Universität Zürich
Institut für Medizinische Mikrobiologie
Direktor: Prof. Dr. med. E. C. Böttger

---

Arbeit unter Leitung von PD Dr. rer. nat. P. Sander

# Generation of unmarked *Mycobacterium bovis* BCG *recA* deletion mutants

**INAUGURAL-DISSERTATION**
zur Erlangung der Doktorwürde der Medizinischen Fakultät
der Universität Zürich

vorgelegt von
**Peter Michael Keller**
**von Mandach und Hottwil AG**

Genehmigt auf Antrag von Prof. Dr. med. E. C. Böttger
Zürich 2008

# Table of Contents

Page

# Abstract

**Objective** *Mycobacterium bovis* BCG is used as a live vaccine against tuberculosis. RecA-dependent homologous recombination is a major driving force of bacterial evolution (*1*). Therefore a RecA phenotype is recommended for all live vaccines (*2-4*). RecA$^-$ life vaccines demonstrate increased genetic stability and are favourable because of a reduced risk of reversion to virulence. Characterization of marked *recA* mutants provided proof of principle that persistence in the host, which is a major criterion of vaccine efficiency, is not affected by the deletion of *recA* (*5*). However, due to the presence of antibiotic resistance markers such a recombinant strain may not be used in human vaccination trials (*6*). The aim of this study was to generate unmarked *recA* deletion mutants in phylogenetically different *Mycobacterium bovis* BCG substrains.

**Methods** Six BCG vaccine substrains were selected based on the following criteria: The substrains had to feature a complete historical documentation with date of lyophilization for the primary seed lot, protective efficiency in case control studies, an official WHO, FDA or government issued permission for human use, and a molecular description (regions of difference, and single nucleotide polymorphisms in comparison to *M. bovis* AF2122/97). Identity of the substrains was confirmed by molecular techniques. Mixed cultures with more than one genotype were excluded. Unmarked inactivation of *recA* was achieved with a suicide (i.e. non-replicative) knock-out plasmid containing a partially deleted *recA* allele flanked by a hygromycin and a kanamycin selectable marker and a levansucrase (*sacB*) counterselectable marker. SacB induces sucrose sensitivity in mycobacteria and allows the isolation of allelic exchange mutants in a two-step selective process (*7*). *M. tuberculosis* H37Rv *recA* was cloned into the suicide vector together with approximately 1.2 kbs of flanking genomic DNA. The genomic sequences of the cloned region are identical in *M. bovis* AF2122/97 and *M. tuberculosis* H37Rv. An in-frame stop codon (following amino acid residue 169) was introduced in the cloned *recA* instead of an internal 1.2-kb *Pst*I fragment. BCG substrains were grown in Middlebrook media; electrocompetent cells were produced and transformed with the suicide knock-out plasmid. Following selection of single-crossover transformants at the *recA* locus on hygromycin B containing media, deletion mutants resulting from intramolecular recombination were isolated after a counterselective step on 10% sucrose containing media. The respective genotypes were determined by PCR and Southern blot analyses.

**Results** Unmarked *recA* deletion mutants were obtained in BCG substrains Frappier, Pasteur, and Denmark. In BCG Tice and BCG Phipps, single-crossover mutants for the *recA* locus could be isolated. After counterselection, only reversion to wild-type allele or persistence of the single-crossover genotype was noted. In BCG Russia only illegitimate recombinant mutants were obtained. Illegitimate integration excludes generation of *recA* deletion mutants as an intramolecular recombination does not occur. Sequencing of the r*ecA* locus revealed a mechanistic explanation for the inability to obtain homologous recombination in BCG Russia. This substrain has an insertional single nucleotide polymorphism (SNP) in *recA* causing translational frame shift and truncation of RecA. Thus, tuberculosis vaccine strain BCG Russia is a natural *recA* mutant.

**Conclusion** The presented strategy allows generation of unmarked allelic exchange mutants for the *recA* locus in *M. bovis* BCG and delivered new *recA*$^-$ vaccine candidate strains that base on three well-documented BCG vaccine substrains. The mutants are free of genomic selectable and counterselectable markers. The constructed suicide vector backbone for the generation of unmarked allelic exchange mutants may be used for other allelic exchange experiments in mycobacteria. BCG Russia *recA* SNP provides an explanation for the observed high degree of *in vitro* evolutional stasis.

**Outlook** The deletion of *recA* in *M. bovis* BCG leads to a stable genome in this live vaccine reducing the risk of further attenuation or reversion of virulence in the process of vaccine production and in vaccinees. The availability of *recA* mutants in *M. bovis* BCG substrains facilitates the construction of heterologous antigen expression vectors for the intracellular compartment with stable expression of recombinant antigens profiting from BCG's long generation time and persistence in the host.

# 1 Introduction

## 1.1 Global impact of tuberculosis

The global burden of tuberculosis infection is a matter of growing concern to the WHO and UNAIDS (*8*). The 2005 estimates according to the WHO fact sheet on tuberculosis (*9*) predict that one third of the total global population is chronically infected with *Mycobacterium tuberculosis* and nearly 9 million people were newly infected in 2005. Mortality due to this disease is devastating, with estimated 1.6 million deaths annually (including those of HIV-infected individuals).

