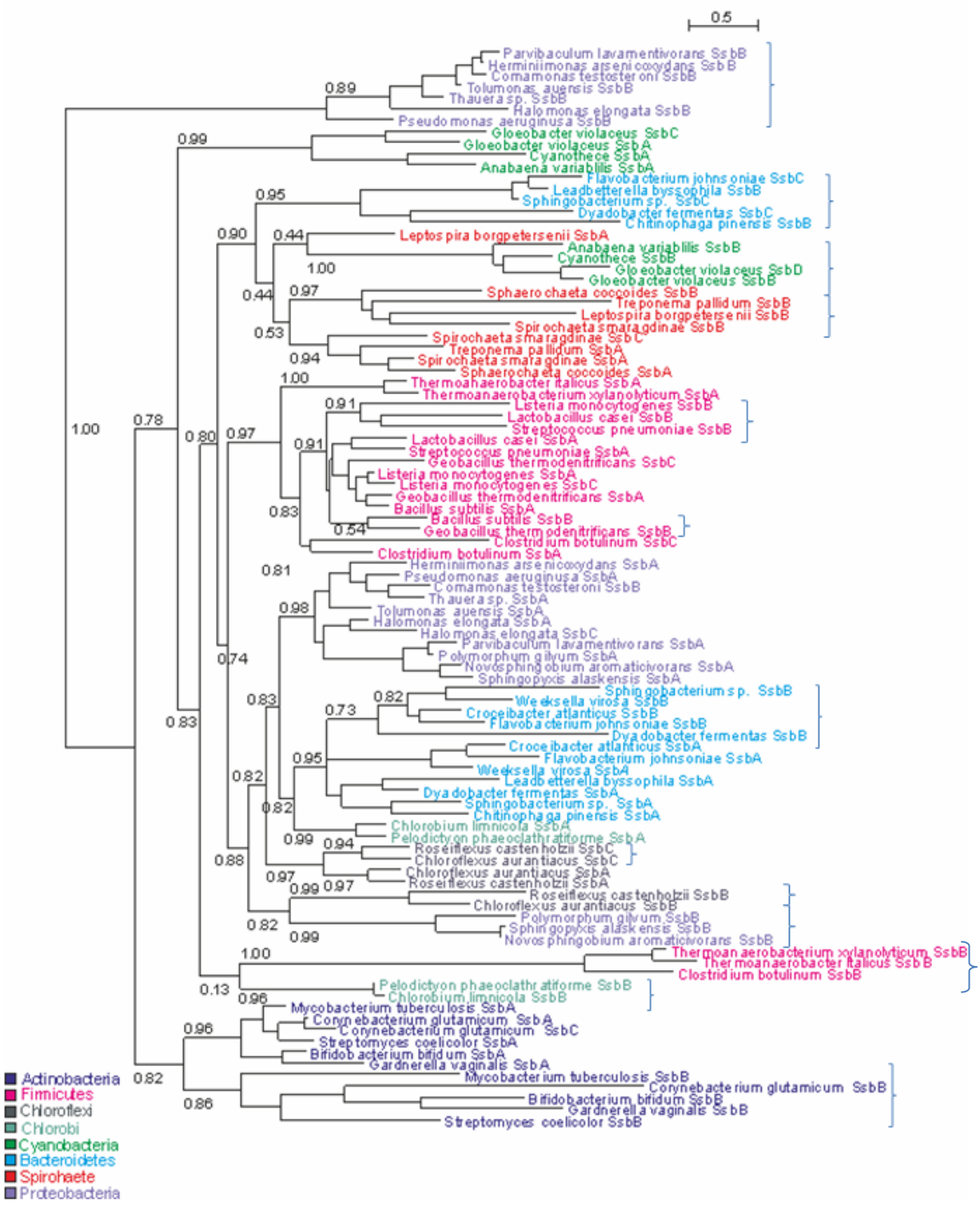


Figure 3.31. Unrooted maximum-likelihood tree of SSB proteins from different phyla. Each phylum is coloured diferently as indicated in the figure legend. Distinctive groups of SsbB proteins are marked with braces. aLRT values are shown for the main branches.

In summary, phylogenetic analysis indicated that in most phyla SsbB sequences have polyphyletic origins. All obtained trees have some common features: in most cases SsbA sequences are clustered together with short branches which indicate slow evolution. Whereas, distribution of the SsbB sequences are more complex and can be divided into: (**i**) SsbBs highly similar to SsbA, as a product of more recent duplications are clustered near SsbA (**ii**) SsbBs that form a distinct group positioned within certain phyla (e.g Actinobacteria), resulting from the duplication of SsbA after phylum divergence (**iii**) SsbBs grouped outside of phylum branch. Described patterns could be seen in all analysed phyla.

3.5.4 Analysis of the recent ssb gene duplications

Recent duplicates are good model to investigate evolution forces shaping the structure of the gene. In this study, phylogentic analysis (branch length of SsbB proteins) showed that SsbB proteins are more divergent. In addition, sequence analysis of SsbBs pointed out greater reduction of the large parts of C-terminal region. Moreover, there are examples where SSB proteins display high similarity (above 80 %). Those are probable candidates for recent duplication event. Therefore to examine selective pressures on recently duplicated *ssb* genes PAL2NAL programme was used. PAL2NAL is converting multiple sequence alignment of proteins and the corresponding DNA (or mRNA) sequences into a codon alignment. Furthermore, it is calculates synonymous (d_S) and non-synonymous (d_N) substitution rates in DNA sequences alignment. d_S is a measure of relative age of the duplication event and when its value is above threshold (>3), such paralogues pairs have to be excluded from further analysis since this is indication that duplication is not a recent event (Kondrashov, Rogozin et al. 2002).

Firstly, the full length proteins were compared by PAL2NAL tool. d_N/d_S ratio was always below 1, implying/indicating that selection on paralogous protein was negative, i.e. non-synonymous changes were not favourable. However, comparison of SSB protein domains has revealed that C-terminal domains evolve much faster as saturation in Ct is achieved while d_S for Nt are still much below threshold in many cases (Table 3.4). Statistical analysis confirmed that these differences were significant (P = 0.02).

As shown in Table 3.4, according to d_N/d_S ratios of given protein domains, changes in nucelotide sequences are predominantly synonymous, which means d_N/d_S ratio is <<1, even when Ct part is taken into account. Therefore, although Ct evolves faster, as can be seen from d_S value (nucleotide level), there is still negative selective pressure present at the amino acid level.

Alignment of recently duplicated SSB paralogous sequences is clearly demonstrating that large parts of C-terminal region of protein are prone to deletions or insertions. (Table 3.4, ins/del), excluding conserved acidic tail (five to ten amino acids at the SSB terminus). Interestingly, approximately 50 % of highly similar paralogous SSBs, have the addition of five or more amino acids to the acidic Ct. The copy gaining this additional amino acids is always the one that is translocated from the original position. Analyses of the more divergent paralogues SSBs have indicated that the acidic Ct is changed as well (examples are: Actinomycetales, Bacilalles, Spirohaetes, Bacteroidetes).

Table 3.4. Analyses of the predicted duplicates from various bacterial species distributed across phyla. d_N/d_S ratio as a measure of evolutionary selective pressure was calculated for N-terminal (first 100 amino acids N_t) and C-terminal region (Ct) of paralogous SSB pairs. d_S value is highlighted in light grey when it is higher for Ct than for Nt. d_N/d_S ratio is highlighted in dark grey when it is higher for Ct than for Nt.

	% positives	N-terminus		C-terminus		ins/del	
		d_S	d_N/d_S	d_S	d_N/d_S	N-terminus	C-terminus
Allochromatium vinosum DSM	71%	1,308	0,0698	61,598	0,0088	0	16
Aggregatibacter aphrophilus NJ8700	82%	2,498	0,0238	1,419	0,3382	0	44
Magnetospirillum magneticum AMB-1	86%	72,356	0,0011	21,765	0,0624	33	20
Sphingomonas wittichii RW1	91%	2,804	0,0387	71,757	0,0052	0	59
Asticcacaulis excentricus CB 48	77%	3,118	0,025	2,753	0,1384	0	12
Burkholderia ambifaria MC40-6	82%	0,469	0,009	9,563	0,0698	0	21
Laribacter hongkongensis HLHK9	74%	2,562	0,018	36,897	0,0052	0	41
Sideroxydans lithotrophicus ES-1	83%	2,453	0,039	7,302	0,0275	0	17
Anaeromyxobacter dehalogenans 2CP-C	95%	0,067	0,001	0,125	0,001	0	2
Desulfatibacillum alkenivorans AK-01	77%	2,369	0,032	48,930	0,006	0	22
Atopobium parvulum DSM 20469	75%	4,015	0,021	2,367	0,096	0	10
Listeria monocytogenes HCC23	73%	1,234	0,033	17,707	0,049	0	19
Gramella forsetti KT0803	92%	0,482	0,125	0,634	0,929	1	1
Laribacter hongkongensis HLHK9	74%	2,562	0,018	37,301	0,005	0	41
Desulfatibacillum alkenivorans AK-01	90%	2,369	0,032	48,931	0,006	0	22
Acidithiobacillus ferrooxidans ATCC 53993	98%	0,216	0,040	0,205	0,0745	0	0
Acidithiobacillus ferrooxidans ATCC 53993	76%	3,862	0,023	4,226	0,081	0	30
Bacteroides vulgatus ATCC 8482	86%	7,054	0,027	60,515	0,0061	0	30
Ammonifex degensii KC4	84%	1,203	0,080	3,929	0,102	1	4
Staphylococcus aureus subsp. aureus str. JKD	78%	2,248	0,040	2,735	0,164	0	14

4. Discussion

The Gram-positive filamentous soil bacterium *S. coelicolor* has two genes encoding for single-stranded DNA-binding (SSB) proteins, designated SsbA and SsbB. Limited studies on SsbA indicate close structural similarities to mycobacterial SSBs, acidic C-terminus (^{194}DEPPF) that mediates protein-protein interaction (Shereda, Kozlov et al. 2008) and eukaryotic posttranslational modification (Mijakovic, Petranovic et al. 2006) that might regulate the activity of this protein. On the contrary, the SsbB protein is a smaller, lacks conserved acid tail and until this study was not investigated. Both genes are positioned in the core region of the linear chromosome of *S. coelicolor* (Figure 3.1) that is defined by the presence of the essential genes. All these suggested that the SsbA protein may be indispensible for life while the retention of duplicated gene, *ssbB*, might be associated with some other but still important cellular function (Bratlie, Johansen et al. 2010).

4.1. SsbA is essential for cell survival while accurate chromosome segregation during sporulation depends on SsbB

Targeted gene disruption allowed the elucidation of the biological role of SsbA and SsbB. As described in Results, we failed to disrupt *ssbA* with Tn*5062.* The *ssbA* lies between *rpsF* and *rpsR*, encoding ribosomal proteins. Conserved synteny of these genes in distantly related bacteria and coupled transcription were reported previously (Lindner, Nijland et al. 2004). Therefore, a polar effect caused by insertional gene replacement was anticipated and a "scar" mutation at the *ssbA* locus was constructed. In spite of this strategy, deletion of *ssbA* locus was obtained only in a strain with an additional copy of the wild type gene thus indicating that *ssbA* is essential for *S. coelicolor* survival, in line with other bacteria (Meyer and Laine 1990; Lindner, Nijland et al. 2004; Shereda, Kozlov et al. 2008).

In contrast, the flanking regions of *ssbB* with neighbouring genes are over 200 bp and a polar effect was not anticipated. The *ssbB* gene was disrupted by replacing the original chromosomal sequence with an *S. coelicolor* mutagenised cosmid. Numerous recombinant colonies were obtained, indicating that this gene is not essential. However, the *ssbB* mutant (TSB01) displayed a white phenotype (Figure 3.6, Figure 3.8). Since we could restore the parental phenotype by introducing an additional copy of *ssbB* (Figure 3.6 A, strain TSB02) we concluded that whi phenotype was entirely a consequence of the mutation in *ssbB*. This was unexpected and intriguing since a role of SSB in morphological development had not been reported previously. Whi phenotype is often associated with many developmental regulators (Noens, Mersinias et al. 2005; Flardh and Buttner 2009). Several mutants displaying this phenotype have been studied over decades. Genes blocked at early stages (*whiG*, *-A*, *–B*, *–H* and *–I*) are defective in sporulation septation, while *whiD* and *whiE* mutants are

blocked in spore maturation (Flardh, Findlay et al. 1999). In addition, two mutants that belong to the SsgA like proteins (SALP family), *ssgA* and *ssgB*, also displayed a non-sporulating white phenotype, while *ssgG* showed, after 6 days of growth, a light grey appearance similar to the *ssbB* mutant. However, in this mutant the unusually large spores, three to even four times normal size, with proportionally segregated chromosomes, were detected (Noens, Mersinias et al. 2005). As shown in Figure 3.9, the *ssbB* mutant produced spore chains with a severe nucleoid segregation defect. Unequal distribution of DNA in spore chains showed that the *ssbB* mutant was impaired in the segregation process rather than in replication. The percentage of spores lacking DNA in analyzed spore chains was around 30 %, and occasionally up to 60 %. This is higher than for the mutants in other genes associated with chromosome segregation (Kim, Calcutt et al. 2000; Kois, Swiatek et al. 2009; Ditkowski, Troc et al. 2010). At present, it is not easy to conclude by what mechanism SsbB participates in chromosomal partitioning. Possibly SsbB interacts and supports the function of other DNA-binding proteins known to be important for chromosomal segregation. Truncated *ssbB* gene did not complement mutant strain phenotype (Figure 3.8). Since the ssDNA binding affinity of SsbBΔC (Figure 3.20 C) is similar to activity of whole length protein we speculate that C-terminus of SsbB is important for its activity during chromosome segregation. SsbB probably interacts and supports the function of other DNA-binding proteins known to be important for chromosomal segregation. For example, ParB binds to numerous *parS* sites near *oriC* while ParA provides support for proper distribution of these nucleoprotein complexes (Jakimowicz, Mouz et al. 2006). FtsK binds to the region where replication terminates and controls the removal of chromosome before septal closure (Wang, Yu et al. 2007), while SMC binds to DNA and promotes its condensation (Kois, Swiatek et al. 2009). Single, double and triple mutations of these genes (Dedrick, Wildschutte et al. 2009) did not completely block nucleoid segregation during sporulation; moreover the percentage of spores lacking DNA in analyzed spore chains was below 25 %. This indicates that SsbB makes an important contribution to chromosome segregation during sporulation.

Our analysis (Table 3.1) showed that the *ssbB* mutant had an increased number of spores in spore chains and slightly increased spore length. It was previously reported that strains with a DNA segregation defect caused by *smc*, *ftsK* or *parB* mutations produced spores that were similar in shape and size to the wild type strain (Dedrick, Wildschutte et al. 2009). Contrary to the *smc*, *ftsK* and *parB* mutants, the *ssbB* mutant also displayed whi phenotype and we hypothesize that our observation are due to more severe chromosome partitioning defect than those reported previously or that *ssbB* mutation causes a distinct defect at a final stage of spore maturation. In any case it will be a key

challenge to clarify the role of the SsbB protein in the complex mechanism of cell division in *Streptomyces*.

4.2. Transcriptional profiles of ssb genes correlate with their proposed biological functions

Differences in transcriptional profile of *ssb* genes were in agreement with their proposed biological functions. The *ssbA*, as a crucial component of DNA metabolism, is expressed throughout bacterial growth, being significantly up-regulated in rich medium during the logarithmic phase, while under starvation conditions its expression is prolonged (Figure 3.12). This result correspond to the cellular need for SsbA, which is higher during intensive DNA replication in fast-growing vegetative hyphae, as reported for SsbA from *B. subtilis* (Lindner, Nijland et al. 2004). Prolonged transcription of *ssbA* in minimal medium could be correlated to nutrient-limiting conditions that trigger aerial mycelium formation. This consequently leads to intensive replication in the apical region of hyphae (Ruban-Osmialowska, Jakimowicz et al. 2006). Co-regulation of transcription with gene(s) encoding ribosomal proteins was observed previously in *B. subtilis* (Lindner, Nijland et al. 2004). Here, we only showed co-transcription of *rpsF-ssbA* (Figure 3.12 A). Since the intergenic region between *ssbA* and the downstream ribosomal protein genes *rpsR* and *rplI* is 45 nt and 18 nt, respectively, co-transcription of the whole operon is predicted. Discrepancy of signals from PCR products obtained with different primers that amplify either *rpsF* or the *rpsF-ssbA* region (Figure 3.12 A2) could be ascribed to faster degradation of longer mRNA under starvation conditions. In general, we noticed that *ssbA* is expressed at a much higher level than *ssbB*. Even when PCR was extend to 35 cycles *ssbB* transcripts were barely detectable in rich medium. In minimal medium the gene is expressed more or less equally throughout 96 hours of growth, with slight increase towards later stages of growth (Figure 3.12 B). Since the minimal media mimics starvation conditions which trigger reproductive phase of growth, these results are consistent with proposed biological role of SsbB protein. In addition, transcriptional profiles obtain from the mycelia grown on sporulation (MS) plates had shown the same pattern of expression (not shown). In comparison, paralogous SsbB from *B. subtilis* is also transcribed in minimal medium but only during the stationary phase (Lindner, Nijland et al. 2004), whereas no expression is observed in rich media. While barely detectable in non-competent cells, SsbB appears to be ~20-fold more abundant than SsbA during competence in other competent bacteria, *Streptococcus pneumoniae* (Attaiech, Olivier et al. 2011). Other highly naturally competent organism, Neisseria *gonorrhoeae* also possesses SsbB, but this SSB, and as it seems, is not involved in DNA uptake (Jain, Zweig et al. 2012). Moreover, this SsbB is expressed both in pilliated (competent)

and non-pilliated (non-competent) cells, as a part of operon involving *TopB* gene, indicating involvement in other processes than competence.

Taking all into account, our results are consistent with the proposed biological role of SsbB during spore formation in the reproductive stage of *Streptomyces* growth.

Transcriptional starts for *ssbA* and *ssbB* (Figure 3.13) were identified as well. The results indicated that *ssbA* has two alternative transcriptional starts. This was in concert with the transcriptional starts identified independently by RNAseq (Dr. K. McDowall, personal communication). Although *ssbA* is transcribed throughout growth, the proposed proximal promoter region is more active (Dr. K. McDowall, personal communication) and does not display a typical *Streptomyces* vegetative σ factor consensus sequence. In contrary, the proposed distal promoter elements most closely resemble SEP like promoters (Strohl 1992; Bourn and Babb 1995). By inspecting promoter regions of *ssbA* we found only short motifs that partially resembled previously described regulatory elements (Ahel, Vujaklija et al. 2002). Altogether this is not surprising since streptomyces promoter regions are poorly understood and highly divergent due to their 6 numerous σ factors encoded by these organisms (Bentley, Chater et al. 2002). No transcriptional signal showing that transcription of *ssbA* occurs from a proximal promoter located in the intergenic (81 bp) or coding region of *rpsF* was detected. This is in concern with data obtained for *rpsF-ssbA-rpsR* operon from *B. subtilis* (Lindner, Nijland et al. 2004).The promoter region of *ssbB* is complex and unusual. It is extremely GC rich (78.9 %), much above the *Streptomyces* average (57- 62 %) reported previously (Strohl 1992). It also possesses various short repeats, which form two higher-order repeats and palindrome sequences that most likely overlap with a not obvious -35 region (Figure 3.13). Since *ssbB* is upregulated under starvation conditions we compared its promoter region with the promoters of *whi* genes known to be expressed during sporulation. However, alignment of these regions did not display common motif(s). Since the biological roles of *parAB* and *ssbB* are important for genome segregation we also compared their regulatory regions. Although the promoter region of parAB is not so rich in GC (63 %) we observed some common features. Upstream of the -35 region, the parAB promoter region possesses many short direct repeats and two copies of a DnaA box motif (TTC/GTCCACA) (Kim, Calcutt et al. 2000). This motif is also found in the *ssbB* promoter region but only in one complete copy (Figure 3.13). The gene-proximal promoter (P1), governing the expression of *parAB* throughout growth on minimal medium, displays a similar pattern of activity to *ssbB*, while the gene-distal promoter (P2) is upregulated during spore chain formation (Jakimowicz, Mouz et al. 2006). All this may reflect intensive DNA-regulatory protein(s) interactions and tight regulation of this region. Unexpectedly, the *ssbB* promoter was active in *E. coli* causing filamentous growth (Figure 3.16) as reported

previously (Meyer and Laine 1990). It is difficult to explain how an extremely GC-rich sequence and no obvious SEP like promoter could act as a regulatory element in a distantly related bacterium.

4.3. SsbA and SsbB proteins are found to localise invegetative and aerial hyphae

Localisation experiments confirmed the presence of labelled SsbA and SsbB within different stages of *Streptomyces* development. Although SsbA should be expressed at much higher levels than SsbB, recombinant protein under control of the same promoter, eGFPSsbA mostly does not show corresponding level of fluorescence. One possibility could be that its structure was disturbed by adding a large tag to relatively small protein. Such protein could be a subjected to rather rapid turnover. Strong support to this explanation is coming from the fact that repeated attempts to introduce "scar" mutation in strain carrying eGFPssbA constructs at *att* site failed (not shown). According to Western blot analyses (Figure 3.17) SsbA and eGFPSsbA are expressed at the approximately same level and most likely fluorescence of the foci is result of localisation of the hybrid tetramers. Diffused, but persistent localisation of SsbA could partly be a reflection of its versatile roles in all processes involving ssDNA. For example, archaeal RNA polymerase interacts with SSB (Richard, Bell et al. 2004). Accordingly, eubacterial SSB possibly bind to ssDNA during transcription since its interaction with RNA polymerase is shown. This possibly contributes to the diffused fluorescence. Most of SsbA bright foci are found within short branches and in the tips of the young aerial mycelia, probably reflecting ongoing processed of DNA replication. In vegetative hyphae, cell wall growth occurs mostly at hyphal tips (Gray, Gooday et al. 1990), suggesting that the tip region might also be the major site of DNA replication (Yang and Losick 2001). This is consistent with the observation that most SsbA bright foci appeared at the tips of vegetative mycelia. Contrary to this, no difference was observed in the distribution or number of foci between the tips and the rest of the vegetative hyphae (including branches) for for DnaN, a replicative DNA polymerase (Ruban-Osmialowska, Jakimowicz et al. 2006). The authors have shown that chromosome replication is taking place in many compartments of both vegetative and aerial hyphae, with the apical compartments of the aerial mycelium exhibiting the highest replication activity. Similary to SsbA, diffused or dispersed fluorescence was observed frequently. This was explanied as a result of the different stages of assembly and disassembly of the replication machinery (Ruban-Osmialowska, Jakimowicz et al. 2006), and could be additional reason for dispersed fluorescence of SsbA as well. Since SsbA has a much wider role in all types of DNA metabolism, not all of diffused localisation should be attributed to disassembly of replication machinery. SsbB protein is present in both types of mycelia as well. This is consistent with the finding that SsbB is continuously expressed during growth

though more strongly in minimal media. Interestingly, SsbB is mostly found in the gaps of DNA staining, indicating its cellular localisation between/or near the nuceloids. Similar localisation is observed for FtsK, a motor protein involved in DNA segregation that localise at the invaginating septa and clear each genome copy to the appropriate side of the septum (Dedrick, Wildschutte et al. 2009). Thus, this in consistency with proposed biological role of SsbB protein. Nevertheless, additional analyses should be performed to look more closely into the localisation of SsbB in the prespore compartments. It was shown by expression analyses that SsbA and SsbB, although differentially expressed, have overlap in expression timing. The colocalisation of these two proteins is rarely observed, possibly showing their differential activities in DNA metabolism. Supprinsingly, when mCherrySsbB was expressed from the plasmid promoter, SsBA was present as regularly positioned bright foci. We hypothesise that over-expression of SsbB somehow influences binding pattern of SsbA although at present it is difficult to elucidate by which mechanism.

In addition, the pattern of mCherry-SsbB foci in filamentous *E. coli* highly resembled the pattern of origins that separate in *E. coli* during cell division (Lau, Filipe et al. 2003). This indicates that during replication SsbB competes with homologous SSB_{Eco} by binding to ssDNA in the replication fork, thus causing the SOS-response in *E. coli.* Moreover, a delay in growth was observed in *E. coli* strain expressing either SsbA and SsbB, although this was much more pronounced when SsbB was overexpressed. This indicates differential activities of these two proteins even in heterologous system, although occupying almost same 3D structure (Paradzik, Ivic et al. 2013). *S. coelicolor* SsbA and SsbB are more similar to each other both on sequence and 3D structure level (not shown), than to *E. coli* SSB. However, C-terminal end of SsbA and SSB are similar (DEPPF vs DDDIPF), and therefore, SsbA could possibly be recoginzed from other *E. coli* proteins, and removed from DNA , at least to some extent, while this is not the case for SsbB.

4.4 SsbA and SsbB bind to ssDNA with different affinity

The EMSA experiment showed that SsbA and SsbB form complexes with circular ΦX1748 ssDNA. The results of this binding experiment indicated that SsbA binds to long ssDNA in a more cooperative way, as reported for SsbA from *S. pneumoniae* and SSB from *E. coli* (Lohman and Ferrari 1994; Grove, Willcox et al. 2005). In addition, cooperative binding of SSB_{EC} is favored at low salt concentration (Lohman and Ferrari 1994). The cooperative binding mode has been proposed to function in DNA replication (Shereda, Kozlov et al. 2008). In agreement with this, our results support the predicted biological role of SsbA. Cooperative binding of SsbB was less pronounced and almost completely abolished under moderate salt concentrations (Figure 3.21). We had shown (Paradzik, Ivic et al.

2013) that even SsbA has more pronounced cooperative binding, SsbB is displaying higher affinity for short oligonucleotides (dT_{35}). This is similar to ssDNA binding affinity of SsbB from *S. pneumoniae* (Grove, Willcox et al. 2005). Interestingly, the opposite result was published recently for SsbA and SsbB from *B. subtilis* (Yadav, Carrasco et al. 2012). This discrepancy was explained by differences in the C-terminal domain. Previous studies showed that deletion of several amino acids in the C-terminal tail enhances ssDNA binding of SSBs from *E. coli* (Kozlov, Eggington et al. 2010) and bacteriophage T7 (Marintcheva, Marintchev et al. 2008). According to the authors, deletion of the acidic C-terminus prevents its binding to the OB-fold, thus increasing binding affinity. In accordance with this, the higher binding activity of SsbB could be explained by the absence of the acidic C-terminus. This explanation does not apply to *S. pneumoniae* SsbB, since it possesses a very acidic tail. However, this could be explained by its extremely short C-terminal region (21aa) compared to the C-terminus of SsbB from *S. coelicolor* (46 aa). However, additional experiments are needed to confirm this hypothesis.

Truncated version of SsbB has shown similar binding activity as whole length SsbB. At high protein/DNA ratio the ssDNA binding activity of SsbΔC is reduced compared to SsbB. This should not have impact on binding to ssDNA in the cell, since we demonstrated that physiological levels of SsbB are low. Therefore, this result confirmed that SsbΔC could not complement *ssbB* mutant strain due to lack of C-terminal region. Alltogether, different binding properties of paralogous SSBs could also reflect their different biological functions.

The presence of disulphide bonds in SsbB protein was recently reported by our group (Paradzik, Ivic et al. 2013). As shown in Figure 3.22 disulphide bonds were detected in SsbB protein samples isolated from homologous and heterologous hosts. Although the possibility that S-S bond formation has occurred during protein purification could not be excluded, the results clearly indicated that binding affinity had decreased significantly in the presence of DTT (Figure 3.22 C). In a conclusion, the binding activity of the SsbB might be regulated during oxidative stress in *S. coelicolor* and the formation of disulphide bridges could increase its binding activity (Kang, Paget et al. 1999).

4.5. Variable copy number and genomic positions of ssb genes

As expected all bacteria analyzed so far have at least one *ssb* gene and this is in accordance with the essential biological role ascribed to SSB protein (Shereda, Kozlov et al. 2008). However, an unexpected and previously unreported finding was that many bacterial species possess more *ssb* copies, mostly of unknown function. The existence and biological role of paralogous SSB proteins attract great deal of attention in more recent period. Majority of the results published so far were

obtained analysing the SSBs from Firmicutes (Lindner, Nijland et al. 2004; Attaiech, Olivier et al. 2011). However, until now, the phylogenetic studies of duplicated SSB proteins were limited. Recent report characterized SsbB homolog within a conserved genetic cluster found in genetic islands of different Proteobacteria (Jain, Zweig et al. 2012). This SSB of unknown function was described as a member of distinct and phylogenetically distant group of SSB proteins.

In this thesis, phylogenetic analyses was performed on large dataset of available SSB sequences, collected from over 5000 bacterial genomes; among these were many paralogous SSBs. SSB proteins are well defined by the presence of OB fold, and relatively high homology. Therefore, these proteins are mostly well annotated in genomic databases, and as a consequence, search using key words in most cases retrieves correct hits, i.e. the same ones as those obtained with BLAST (similarity search).

Unexpectedly, 42 % of analysed species across different phyla showed to possess additional copy(s) of *ssb* genes on their chromosomes. In addition, members of most analysed phyla did not have uniform number of *ssb* genes (See Results, Table 3.3). This observation is not surprising, since it has been published recently that small divergences of 16S rRNA gene sequences can be accompanied by large differences in total genome repertoire (Kuo and Ochman 2009; Francino 2012) and that even populations of the identical 16S rRNA sequences can vary greatly at the genomic level (Kuo and Ochman 2009). Variability in the number of the *ssb* genes reported in this study could be further explained by the notion that gene amplification contributes to genome plasticity thus allowing better adaptation of prokaryotes to constantly changing environmental conditions (Romero and Palacios 1997). Therefore, duplicated gene often fulfil some new functional niche or help in gene dose effect. On the contrary these genes could easily be lost from the genome repertoire if change in life style cause paralogous gene function to become redundant (Francino 2012).

As shown in Section Results, most of the bacterial species display conserved syntheny of *ssbA* genes. In most cases this gene is placed between ribosomal protein genes (*rpsF-ssbA-rpsR*). It is a reasonable to conclude that this is an ancestral position of the *ssb* gene. Preservation of this position was observed previously, 40 % of analysed *ssb* genes (35/87) were positioned within ribosomal operon (Lindner, Nijland et al. 2004). In this thesis we examined members of 28 phyla and in 25 phyla *ssbA* position was conserved at high percentages (often above 90 %). Possibly this discrepancy with previously published results (Lindner, Nijland et al. 2004) could be explained by the bias that is introduced when taxonomic coverage of sequenced genomes is not representative. For example, Proteobacteria species represent most sequenced phyla due to their impact to human health. Their *ssbA* gene is allocated adjacent to *UvrA* gene and if these sequences are predominantly present in analysed dataset, the results would be significantly altered.

According to summarized data in the Table 3.3, only few phyla (Cyanobacteria, Bacteroides, Chlamydia and some Proteobacteria) have *ssbA* within other genomic context than ribosomal operon. However, these positions seem random and we argue that translocations of *ssbA* genes were probably separate evolutionary events since these groups are not closely related. Example of Proteobacteria speaks in favour of this; although all subdivisions share last common ancestor, only members of ε subdivision has *rpsF-ssbA-rpsR* cluster preserved. This could be explained by a translocation of *ssbA* in the ancestor of the subdivisions that diverged latter.

Interesting example of *ssb* translocation and subsequent subfunctionalisation is proposed for *priB* in β and γ subdivisions of Proteobacteria. As suggested (Ponomarev, Makarova et al. 2003), PriB protein probably evolved from SSB. In this case, paralogous *ssb* in the new position (near *UvrA* gene) retained original function, while *ssb* in original position (*rpsF-ssbA-rpsR*) adopted new, specialized function. Similar example was observed in this study; in the phylum Chlorobi, *ssbA* is duplicated near *DnaF* gene (encoding for delta subunit of DNA polymerase III), while *ssb* gene between *rpsF* and *rpsR* genes lost its acidic tail, and most likely adopted new function.

While the syntheny of *ssb*A genes is preserved in most bacterial groups (Figure 3.23), it seems that positioning of additional *ssb* copies is random. The newly duplicated genes are often clustered within a genome (Reams and Neidle 2004). Although present in some cases, this pattern is rarely observed among *ssb* genes. One striking example is *ssb* gene from *Deinococus radiodurrans*, which consists of two OB-folds/per dimer and probably evolved from the cluster of duplicated *ssb* genes (Bernstein, Eggington et al. 2004). As mentioned above, *ssb* paralogues are scattered around genomes, but are frequently found near genes involved in DNA metabolism. This random distribution of paralogous *ssb* genes speaks in favor of relaxed selection i.e. while essential *ssb* gene is performing vital function in all aspects of DNA metabolism, additional copies mutate faster and evolve specialized functions. This specialization could also be acheived through increase in the expression of one gene copy through regulatory changes (Francino 2012). The examples of such divergence are experimentally determined - differential expression is shown for SsbB from *S. coelicolor* (this study), or SsbB from *B. subtilis* (Lindner, Nijland et al. 2004). This time-separated expression could allow paralogues to adapt to different functional niches.

In conclusion, retention of *ssbB* genes and additional *ssb* copies, as well as their position within certain larger groups, indicates that its role is important and well established to species harbouring it (e.g. Actinobacteria, Cyanobacteria, some Proteobacteria, etc.). In addition, it cannot be neglected that large number of bacteria does not possesses duplicated *ssb* gene(s). As mentioned above, this fact could be explained by the loss of duplicated gene in certain ecological niches**.**

4.6. Paralogouss SSB proteins in Eubacteria have polyphyletic origins

The results clearly showed that two main groups of SSB sequences, SsbA and SsbB (Figure 3.25) are not clustered into two distinct subgroups by CLANS programme. Furthermore, phylogenetic analyses confirmed polyphyletic origin of various SsbB proteins. While in CLANS SsbAs are clustering much closer and according to affiliation to certain group of Bacteria (Figure 3.25), SsbB proteins are forming only few distinctive groups. However, it should be taken into account that all dispersed members could theoretically be ascribed to some specific group(s) that did not emerge due to the underrepresentation of the particular sequences. Disperse dots from CLANS clustering are also positioned outside the main groups of SSBs in the phylogenetic trees. Similar result was obtained even within the phylogenetic tree with broaden taxonomic coverage. In addition, 3D graph of CLANS analysis revealed that many SsbB sequences were placed together with SsbA dots. In this study it was confirmed that these sequences are usually highly similar to SsbAs and therefore are branching together within the trees (Figures 3.25, 3.26, 3.28...). These results could be explained as more recent duplication events.

SsbA sequences from Actinomycetes form a cluster separate (Figure 3.26) from the other phyla. This might reflect their adaptations to high GC content, as reported previously (Stefanic, Vujaklija et al. 2009) for 3D structures of SsbA from few Actinomycetes. In this study it was shown that two paralogous genes have adopted different biological functions through evolution that possibly correlate with the complex life-cycle of streptomycetes. To see if this was confined to the genus *Streptomyces* we determined the phylogeny of SSB proteins by analyzing representative SSBs from related orders of the phylum Actinobacteria. Phylogenetic analyses (Figure 3.26) revealed that SsbA and SsbB from orders Actinomycetales and Bifidobacteriales are clustered together, indicating that duplication occurred in a common ancestor. Retention of paralogues in most members of these orders indicates that these gene products have well established roles (Bratlie, Johansen et al. 2010). Although the SsbB proteins branched together they evolved much faster, as indicated by branch lengths. In contrast, among SsbA proteins, primary structure is conserved, probably reflecting their essential biological role. In the case of SSBs from *S. nausensis* and *C. glutamicum,* the third SSB (named SsbC) could be the result of new duplication (*C. Glutamicum)* or it could be a consequence of recent horizontal transfer (*S. nausensis)*. Some Actinobacteria SSBs are clustered separately, both in CLANS and phylogenetic tree (Figure 3.25, 3.26). Representatives of this group belong to deepest branches of this phylum (Zhi, Li et al. 2009) such as radiotolerant bacteria *Rubrobacter xyllanophilus.* Interestingly, clustering of orders Coriobacterales might indicate some recent duplication events

since their SSBs grouped together. Some genomes of these orders did not retain duplicated genes or duplication did not occur.

All analyzed Cyanobacteria species have at least two SSB proteins. Two distinguished groups of their SSBs are present both in CLANS clustering (Figure 3.25) and phylogenetic trees (Figure 3.27 and 3.31). This strongly suggests that the role of both SSBs might be very important. However, proteins annotated as SsbA are forming evolutionary distant cluster (Figure 3.25 and Figure 3.30) that display higher divergence. Nevertheless, two criteria justify obtained annotation: (i) these proteins possess highly conserved C-terminus and (ii) in the genome of *Gloeobacter vioalceus,* SsbA occupies ancestral position near ribosomal gene. Although acidic C-terminus is present to some extent in the group of the proteins annotated as SsbBs, their C-tail is much less conserved (when compared to SsbAs), often containing basic residues, such as asparagine. However, without experimental proof, it is not possible to claim with certainty which SSB has essential role in Cyanobacteria.

Members of the phylum Bacteroidetes has very divergent SSB sequences (see disperse dots in CLANS (Figure 3.25) and long branches in phylogenetic tree (Figure 3.28). Moreover, as shown, these sequences are forming third, well supported group of SSBs. Many species of this phylum have SSBs that belong to all three groups, and many possess only two SSBs. However, each species always has one member of SsbA indicating that this group is essential. In addition, *ssb* genes encoding group of SsbA proteins are predominantly located near *mutY* (65 %). Although phyla Bacteroidetes and Chlorobi are described in literature as closely related (Gupta and Lorenzini 2007), only SsbA from Chlorobi is branched with SsbA and SsbB from Bacteroidetes. As discussed in section 4.5, SsbA within Chlorobi is translocated near *dnaF* gene. The annotated SsbBs from Chlorobi formed evolutionary more distant group (Figure 3.28). Although all genes encoding these proteins are positioned within ribosomal genes (*rpsF-ssb-rpsR*) their C-terminus is not conserved.

Firmicutes represent phylum with very complex SSB relationships (Figure 3.29). The role of SsbB is very well described for two members of this group showing the importance of this protein in natural competence (Lindner, Nijland et al. 2004). Still, many other bacteria of this phylum lack SsbB, as most available genomes have one *ssb* gene. Some SsbBs from class Clostridia formed more distinct groups in the phylogenetic trees (Figure 3.31) and possesses unique domain (DUF) at its N-terminus. It is likely that this domain has some important although still unidentified role that contributes to specialised function of these SsbBs. SsbBs from the orders Bacillales and Lactobacillales although clustered separately, display less divergence. Above mentioned SsbBs from *B. subtilis* (Bacillales) and *S. pneumoniae* (Lactobacillales) (Lindner, Nijland et al. 2004; Attaiech, Olivier et al. 2011) that perform similar roles are clustered separately but this branching is not supportive enough (0.68).

Similar to other groups of Bacteria, SsbA sequences from Proteobacteria are grouped together in CLANS clustering. δ proteobacteria are particularly interesting group, as their SSB paralogues are not forming any distinct clusters although almost half of the representatives have more than one *ssb* genes. Furthermore, syntheny of their *ssb* genes is not observed. Other classes of Proteobacteria (α, β and γ) have versatile groups of SSB proteins. As previously in this study SsbA sequences are clustered together, both in CLANS clustering (Figure 3.25) and phylogenetic trees (short branches, Figure 3.31) and their genomic context is preserved (Table 3.3). Some representatives of α, γ and mostly β Proteobacteria have SsbB that form a distinguish cluster (Figure 3.31). All these SsbBs are positioned near *TopB* gene encoding DNA topoisomerase III (Topo III). Topo III is present only in three phyla: Proteobacteria, Firmicutes and Bacteroides. The scattered distribution of Topo III among bacterial phyla suggests that this enzyme was not present in the last common ancestor of all Bacteria (Forterre and Gadelle 2009). The bacterial Topo III does not seem to play a role in the control of DNA topology, but instead cooperates with the RecQ helicase to resolve Holliday junctions and converging replication forks, suppressing cross-over formation and stabilizing stalled replication forks (Suski and Marians 2008). Bacterial-like Topo III are encoded by several plasmids, suggesting that the bacterial Topo III enzymes may be ultimately of plasmid/viral origin (Forterre, Gribaldo et al. 2007). Extensive analysis in this work indicates possible cooperation of SsbB and DNA Topo III, since Topo III needs ssDNA for its activity. Recently SsbB of this group (from *N. gonorrheae*) was experimentally characterized (Jain, Zweig et al. 2012). It was shown that this protein is co-expressed with DNA *TopB.* However, the role of this SsbB still remains to be elucidated. In addition, there are *ParAB* genes always present in the same cluster. ParA and B proteins work in tandem during chromosome segregation (Kim, Calcutt et al. 2000). Possibly SsbB participate in this process as well.

Another cluster of Proteobacteria SsbBs belongs to α subdivision (Figure 3.31). This SsbB could be of plasmid origin due to the following: plasmid pLB1 was isolated from the soil (NCBI) carries gene that encodes SSB with over 90 % of similarity with SsbB from α subdivision. This gene on plasmid pLB1 and chromosomes is also surrounded by *ParA* and *ParB* like genes and *Xre* transcriptional regulator. Interestingly, many SsbBs in this cluster underwent multiple duplication events, although the reason for this can not be explained at the moment. Contrary to this, most of the SsbBs from γ subdivision branch with SsbA sequences (data not shown) what could indicate occurrence of recent duplications and/or frequent gene conversion events.

In almost all obtained trees SsbA sequences are clustering together having short branches and their grouping are in concert with their taxonomic position. This is not surprising since SsbA proteins are essential for bacterial survival (Shereda, Kozlov et al. 2008) and as such are subjected to high

evolutionary pressure against mutations. General notion is that cellular functions essential for life are usually encoded by ancestral genes. These genes, including *ssbA*, are associated with their host genomes throughout evolutionary history and are inherited through generations via vertical transfer (Wellner, Lurie et al. 2007). With respect to SsbB sequences, phylogenetic analysis indicated that in most phyla SsbB sequences are of polyphletic origin.

According to our analyses there are no traces to support early evolutionary duplication of *ssbA* gene (at the root of Eubacterial tree). Present results support idea of frequent duplication of *ssb* genes and does not exclude the possibility that this event occured in early evolution as well. If such event happened, ancestral *ssbB* gene was lost or subjected to concerted evolution and preserved high similarity to its paralogous gene. However, according to other studies it seems unlikely that two or more identical genes would be maintained in long-term evolution for the purpose of increasing the amount of gene product (Francino 2012). There is an example of frequent gene conversion between *tufA* and *tufB* genes (Elongation factor Tu) in γ proteobacteria (Lathe and Bork 2001). Compared to *ssb* genes, *tufA* and *tufB* genes, when both present, are always located at the same position on the chromosome. Moreover, this duplication is present in the same genomic context in *Deincoccus* and therefore this duplication represent ancient event which was preserved only in few groups of bacteria. More recent study (Isabel, Leblanc et al. 2008) on 72 Yersinia strains reveals that *tuf* genes are clustering either to *tufA* or *tufB* clade, and are probably not evolving through mechanism of gene conversion. Some level of gene conversion for *ssb* genes can be assumed, at least within pairs which seem to be recent duplicates, but there is strong evidence that this is not the trace of ancient duplication. First of all, genomic position of paralogous *ssb* genes is not conserved. Furthermore, even within groups that have divergent SSB paralogues (Actinobacteria, Cyanobacteria, Proteobacteria, Firmicutes, etc), there are lot of examples of additional SSBs (*Corynebacterium*, *Gloeobacter*, *Geobacillus, Lactococcus, Listeria*, *Erythrobacter...*), moreover, no phylum can be found where such duplicates are not present. Generally, the rates of duplication and deletion in bacteria are high, with only a small fraction of retained duplications. In any case, steady accumulation of paralogues seems to be of a similar magnitude in most of the genomes. It is believed that highly expressed genes evolve slowly and that their duplication is avoided or counter-selected (Drummond, Bloom et al. 2005). Furthermore, it is found that genes of high expression level, as measured by their codon bias, are strongly underrepresented among the most recent duplications (Hooper and Berg 2003). However, there are examples where highly expressed genes, e.g. trGTPases genes, *tuf* (EF-Tu) and *fus* (EFG), are represented by multiple copies (Lathe and Bork 2001; Lopez-Ramirez, Alcaraz et al. 2011; Margus, Remm et al. 2011). Duplications of *fus* gene are represented among all phyla, while *tuf*

genes duplications are restricted to few phylogenetic groups as discussed above. Furthermore, there are thirteen genomes in which recent duplications of second EGF gene were identified. Similary to *ssb*, it was found that bacterial genomes possess from at least one up to 12 copies of DEAD-box helicase genes (Lopez-Ramirez, Alcaraz et al. 2011). Expansions of *ssb* genes are detected in many groups. While two or maybe three copies of highly similar *ssb* genes could be explained by increased dosage, it is hard to apply this explanation to higher number of *ssb* copies in the cell. In addition, these SSBs are often very short (less than 100 amino acids) and could be pseudogenes. Such example are Phytoplasmas, which have very small genomes but consist of 26,9 % paralogues (Voon L. Chan 2006) and have 11 and 14 copies of *ssb* gene. This group survived great genome reductions and multiple horizontal gene transfer events (Sirand-Pugnet, Lartigue et al. 2007). Aster yellows witches'-broom phytoplasma possesses *ssb* gene between ribosomal protein genes, but its product is reduced to only 59 amino acids. Other copies are longer, but still short when compared to SSBs from other bacteria. Usually domain for ssDNA binding (OB fold) is conserved within first 100 amino acids long so it is possible that such short SSBs found in Phytoplasmas are not functional i.e. they are becoming pseudogenes. It was reported that shortening of C-terminal region influences DNA binding abilities (Kozlov, Cox et al. 2010), therefore it would be interesting to experimentally determine ssDNA binding abilities of short SSBs from Phytoplasmas. Taking into account all these, there is still important biological question: What are the roles of these multiple *ssb* copies in the cell?

In a conclusion, Eubacterial SSBs paralogous probably have multiple evolutionary origins. Gaining additional copies of *ssb* genes is somehow beneficial, although they can be lost later with no (significant) influence on survival (Lindner, Nijland et al. 2004; Paradzik, Ivic et al. 2013). Many species contain only one copy of *ssb,* and yet there are numerous examples of that phylogenetically closely related species have two or even more *ssb* copies in their genomes. These additional *ssb* genes are obtained through recent or ancient duplication events i.e. they are paralogues to original chromosomal *ssb.* In some cases, there is evidence that *ss*b genes might have been gained through horizontal transfer. Furthermore, the abundance of *ssb* genes is random across species and it is hard, if not impossible, to attribute certain life style/ecological niche to acquisition and preservation of additional *ssb* copies.

4.7. Analyses of recent duplicates revealed different pressures acting on SSB domains

As discussed, eubacterial *ssb* genes are frequently duplicated. As mentioned, criterion for such statement is that two paralogues branch together and that they are more similar to each other than to any other available sequence. The evolution of the new function within these duplicated gens may

start with the amplification of a gene having some level of preadaptation for that function (Francino 2012).

Recent *ssb* duplicates were analyzed using PAL2NAL tool. During this analysis many paralogue SSB pairs had to be excluded from d_S/d_N analysis as their d_S value were over saturation threshold (>3). This is the common problem of such type of analysis (Huerta-Cepas and Gabaldon 2011). When N-terminal region and C-terminal region are compared in terms of synonymous substitutions (d_S as a measure of relative age of the duplication event), there is almost always higher d_S value (often over level of saturation) for the C-terminal region. This is in discrepancy to common understanding of synonymous substitutions as relatively free of selective pressures. Probably, some forces, like codon usage, are acting at the DNA level and are responsible for these differences. Results of this analyses indicated that selection pressure is acting differently on different SSB domains. Nevertheless, when d_N/d_S ratio was compared for those pair which did not pass saturation threshold, there is slightly relaxed pressure on Ct region, although this is not statistically significant. PAL2NAL programme is not taking into account the existence of deletions in C-terminal regions. Alignment of SsbA proteins sequences with its recently duplicated paralogues (i.e. they are branched together) is clearly demonstrating that large parts of C-terminal region of protein are prone to deletions, with exception of acidic tails. Last five amino acids are well known to function in interacting with other proteins (Shereda, Kozlov et al. 2008) or with OB fold (Marintcheva, Marintchev et al. 2008). Although the acidic Ct tail is present in many cases in both paralogues, it is obviously not obligatory feature for both. This is well supported with the fact that much recently duplicated SSB protein often has addition of small stretch of amino acids to their C-terminus. In general, C-terminal region is usually rich in asparagine, glutamine, glicine or proline repeats and this could be a reason why mutation events in this region are facilitated and frequent. Taking all into account, Ct region is hot-spot for SSB divergence, which can be observed both in recent duplicates and more diverged paralogous SSB proteins, while OB fold (Nt) tends to preserve its sequence and consequently, function in binding single-stranded DNA.

5. Conclusions

1. SsbA is essential for cell survival while accurate chromosome segregation during sporulation depends on SsbB.

2. Two *ssb* genes from *S. coelicolor* are expressed differently; in minimal medium gene expression was prolonged for both genes and significantelly upregulated for *ssbB*. Gene expression profile correlate with their proposed biological functions.

3. Determination of promoter regions: *ssbA* gene is transcribed as a part of polycistronic mRNA from two alternative transcripional sites 163 bp and 75 bp upstream of *rpsF* translational start codon whereas *ssbB* is a trancribed as a monocinstronic mRNA from unusual promoter region 73 bp upstream of the AUG codon.

3. Fluorescently labelled SsbA and SsbB proteins are detected in vegetative and aerial mycelia. While SsbA is mostly found to cololocalise with DNA, SsbB is always present between nuceloids within *S. coelicolor* hyphae.

4. SsbA and SsbB proteins bind to single stranded DNA with different affinity. SsbA showes highly cooperative mode of binding whereas cooperative binding of SsbB is less pronounced and significantly decreased under moderate salt conditions.

5. Evolution of SSB paralogous proteins in Eubacteria is highly dinamic. Phylogenetic analyses indicate frequent duplication and loss of *ssb* genes and polyphyletic origins of SsbB proteins. Distinct groups of SsbB proteins observed in various eubacterial phyla could suggest their adaptation to a specific functions.

6. References

Abascal, F., R. Zardoya, et al. (2005). "ProtTest: selection of best-fit models of protein evolution." Bioinformatics **21**(9): 2104-2105.

Adams, D. W. and J. Errington (2009). "Bacterial cell division: assembly, maintenance and disassembly of the Z ring." Nat Rev Microbiol **7**(9): 642-653.

Ahel, I., D. Vujaklija, et al. (2002). "Transcriptional analysis of the recA gene in Streptomyces rimosus: identification of the new type of promoter." FEMS Microbiol Lett **209**(1): 133-137.

Ainsa, J. A., N. J. Ryding, et al. (2000). "WhiA, a protein of unknown function conserved among gram-positive bacteria, is essential for sporulation in Streptomyces coelicolor A3(2)." J Bacteriol **182**(19): 5470-5478.

Attaiech, L., A. Olivier, et al. (2011). "Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: maintenance of a reservoir for genetic plasticity." PLoS Genet **7**(6): e1002156.

Baker, T. A. and S. H. Wickner (1992). "Genetics and enzymology of DNA replication in Escherichia coli." Annu Rev Genet **26**: 447-477.

Bao, K. and S. N. Cohen (2003). "Recruitment of terminal protein to the ends of Streptomyces linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication." Genes Dev **17**(6): 774-785.

Bayer, I., A. Fliess, et al. (1989). "Modulation of the affinity of the single-stranded DNA-binding protein of Escherichia coli (E. coli SSB) to poly(dT) by site-directed mutagenesis." Eur J Biochem **179**(2): 399-404.

Benkovic, S. J., A. M. Valentine, et al. (2001). "Replisome-mediated DNA replication." Annu Rev Biochem **70**: 181-208.

Bentley, S. D., K. F. Chater, et al. (2002). "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)." Nature **417**(6885): 141-147.

Bernstein, D. A., J. M. Eggington, et al. (2004). "Crystal structure of the Deinococcus radiodurans single-stranded DNA-binding protein suggests a mechanism for coping with DNA damage." Proc Natl Acad Sci U S A **101**(23): 8575-8580.

Bielen, A., H. Cetkovic, et al. (2009). "The SGNH-hydrolase of Streptomyces coelicolor has (aryl)esterase and a true lipase activity." Biochimie **91**(3): 390-400.

Bielen, A. and D. Vujaklija (2007). Priprema i transformacija protoplasta. Metode u molekularnoj biologiji. A. A. ristov. Zagreb**:** 923-926.

Bierman, M., R. Logan, et al. (1992). "Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp." Gene **116**(1): 43-49.

Bourn, W. R. and B. Babb (1995). "Computer assisted identification and classification of streptomycete promoters." Nucleic Acids Res **23**(18): 3696-3703.

Bratlie, M. S., J. Johansen, et al. (2010). "Gene duplications in prokaryotes can be associated with environmental adaptation." BMC Genomics **11**: 588.

Bujalowski, W. and T. M. Lohman (1986). "Escherichia coli single-strand binding protein forms multiple, distinct complexes with single-stranded DNA." Biochemistry **25**(24): 7799-7802.

Casas-Finet, J. R., M. I. Khamis, et al. (1987). "Optically detected magnetic resonance of tryptophan residues in Escherichia coli ssb gene product and E. coli plasmid-encoded single-stranded DNA-binding proteins and their complexes with poly(deoxythymidylic) acid." J Biol Chem **262**(18): 8574-8583.

Chan, K. W., Y. J. Lee, et al. (2009). "Single-stranded DNA-binding protein complex from Helicobacter pylori suggests an ssDNA-binding surface." J Mol Biol **388**(3): 508-519.

Chase, J. W., J. J. L'Italien, et al. (1984). "Characterization of the Escherichia coli SSB-113 mutant single-stranded DNA-binding protein. Cloning of the gene, DNA and protein sequence analysis, high pressure liquid chromatography peptide mapping, and DNA-binding studies." J Biol Chem **259**(2): 805-814.

Chater, K. F. (2001). "Regulation of sporulation in Streptomyces coelicolor A3(2): a checkpoint multiplex?" Curr Opin Microbiol **4**(6): 667-673.

Cherepanov, P. P. and W. Wackernagel (1995). "Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant." Gene **158**(1): 9-14.

Clark, A. J. and S. J. Sandler (1994). "Homologous genetic recombination: the pieces begin to fall into place." Crit Rev Microbiol **20**(2): 125-142.

Costes, A., F. Lecointe, et al. (2010). "The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks." PLoS Genet **6**(12): e1001238.

Cox, M. M. (2007). "Regulation of bacterial RecA protein function." Crit Rev Biochem Mol Biol **42**(1): 41-63.

Curth, U., J. Genschel, et al. (1996). "In vitro and in vivo function of the C-terminus of Escherichia coli single-stranded DNA binding protein." Nucleic Acids Res **24**(14): 2706-2711.

Das, D., H. Hyun, et al. (2007). "Crystal structure of a novel single-stranded DNA binding protein from Mycoplasma pneumoniae." Proteins **67**(3): 776-782.

Davydova, E. K. and L. B. Rothman-Denes (2003). "Escherichia coli single-stranded DNA-binding protein mediates template recycling during transcription by bacteriophage N4 virion RNA polymerase." Proc Natl Acad Sci U S A **100**(16): 9250-9255.

Dedrick, R. M., H. Wildschutte, et al. (2009). "Genetic interactions of smc, ftsK, and parB genes in Streptomyces coelicolor and their developmental genome segregation phenotypes." J Bacteriol **191**(1): 320-332.

DeSanti, C. L. and W. R. Strohl (2003). "Characterization of the Streptomyces sp. strain C5 snp locus and development of snp-derived expression vectors." Appl Environ Microbiol **69**(3): 1647-1654.

Dharmaraj, S. and A. Sumantha (2009). "Bioactive potential of Streptomyces associated with marine sponges." World Journal of Microbiology & Biotechnology **25**(11): 1971-1979.

Din, S., S. J. Brill, et al. (1990). "Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells." Genes Dev **4**(6): 968-977.

Ditkowski, B., P. Troc, et al. (2010). "The actinobacterial signature protein ParJ (SCO1662) regulates ParA polymerization and affects chromosome segregation and cell division during Streptomyces sporulation." Mol Microbiol **78**(6): 1403-1415.

Donachie, W. D. (1993). "The cell cycle of Escherichia coli." Annu Rev Microbiol **47**: 199-230.

Drummond, D. A., J. D. Bloom, et al. (2005). "Why highly expressed proteins evolve slowly." Proc Natl Acad Sci U S A **102**(40): 14338-14343.

Erickson, H. P. (1995). "FtsZ, a prokaryotic homolog of tubulin?" Cell **80**(3): 367-370.

Erickson, H. P. and M. Osawa (2010). "Cell division without FtsZ--a variety of redundant mechanisms." Mol Microbiol **78**(2): 267-270.

Fedorov, R., G. Witte, et al. (2006). "3D structure of Thermus aquaticus single-stranded DNA-binding protein gives insight into the functioning of SSB proteins." Nucleic Acids Res **34**(22): 6708-6717.

Fernandez-Martinez, L. T., R. Del Sol, et al. (2011). "A transposon insertion single-gene knockout library and new ordered cosmid library for the model organism Streptomyces coelicolor A3(2)." Antonie Van Leeuwenhoek **99**(3): 515-522.

Ferrari, M. E., J. Fang, et al. (1997). "A mutation in E. coli SSB protein (W54S) alters intra-tetramer negative cooperativity and inter-tetramer positive cooperativity for single-stranded DNA binding." Biophys Chem **64**(1-3): 235-251.

Flardh, K. (2003). "Growth polarity and cell division in Streptomyces." Curr Opin Microbiol **6**(6): 564-571.

Flardh, K. and M. J. Buttner (2009). "Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium." Nat Rev Microbiol **7**(1): 36-49.

Flardh, K., K. C. Findlay, et al. (1999). "Association of early sporulation genes with suggested developmental decision points in Streptomyces coelicolor A3(2)." Microbiology **145 (Pt 9)**: 2229-2243.

Forterre, P. and D. Gadelle (2009). "Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms." Nucleic Acids Res **37**(3): 679-692.

Forterre, P., S. Gribaldo, et al. (2007). "Origin and evolution of DNA topoisomerases." Biochimie **89**(4): 427-446.

Francino, M. P. (2012). "The ecology of bacterial genes and the survival of the new." Int J Evol Biol **2012**: 394026.

Frickey, T. and A. Lupas (2004). "CLANS: a Java application for visualizing protein families based on pairwise similarity." Bioinformatics **20**(18): 3702-3704.

George, N. P., K. V. Ngo, et al. (2012). "Structure and cellular dynamics of Deinococcus radiodurans single-stranded DNA (ssDNA)-binding protein (SSB)-DNA complexes." J Biol Chem **287**(26): 22123-22132.

Glassberg, J., R. R. Meyer, et al. (1979). "Mutant single-strand binding protein of Escherichia coli: genetic and physiological characterization." J Bacteriol **140**(1): 14-19.

Goldman, N. and Z. Yang (1994). "A codon-based model of nucleotide substitution for protein-coding DNA sequences." Mol Biol Evol **11**(5): 725-736.

Gouy, M., S. Guindon, et al. (2010). "SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building." Mol Biol Evol **27**(2): 221-224.

Gray, D. I., G. W. Gooday, et al. (1990). "Apical hyphal extension in Streptomyces coelicolor A3(2)." J Gen Microbiol **136**(6): 1077-1084.

Gregory, M. A., R. Till, et al. (2003). "Integration site for Streptomyces phage phiBT1 and development of site-specific integrating vectors." J Bacteriol **185**(17): 5320-5323.

Grove, D. E., S. Willcox, et al. (2005). "Differential single-stranded DNA binding properties of the paralogous SsbA and SsbB proteins from Streptococcus pneumoniae." J Biol Chem **280**(12): 11067-11073.

Guindon, S., J. F. Dufayard, et al. (2010). "New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0." Syst Biol **59**(3): 307-321.

Gupta, R. S. and E. Lorenzini (2007). "Phylogeny and molecular signatures (conserved proteins and indels) that are specific for the Bacteroidetes and Chlorobi species." BMC Evol Biol **7**.

Gust, B., G. L. Challis, et al. (2003). "PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin." Proc Natl Acad Sci U S A **100**(4): 1541-1546.

Gust, B., G. Chandra, et al. (2004). "Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces." Adv Appl Microbiol **54**: 107-128.

Hamon, L., D. Pastre, et al. (2007). "High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein--DNA complexes." Nucleic Acids Res **35**(8): e58.

Han, E. S., D. L. Cooper, et al. (2006). "RecJ exonuclease: substrates, products and interaction with SSB." Nucleic Acids Res **34**(4): 1084-1091.

Hobbs, M. D., A. Sakai, et al. (2007). "SSB protein limits RecOR binding onto single-stranded DNA." J Biol Chem **282**(15): 11058-11067.

Hooper, S. D. and O. G. Berg (2003). "On the nature of gene innovation: duplication patterns in microbial genomes." Mol Biol Evol **20**(6): 945-954.

Hopwood, D. A. (1967). "Genetic analysis and genome structure in Streptomyces coelicolor." Bacteriol Rev **31**(4): 373-403.

Hopwood, D. A. (2006). "Soil to genomics: the Streptomyces chromosome." Annu Rev Genet **40**: 1-23.

Hopwood, D. A. (2007). Streptomyces in nature and medicine : the antibiotic makers. Oxford ; New York, Oxford University Press.

Huerta-Cepas, J. and T. Gabaldon (2011). "Assigning duplication events to relative temporal scales in genome-wide studies." Bioinformatics **27**(1): 38-45.

Innan, H. and F. Kondrashov (2010). "The evolution of gene duplications: classifying and distinguishing between models." Nat Rev Genet **11**(2): 97-108.

Isabel, S., E. Leblanc, et al. (2008). "Divergence among genes encoding the elongation factor Tu of Yersinia Species." J Bacteriol **190**(22): 7548-7558.

Jain, S., M. Zweig, et al. (2012). "Characterization of the single stranded DNA binding protein SsbB encoded in the Gonoccocal Genetic Island." PLoS One **7**(4): e35285.

Jakimowicz, D., J. Majkadagger, et al. (2000). "Architecture of the Streptomyces lividans DnaA protein-replication origin complexes." J Mol Biol **298**(3): 351-364.

Jakimowicz, D., S. Mouz, et al. (2006). "Developmental control of a parAB promoter leads to formation of sporulation-associated ParB complexes in Streptomyces coelicolor." J Bacteriol **188**(5): 1710-1720.

Jakimowicz, D. and G. P. van Wezel (2012). "Cell division and DNA segregation in Streptomyces: how to build a septum in the middle of nowhere?" Mol Microbiol **85**(3): 393-404.

Jakimowicz, D., P. Zydek, et al. (2007). "Alignment of multiple chromosomes along helical ParA scaffolding in sporulating Streptomyces hyphae." Mol Microbiol **65**(3): 625-641.

Jedrzejczak, R., M. Dauter, et al. (2006). "Structure of the single-stranded DNA-binding protein SSB from Thermus aquaticus." Acta Crystallogr D Biol Crystallogr **62**(Pt 11): 1407-1412.

Johnson, B. F. (1984). "Two-dimensional electrophoretic analysis of the regulation of SOS proteins in three ssb mutants." Arch Microbiol **138**(2): 106-112.

Kang, J. G., M. S. Paget, et al. (1999). "RsrA, an anti-sigma factor regulated by redox change." EMBO J **18**(15): 4292-4298.

Kantake, N., M. V. Madiraju, et al. (2002). "Escherichia coli RecO protein anneals ssDNA complexed with its cognate ssDNA-binding protein: A common step in genetic recombination." Proc Natl Acad Sci U S A **99**(24): 15327-15332.

Kelemen, G. H., P. H. Viollier, et al. (2001). "A connection between stress and development in the multicellular prokaryote Streptomyces coelicolor A3(2)." Mol Microbiol **40**(4): 804-814.

Kerr, I. D., R. I. Wadsworth, et al. (2003). "Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein." EMBO J **22**(11): 2561-2570.

Kieser, T. B., M.J.; Buttner, M.J.; Chater, K.F.; . Hopwood, D.A (2000). Practical Streptomyces Genetics. Norwich, The John Innes Foundation.

Kim, H. J., M. J. Calcutt, et al. (2000). "Partitioning of the linear chromosome during sporulation of Streptomyces coelicolor A3(2) involves an oriC-linked parAB locus." J Bacteriol **182**(5): 1313-1320.

Kois, A., M. Swiatek, et al. (2009). "SMC protein-dependent chromosome condensation during aerial hyphal development in Streptomyces." J Bacteriol **191**(1): 310-319.

Kondrashov, F. A., I. B. Rogozin, et al. (2002). "Selection in the evolution of gene duplications." Genome Biol **3**(2): RESEARCH0008.

Kowalczykowski, S. C., D. A. Dixon, et al. (1994). "Biochemistry of homologous recombination in Escherichia coli." Microbiol Rev **58**(3): 401-465.

Kozlov, A. G., M. M. Cox, et al. (2010). "Regulation of single-stranded DNA binding by the C termini of Escherichia coli single-stranded DNA-binding (SSB) protein." J Biol Chem **285**(22): 17246-17252.

Kozlov, A. G., J. M. Eggington, et al. (2010). "Binding of the dimeric Deinococcus radiodurans single-stranded DNA binding protein to single-stranded DNA." Biochemistry **49**(38): 8266-8275.

Kuo, C. H. and H. Ochman (2009). "Deletional bias across the three domains of life." Genome Biol Evol **1**: 145-152.

Kuzminov, A. (1999). "Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda." Microbiol Mol Biol Rev **63**(4): 751-813, table of contents.

Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and Clustal X version 2.0." Bioinformatics **23**(21): 2947-2948.

Lathe, W. C., 3rd and P. Bork (2001). "Evolution of tuf genes: ancient duplication, differential loss and gene conversion." FEBS Lett **502**(3): 113-116.

Lau, I. F., S. R. Filipe, et al. (2003). "Spatial and temporal organization of replicating Escherichia coli chromosomes." Mol Microbiol **49**(3): 731-743.

Lecointe, F., C. Serena, et al. (2007). "Anticipating chromosomal replication fork arrest: SSB targets repair DNA helicases to active forks." EMBO J **26**(19): 4239-4251.

Lindner, C., R. Nijland, et al. (2004). "Differential expression of two paralogous genes of Bacillus subtilis encoding single-stranded DNA binding protein." J Bacteriol **186**(4): 1097-1105.

Liu, J. H., T. W. Chang, et al. (2004). "Crystal structure of PriB, a primosomal DNA replication protein of Escherichia coli." J Biol Chem **279**(48): 50465-50471.

Lohman, T. M. and M. E. Ferrari (1994). "Escherichia coli single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities." Annu Rev Biochem **63**: 527-570.

Lopez-Ramirez, V., L. D. Alcaraz, et al. (2011). "Phylogenetic distribution and evolutionary history of bacterial DEAD-Box proteins." Journal of Molecular Evolution **72**(4): 413-431.

Lu, D. and J. L. Keck (2008). "Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I." Proc Natl Acad Sci U S A **105**(27): 9169-9174.

MacNeil, D. J., K. M. Gewain, et al. (1992). "Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector." Gene **111**(1): 61-68.

Margus, T., M. Remm, et al. (2011). "A computational study of elongation factor G (EFG) duplicated genes: diverged nature underlying the innovation on the same structural template." PLoS One **6**(8): e22789.

Marintcheva, B., A. Marintchev, et al. (2008). "Acidic C-terminal tail of the ssDNA-binding protein of bacteriophage T7 and ssDNA compete for the same binding surface." Proc Natl Acad Sci U S A **105**(6): 1855-1860.

McCormick, J. R. (2009). "Cell division is dispensable but not irrelevant in Streptomyces." Curr Opin Microbiol **12**(6): 689-698.

McCormick, J. R., E. P. Su, et al. (1994). "Growth and viability of Streptomyces coelicolor mutant for the cell division gene ftsZ." Mol Microbiol **14**(2): 243-254.

Meyer, R. R., J. Glassberg, et al. (1980). "A temperature-sensitive single-stranded DNA-binding protein from Escherichia coli." J Biol Chem **255**(7): 2897-2901.

Meyer, R. R. and P. S. Laine (1990). "The single-stranded DNA-binding protein of Escherichia coli." Microbiol Rev **54**(4): 342-380.

Mijakovic, I., D. Petranovic, et al. (2006). "Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine." Nucleic Acids Res **34**(5): 1588-1596.

Moreira, D. (2000). "Multiple independent horizontal transfers of informational genes from bacteria to plasmids and phages: implications for the origin of bacterial replication machinery." Mol Microbiol **35**(1): 1-5.

Murzin, A. G. (1993). "OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences." EMBO J **12**(3): 861-867.

Noens, E. E., V. Mersinias, et al. (2005). "SsgA-like proteins determine the fate of peptidoglycan during sporulation of Streptomyces coelicolor." Mol Microbiol **58**(4): 929-944.

Noens, E. E., V. Mersinias, et al. (2007). "Loss of the controlled localization of growth stage-specific cell-wall synthesis pleiotropically affects developmental gene expression in an ssgA mutant of Streptomyces coelicolor." Mol Microbiol **64**(5): 1244-1259.

Paradzik, T., N. Ivic, et al. (2013). "Structure-function relationships of two paralogous single-stranded DNA-binding proteins from Streptomyces coelicolor: implication of SsbB in chromosome segregation during sporulation." Nucleic Acids Res.

Pei, J. and N. V. Grishin (2007). "PROMALS: towards accurate multiple sequence alignments of distantly related proteins." Bioinformatics **23**(7): 802-808.

Petranovic, D., O. Michelsen, et al. (2007). "Bacillus subtilis strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication." Mol Microbiol **63**(6): 1797-1805.

Ponomarev, V. A., K. S. Makarova, et al. (2003). "Gene duplication with displacement and rearrangement: origin of the bacterial replication protein PriB from the single-stranded DNA-binding protein Ssb." J Mol Microbiol Biotechnol **5**(4): 225-229.

Posada, D. (2008). "jModelTest: phylogenetic model averaging." Mol Biol Evol **25**(7): 1253-1256.

Raghunathan, S., A. G. Kozlov, et al. (2000). "Structure of the DNA binding domain of E. coli SSB bound to ssDNA." Nat Struct Biol **7**(8): 648-652.

Raghunathan, S., C. S. Ricard, et al. (1997). "Crystal structure of the homo-tetrameric DNA binding domain of Escherichia coli single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution." Proc Natl Acad Sci U S A **94**(13): 6652-6657.

Reams, A. B. and E. L. Neidle (2004). "Selection for gene clustering by tandem duplication." Annu Rev Microbiol **58**: 119-142.

Reddy, M. S., N. Guhan, et al. (2001). "Characterization of single-stranded DNA-binding proteins from Mycobacteria. The carboxyl-terminal of domain of SSB is essential for stable association with its cognate RecA protein." J Biol Chem **276**(49): 45959-45968.

Richard, D. J., S. D. Bell, et al. (2004). "Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase." Nucleic Acids Res **32**(3): 1065-1074.

Romero, D. and R. Palacios (1997). "Gene amplification and genomic plasticity in prokaryotes." Annu Rev Genet **31**: 91-111.

Roy, R., A. G. Kozlov, et al. (2007). "Dynamic structural rearrangements between DNA binding modes of E. coli SSB protein." J Mol Biol **369**(5): 1244-1257.

Ruban-Osmialowska, B., D. Jakimowicz, et al. (2006). "Replisome localization in vegetative and aerial hyphae of Streptomyces coelicolor." J Bacteriol **188**(20): 7311-7316.

Ryding, N. J., M. J. Bibb, et al. (1999). "New sporulation loci in Streptomyces coelicolor A3(2)." J Bacteriol **181**(17): 5419-5425.

Saikrishnan, K., J. Jeyakanthan, et al. (2003). "Structure of Mycobacterium tuberculosis single-stranded DNA-binding protein. Variability in quaternary structure and its implications." J Mol Biol **331**(2): 385-393.

Saikrishnan, K., G. P. Manjunath, et al. (2005). "Structure of Mycobacterium smegmatis single-stranded DNA-binding protein and a comparative study involving homologus SSBs: biological implications of structural plasticity and variability in quaternary association." Acta Crystallogr D Biol Crystallogr **61**(Pt 8): 1140-1148.

Sambrook, J. F., E.F.; Maniatis, T. (1989). Molecular Cloning - A Laboratory Manual New York, Cold Spring Harbor Laboratory Press.

Savvides, S. N., S. Raghunathan, et al. (2004). "The C-terminal domain of full-length E. coli SSB is disordered even when bound to DNA." Protein Sci **13**(7): 1942-1947.

Schwedock, J., J. R. McCormick, et al. (1997). "Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of Streptomyces coelicolor." Mol Microbiol **25**(5): 847-858.

Shamoo, Y., A. M. Friedman, et al. (1995). "Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA." Nature **376**(6538): 362-366.

Shaner, N. C., R. E. Campbell, et al. (2004). "Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein." Nat Biotechnol **22**(12): 1567-1572.

Shereda, R. D., D. A. Bernstein, et al. (2007). "A central role for SSB in Escherichia coli RecQ DNA helicase function." J Biol Chem **282**(26): 19247-19258.

Shereda, R. D., A. G. Kozlov, et al. (2008). "SSB as an organizer/mobilizer of genome maintenance complexes." Crit Rev Biochem Mol Biol **43**(5): 289-318.

Sirand-Pugnet, P., C. Lartigue, et al. (2007). "Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome." PLoS Genet **3**(5): e75.

Smulczyk-Krawczyszyn, A., D. Jakimowicz, et al. (2006). "Cluster of DnaA boxes involved in regulation of Streptomyces chromosome replication: from in silico to in vivo studies." J Bacteriol **188**(17): 6184-6194.

Stefanic, Z., D. Vujaklija, et al. (2007). "Preliminary crystallographic study of Streptomyces coelicolor single-stranded DNA-binding protein." Croatica Chemica Acta **80**(1): 35-39.

Stefanic, Z., D. Vujaklija, et al. (2009). "Structure of the single-stranded DNA-binding protein from Streptomyces coelicolor." Acta Crystallogr D Biol Crystallogr **65**(Pt 9): 974-979.

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

Suski, C. and K. J. Marians (2008). "Resolution of converging replication forks by RecQ and topoisomerase III." Mol Cell **30**(6): 779-789.

Suyama, M., D. Torrents, et al. (2006). "PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments." Nucleic Acids Research **34**: W609-W612.

Szczepankowska, A. K., E. Prestel, et al. (2011). "Phylogenetic and complementation analysis of a single-stranded DNA binding protein family from lactococcal phages indicates a non-bacterial origin." PLoS One **6**(11): e26942.

Szklarczyk, D., A. Franceschini, et al. (2011). "The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored." Nucleic Acids Res **39**(Database issue): D561-568.

Theobald, D. L. and D. S. Wuttke (2005). "Divergent evolution within protein superfolds inferred from profile-based phylogenetics." J Mol Biol **354**(3): 722-737.

Tseng, Y. C., J. L. Hung, et al. (1994). "Involvement of RecF pathway recombination genes in postreplication repair in UV-irradiated Escherichia coli cells." Mutat Res **315**(1): 1-9.

Umezu, K., N. W. Chi, et al. (1993). "Biochemical interaction of the Escherichia coli RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein." Proc Natl Acad Sci U S A **90**(9): 3875-3879.

Umezu, K. and R. D. Kolodner (1994). "Protein interactions in genetic recombination in Escherichia coli. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein." J Biol Chem **269**(47): 30005-30013.

Vassin, V. M., R. W. Anantha, et al. (2009). "Human RPA phosphorylation by ATR stimulates DNA synthesis and prevents ssDNA accumulation during DNA-replication stress." J Cell Sci **122**(Pt 22): 4070-4080.

Ventura, M., C. Canchaya, et al. (2007). "Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum." Microbiol Mol Biol Rev **71**(3): 495-548.

Voon L. Chan, P. M. S., Billy Bourke (2006). Bacterial genomes and infectious diseases, Humana Press.

Wang, L., Y. Yu, et al. (2007). "Role of an FtsK-like protein in genetic stability in Streptomyces coelicolor A3(2)." J Bacteriol **189**(6): 2310-2318.

Wellner, A., M. N. Lurie, et al. (2007). "Complexity, connectivity, and duplicability as barriers to lateral gene transfer." Genome Biology **8**(8).

Willemse, J., J. W. Borst, et al. (2011). "Positive control of cell division: FtsZ is recruited by SsgB during sporulation of Streptomyces." Genes Dev **25**(1): 89-99.

Willey, J., J. Schwedock, et al. (1993). "Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by Streptomyces coelicolor." Genes Dev **7**(5): 895-903.

Wold, M. S. (1997). "Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism." Annu Rev Biochem **66**: 61-92.

Yadav, T., B. Carrasco, et al. (2012). "Genetic recombination in Bacillus subtilis: a division of labor between two single-strand DNA-binding proteins." Nucleic Acids Res **40**(12): 5546-5559.

Yang, C., U. Curth, et al. (1997). "Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 A resolution." Nat Struct Biol **4**(2): 153-157.

Yang, M. C. and R. Losick (2001). "Cytological evidence for association of the ends of the linear chromosome in Streptomyces coelicolor." J Bacteriol **183**(17): 5180-5186.

Yuzhakov, A., Z. Kelman, et al. (1999). "Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase." Cell **96**(1): 153-163.

Zhi, X. Y., W. J. Li, et al. (2009). "An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class Actinobacteria, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa." Int J Syst Evol Microbiol **59**(Pt 3): 589-608.

Other sources:

www.pdbj.org/eprots/index_en.cgi?PDB%3A3BEP

http://www.rcsb.org/pdb/explore/explore.do?structureId=1EYG

http://streptomyces.org.uk/redirect/RecombineeringFEMSMP-2006-5.pdf

7. Summary

SSB (**s**ingle **s**tranded DNA-**b**inding) proteins are essential in all domains of life. These proteins bind single stranded DNA during DNA replication, recombination and repair. Although genome sequencing revealed presence of the *ssb* parologues in many bacteria, the role of these duplicated genes is poorly studied. Linear choromsome of model bacterium *Streptomyces coelicolor* contains two *ssb* genes. The aims of this study were: (**i**) to investigate the cellular roles of paralogous SSB proteins from model bacterium *S. coelicolor*, starting with the hypothesis that only one of them is essential for bacterial growth; (**ii**) next aim was to characterize both SSB proteins using different molecular, biochemical and biophysics methods by determination of transcriptional profiles during growth, localisation in *S. coelicolor* mycelia and the analysis of their ssDNA binding affinities; and (**iii**) to perform extensive bioinformatics analyses on all available SSB protein sequences and to analyze their phylogenetic relationships.

Analysis of *S. coelicolor ssb* mutants has shown that *ssbA* is essential for bacterial growth, whereas *ssbB* displayed a whi phenotype, which is characteristic for streptomycetes developmental mutants. Moreover, analysis of mutant strain by fluorescent microscopy revealed that around 30 % of spores lacked DNA, while 23 % spores had excessive amount of DNA. Thus, *ssbB* probably plays a key role in chromosome segregation during sporulation. RT-PCR analyses has shown that two *ssb* genes from *S. coelicolor* are expressed differently, which correlate with their proposed biological functions. *ssbA* gene was transcribed as a part of policistronic mRNA from two alternative transcripional sites 163 bp and 75 bp upstream of *rpsF* translational start codon whereas *ssbB* was a trancribed as a monocinstronic mRNA from unusual promoter region 73 bp upstream of the AUG codon. Flourescently labelled SsbA and SsbB proteins were found to localise both in vegetative and aerial mycelia. While SsbA was predominantly cololocalizing with DNA, SsbB was always present between nuceloids within *S. coelicolor* hyphae. Distinctive DNA binding affinites monitored by electrophoretic mobility shift were observed. SsbA showed highly cooperative mode of binding whereas cooperative binding of SsbB was less pronounced and almost completely abolished under moderate salt conditions. Bioinformatic analyses has shown that evolution of SSB proteins in Eubacteria is highly dinamic. Phylogenetic anaylses of *ssb* genes revealed frequently duplication and loss, and indicated polyphyletic origins of SsbB protiens in various eubacterial phyla. Distinct groups of SsbB proteins observed in various eubacterial phyla could suggest their adaptation to a specific functions.

8. Sažetak

Proteini SSB (eng. **s**ingle **s**tranded DNA-**b**inding- proteini koji se vežu na jednolančanu DNA) su esencijalni u svim domenama života. Ovi proteini vežu jednolančanu DNA (jlDNA) u procesima DNA replikacije, rekombinacije i popravka. Proteini SSB vežu jlDNA preko domene koja veže DNA i naziva se OB domena i nalazi se na N-terminalnom kraju proteina. Nedavno je dokazano da uloga proteina SSB nije pasivna već izrazito kompleksna i dinamična jer tijekom interakcija s velikim brojem važnih staničnih proteina uključenih u DNA metabolizam modulira njihove aktivnosti. Pokazano je da protein SSB ostvaruje interakcije s drugim staničnim proteinima preko kiselo nabijenih aminokiselina koje se nalaze na samom C-kraju proteina. Cjelovite analize porodice SSB pokazale su da je C-terminalni kraj izuzetno sačuvan kod SSB proteina. Pretrage bakterijskih genoma ukazale su na veliki broj bakterijskih kromosoma s više gena *ssb.* No, stanična uloga paralognih gena i razlog njihovog zadržavanja u genomu nakon duplikacije ostaje nepoznanica.

Bakterija *Streptomyces coelicolor* ima dva gena *ssb* koji su smješteni u esencijalnoj regiji kromosoma. Streptomiceti su kompleksne bakterije koje tijekom rasta diferenciraju, zbog čega su važne i za temeljna istraživanja. Uz to, zbog sinteze antibiotika i drugih bioaktivnih spojeva (imunosupresora ili protutumorskih lijekova) ove bakterije su izuzetno važne i u biotehnologiji.

Ciljevi ovog istraživanja bili su: (**i**) istražiti biološke uloge paralognih proteina SSB iz bakterije *S. coelicolor* polazeći od hipoteze da je samo jedan od njih esencijalan za preživljenje bakterije; (**ii**) molekularno-biološkim, biokemijskim i biofizičkim metodama za oba proteina SSB ispitati: transkripcijske profile, sintezu i lokalizaciju tijekom rasta *S. coelicolor*, te im nakon pročišćavanja odrediti afinitete vezanja za jednolančanu DNA različitim tehnikama; (**iii**) načiniti opsežne bioinformatičke analize na svim dostupnim sekvencama bakterijskih proteina SSB te odrediti njihove filogenetske odnose, odnosno poziciju proteina SSB iz bakterije *S. coelicolor* u odnosu na ostale bakterijske proteine SSB.

Metodom "REDIRECT" (ciljanom PCR mutagenezom razvijenom za streptomicete) ustanovljeno je da je gen *ssbA* esencijalan za preživljenje bakterije *S. coelicolor*, dok je metodom transpozonske mutageneze gena *ssbB* određena biološka uloga proteina SsbB. Ovaj mutant je pokazao bijeli (*whi*) fenotip karakterističan za razvojne mutante streptomiceta, te je fluorescentnom mikroskopijom utvrđeno kako ima oko 30 % spora kojima nedostaje DNA i oko 23 % spora koje imaju DNA u suvišku. Stoga gen *ssbB* vjerojatno ima ulogu u segregaciji kromosoma prilikom sporulacije. Divlji tip gena komplementirao je *ssbB* mutant, za razliku od mutiranog gena (*ssbB*ΔC), čijem produktu nedostaje 30 aminokiselina na C-kraju. Ovaj rezultat sugerira da aminokiseline C-kraja proteina SsbB vjerojatno rade interakciju(e) s drugim proteinom(ima) čija je biološka uloga važna za točno odvajanje krromosoma tijekom sporulacije.

Metodom RT-PCR određeni su transkripcijski profili za oba gena *ssb* te je pokazano da oni imaju diferencijalnu ekspresiju. Gen *ssbA* prepisuje se u operonu s okolnim genima koji kodiraju za ribosomske proteine. Raznima ekspresije ovog operona najveća je u ranijima fazama rasta, u oba ispitana hranjiva medija (bogati i minimalni), što odgovara ulozi gena *ssbA* u procesima replikacije koji su intenzivniji u eksponencijalnoj fazi rasta. Isto tako, u minimalnom mediju ekspresija operona je produžena do kasnijih faza rasta u odnosu na bogati medij. Taj rezultat je očekivan jer minimalni medij oponaša uvjete gladovanja u kojem streptomiceti ulaze u reproduktivnu fazu rasta, gdje se također intenzivno odvijaju procesi replikacije DNA. Gen *ssbB* prepisuje se kao monocistronska mRNA; u minimalnom mediju transkript je prisutan u svim fazama rasta, s blagim povećanjem količine transkripta u stacionarnoj fazi. U bogatom mediju prisutnost transkripta je jedva vidljiva.

Određena su i mjesta početka transkripcije metodom 5' RACE (eng. **R**apid **a**mplification of **c**DNA **e**nds- brzo umnažanje krajeva cDNA). Utvrđeno je postojanje dva alternativna početka prepisivanja gena *ssbA*, koji se nalaze 163 pb i 75 pb uzvodno od početnog kodona gena *rpsF.* Transkripcija gena *ssbB* počinje 73 pb uzvodno od početnog kodona istog gena. Promotorska regija gena *ssbB* izuzetno je složena te obiluje ponavljajućim elementima. Sastav GC parova ove regije iznosi 78.9 %, što je daleko iznad prosjeka sastava GC parova streptomicetnih promotorskih regija.

Proteini SSB su obilježeni fluorescentnim oznakama kako bi se utvrdila njihova lokalizacija u miceliju streptomiceta. SsbA je obilježen zelenim fluorescentnim proteinom (eGFP), dok je SsbB obilježen crvenim fluorescentnim proteinom (mCherry). Fluorescencijskom mikroskopijom pokazano je da se oba proteina mogu pronaći i u vegetativnom i u zračnom miceliju. Proteini SsbA i SsbB uglavnom ne kolokaliziraju; dok SsbA lokalizira s DNA, SsbB se nalazi između nukleoida.

Metodom zaostajanja DNA u gelu (EMSA, eng. **e**lectrophoretic **m**obility **s**hift **a**ssay) određeni su afiniteti vezanja za oba proteina SSB. Ovaj eksperiment je potvrdio vezanje proteina SSB iz streptomiceta na dugu jlDNA. Protein SsbA veže jednolanačnu DNA na kooperativni način. Kooperatino vezanje je slabije kod proteina SsbB, štoviše pri povećanoj koncentraciji soli navedeni način vezanja je značajno inhibiran.

Za bioinformatičku analizu prikupljene su sekvence proteina SSB iz preko 5000 bakterijskih genoma. Analiza distribucije gena *ssb* po taksonomski različitim bakterijskim skupinama ukazala je na učestale duplikacije i delecije kod svih bakterija. Sve bakterije imaju najmanje jedan gen *ssb (ssbA)*, tako potvrđujući njihovu esencijalnu ulogu. Genomski položaj *ssbA* očuvan je u većini koljena bakterija. Gen *ssbB* nalazi se na različitim položajima; sintenija ovog gena najčešće nije konzervirana ni kod bliskih vrsta. Filogenetska analiza pokazala je da SsbA proteini evoluiraju sporo, iznova potvrđujući kako je njihova stanična uloga važna i očuvana. Filogenetska stabla ukazuju da dodatne

kopije gena *ssb* imaju polifiletsko podrijetlo. U nekoliko bakterijskih koljena očuvane su skupine SsbB proteina, time sugerirajući da su navedeni proteini evoluirali te da imaju važnu važnu biološku ulogu u tim skupinama bakterija. Nadalje, bliski paralozi pronađeni su u svim skupinama bakterija te su takvi parovi analizirani kako bi se ustanovili načini na koje evolucijskih pritisci djeluju na nedavno duplicirane gene *ssb*. Pokazano je kako su ovakvi genski parovi uvijek pod negativnom selekcijom, odnosno promjene nukleotida koji dovode do promjene u aminokiselinskom sastavu nisu evolucijski favorizirane. Stoga, nedavno duplicirani geni *ssb* vjerojatno imaju ulogu u povećanju koncentracije proteina u bakterijskoj stanici. Odvojena usporedba N-terminalnog kraja i C-terminalnog kraja visoko homolognih proteina SSB također je pokazala da su oba dijela proteina uvijek pod negativnom selekcijom. Međutim na razini DNA sekvence, C-terminalni kraj proteina puno brže evoluira.

Zaključci ovog istraživanja su slijedeći: (**i**) gen *ssbA* je esencijalan za stanično preživljenje, dok gen *ssbB* ima ulogu u kromosomskoj segregaciji tijekom sporulacije, (**ii**) geni *ssb* imaju diferencijalnu gensku ekspresiju što korelira sa predloženim biološkim ulogama, (**iii**) proteini SSB lokaliziraju u svim djelovima micelija streptomiceta, a pritom najčešće ne kolokaliziraju, (**iv**) proteini SSB imaju različite afinitete vezanja jednolančane DNA, te (**v**) analize sekvenci proteina SSB iz bakterija pokazale su dinamičnu evoluciju paraloga SSB, pri čemu učestalo dolazi do duplikacija i delecija gena *ssb*; iz čega proizlazi kako SsbB proteini imaju polifiletsko podrijetlo u različitim skupinama bakterija.

9. Abbreviations

Abbreviations

(M)bp	(**M**ega)**b**ase **p**air
aa	**amino acid**
AFM	**A**tomic **f**orce **m**icroscopy
APS	**A**mmonium **p**ersulfate
ATP	**A**denosine **t**ri**p**hosphate
CRM	**c**omplete **r**egeneration **m**edium
Ct	**C**-**t**erminus
DMSO	**D**imethyl **s**ulfoxide
dNTP	(**D**eoxy) **N**ucleoside **t**ri**p**hosphate
dsDNA	**d**ouble-**s**tranded DNA
DTT	**D**i**t**hio**t**hreitol
EDTA	**E**thylene**d**iamine**t**etraacetic **a**cid
Glu	**G**lutamine
HRP	**H**orse**r**adish **p**eroxidase
IPTG	**I**so**p**ropyl β-D-1-**t**hio**g**alactopyranoside
kDa	**k**ilo**d**altons
LB	**L**uria-**B**ertani broth
Lys	**Lys**ine
min	**min**ute(s)
MM	**m**inimal **m**edium
NCBI	**N**ational **C**enter for **B**iotechnology **I**nformation
Nt	**N**-**t**erminus
OB	**o**ligosaccharide/oligonucleotide-**b**inding fold
OD	**o**ptical **d**ensity
ORF	**o**pen **r**eading **f**rame
Ori	**ori**gin of replication
PCR	**P**olymerase **C**hain **R**eaction
Phe	**Phe**nylalanine
Pro	**Pro**line
RBS	**R**ibosome **b**inding **s**ite
RT	**r**oom **t**emperature
SB	**S**ample **B**uffer

SDS	**S**odium **d**odecyl **s**ulfate
SSB	**s**ingle **s**tranded DNA-**b**inding protein
ssDNA	**s**ingle-**s**tranded DNA
TAE	**T**ris **A**cetate **E**DTA
TBST	**T**ris-**b**uffered **s**aline with **T**ween
TEMED	N, N, N', N'-**te**tra**m**etil**e**tilen**d**iamin
Trp	**Tri**p**p**tophan
UV	**U**ltra**v**iolet

10. Supplement

The list of genes presented by white boxes in Figure 3.24

1 – hypothetical protein

2 – NAD-dependent glutamate dehydrogenase

3 – engD – GTP-dependent nucleic acid binding protein

4 – exoAA – apurinic/apyrimidinic exonuclease

5 – ywpG – gene of unknown function

6 – DHH family protein

7 – Lon-like ATP-dependent protease

8 – ferric uptake regulation

9 – phosphate ABC transporter ATPase

10 – peptidyl tRNA hydrolase

11 – ribosomal 5S rRNA E-loop binding protein

12 – malate synthase A

13 - selenocysteine synthase

14 - 4-diphosphocytidyl-2C-methyl-D-erythritol synthase

15 - PpiC - peptidyl-prolyl *cis-trans* isomerase

16 - AsnC/Lrp family regulatory protein

17 - histidine kinase

18 - rpsD - 30S ribosomal protein S4

19 - tprE

20 - deoC - 2-deoxyribose-5-phosphate aldolase

21 - deoD - purine-nucleoside phosphorylase

22 - peptidase M16 domain-containing protein

23 - Zn-dependent peptidase

24 - dipeptidase pevP argE

25 - thyA - thymidilate synthase

26 - precorrin-6X-reductase

27 - mreB - rod-shape determining protein

28 - ABC transporter like

29 - DNA binding protein HLI

30 - mutY - A/G specific adenine glycosylase

31 - CorC/HlyC family transporter associated protein

32 - ATP-dependent helicase

33 - leucyl aminopeptidase

34 - ftsQ - cell division protein

35 - phosphoesterase RecJ-like

36 - tRNA delta-2-isopentenylpyrophosphate transferase

37 - Zinc finger SWIM-domain containing protein

38 - UvrD/REP helicase

39 - recA

40 - mgtE2 - Mg^{2+} transporter

41 - pmbA

42 - parA

43 - parB

44 - yfeB

45 - yfeA

46 - yfd

47 - yfb

48 - yfa

49 - yeh

50 - alpA - phage transcriptional regulator

51 - cobyrinic acid a,c-diamide synthase

52 - UTP - glucose-1-phosphate uridylyltransferase

53 - DNA recombinase (plasmid related)

11. Curriculum vitae

Education

• J.J. Strossmayer *University* of Osijek, University of Dubrovnik and Ruđer Bošković Institute: PhD Study of *Molecular Biosciences* (2008. to date, pHD thesis submitted- title Molecular characterization, functional and phylogenetic analysis of paralogous SSB proteins from *Streptomyces coelicolor*)

• Faculty of science, University of Zagreb, undergraduate study-molecular biology (2001-2006)

• Gymnasium Juraj Baraković, Zadar (1997-2001)

Working experience

• Research visit University of Stratchclyde, Glasgow, UK (April 2011-July 2011)- FEMS scholarship

• PhD student at Ruđer Bošković Institute (since February 2007), in Laboratory of molecular genetics

• Work on diploma thesis in Tissue typing center, University hospital Zagreb, Kišpatićeva 12 (January 2006- June 2006)

Courses

• FEBS Advanced Lecture Course: «Analysis and Engineering of Biomolecular Systems», Spetses, Greece, September 2010

• 10th international Summer School on Biophysics: "Supramolecular structure and function", Rovinj, Croatia, September 2009

• EMBO Young Scientist forum, Zagreb, Croatia, May 2009

• John Innes/Rudjer Bošković Summer School in Applied Molecular Microbiology, : «Microbial Secondary Metabolites: Genomes, Signals and Communities», Dubrovnik, Croatia, August 2008.

Oral presentations

• Paradžik T., Ivić N., Filić Ž., Herron P., Luić M., Vujaklija D. Structure-function relationship of two paralogous SSB proteins from *Strepomyces coelicolor, Book of abstracsts of 3rd congress of Croatian geneticts with international participation,* Croatia, 2012

• Invited lecture from Croatian microbiological society, title: «New understandings about SSB protein roles in bacterial cell», Zagreb, Croatia 2010.

Publications

• Paradzik, T., Ivic, N., Filic, Z., Manjasetty, B.A., Herron, P., Luic, M. and Vujaklija, D. (2013) Structure-function relationships of two paralogous single-stranded DNA-binding proteins from *Streptomyces coelicolor*: implication of SsbB in chromosome segregation during sporulation. *Nucleic Acids Res*.

Posters

• Šimunov T., Razdorov G., Castaldo G., Vujaklija D. New interacting partners and possible roles of SSB protein from *Streptomyces coelicolor*, FEBS Advance lecture course, Greece, 2010.

• Šimunov T., Razdorov G., Castaldo G, Vujaklija D. Interacting partners of *Streptomyces coelicolor* SSB protein, *Book of abstracts of the Congress «Biology of Streptomycetes»,* Münster, Germany, October 2009.

• Šimunov T., Štefanić Z., Razdorov G., Castaldo G., Luić M., Vujaklija D. SSB protein from antibiotic-producing bacteria, *Book of Abstracts of CESAR 2009 Central European Symposium on Antimicrobial Resistance,* Croatia, 2009

• Šimunov T., Razdorov G., Castaldo G, Vujaklija D. SSB protein from *Streptomyces coelicolor* and its interacting proteins, *Book of Abstracts of 10th international Summer School on Biophysics* 2009.

• Šimunov T., Vlašić I., Ivančić-Baće I., Brčić-Kostić K., Vujaklija D. Tyrosine kinases in prokaryotic organisms and SSB proteins. *Book of Abstracts of the Congress "50 Years of Molecular Biology in Croatia"*. Zagreb 2008.

• Šimunov T., Vlašić I., Ivančić-Baće I., Brčić-Kostić K., Vujaklija D. Tyrosine kinases in prokaryotic organisms and SSB proteins. *Book of Abstracts of Fourth Croatian Congress of Microbiology with International Participation*. Zadar 2008

• Šimunov T., Kuzmanić A., Vujaklija D**.** Tyrosine phosphorylation in *streptomyces* and single stranded DNA binding protein *Book of Abstracts of Microbiological Congress „Power of microbes in industry and environment"*, Zadar, 2007.

Rewards

- FEMS schoolarship for young scientists, 2007
- FEBS Youth Travel Grant, 2010
- FEMS short-term scholarship for visiting scientist, 2011

Memberships

- Croatian society for biochemistry and molecular biology
- Croatian microbiological society

Other skills and competences

- Computer skills: Microsoft OfficeTM (WordTM, ExcelTM and PowerPointTM), Adobe Photoshop, phylogeny bioinformatic tools
- Languages: english (excellent), italian (good), spanish (good)