In addition to rising epidemiological concerns on the spread of *M. tuberculosis* (*10*), there is a new threat to the global health: An increasing number of clinical strains showing resistance towards at least rifampicin and isoniacid have been isolated in the last few years. These strains fulfil the definition of multi-drug resistant tuberculosis and the treatment of these infections requires expensive second-line antibiotics and is often condemned to failure (*11-13*). Estimates are that more than 4% of all patients with tuberculosis worldwide have an infection with a multi-drug resistant strain (*14*). Even more concerning strains that are described as extensively drug resistant (XDR), are on the rise. XDR strains are defined as being resistant to fluoroquinolones and one or more of three injectable second-line drugs (capreomycin, kanamycin and amikacin) in addition to the previously mentioned first-line antibiotics (isoniazid and rifampicin). XDR tuberculosis cases have been found in at least 17 countries on all continents (*15*).

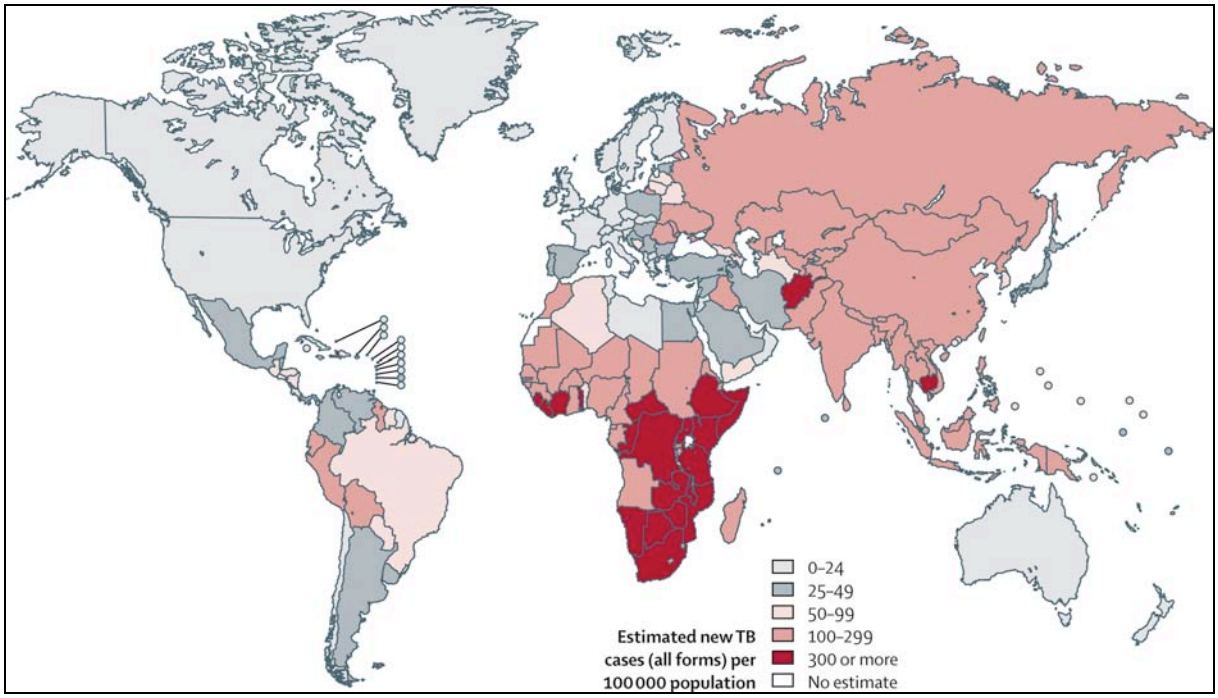



**Figure 1.** Worldwide incidence of tuberculosis in 2005. Source: WHO (*16*).

The role of vaccination in the prevention of tuberculosis is a field of ongoing debate and research: Since 1997, Ian Orme, David McMurray et al. have tested over 170 vaccine candidate strains in a NIH-funded program (*17, 18*). Even though four different concepts of vaccination strategies (subunit vaccines, DNA vaccines, live

attenuated mycobacteria and non mycobacterial-antigen delivery systems) were pursued, so far only the strategy of a standard intracutaneous BCG vaccination followed by protein antigen boost has proofed better protective efficacy in a guinea pig and mouse TB aerosol exposure model than the current Danish BCG vaccination strain that is recommended by the WHO (*10*).

**Table 1.** Novel vaccines against human tuberculosis, based on (*10*).

| Name | Type | Effectivness evidence | Developmental state | References |
|---|---|---|---|---|
| Modified vaccinia Ankara 85A | Live attenuated vector | Mice, guineapigs | Phase II | (*19-22*) |
| 72f fusion protein | Subunit | Mice, guineapigs, macaques | Phase I | (*23*) |
| rBCG30 | Recombinant BCG | Guineapigs | Phase I | (*24-26*) |
| ESAT6-85B fusion | Subunit | Guineapigs, macaques | Phase I | (*27*) |
| ESAT6-TB10.4 fusion | Subunit | Mice | Preclinical | (*28*) |
| Δ*ureC hly*$^+$ rBCG | Attenuated recombinant BCG | Mice | Preclinical | (*29*, *30*) |
| Δ*recA* rBCG | Recombinant BCG | Mice | Preclinical | (*5*, *31*) |

## 1.2 The BCG vaccine

The live attenuated strain of *Mycobacterium bovis*, which was developed by Albert Calmette and Camille Guérin between 1906 and 1919 at the Institut Pasteur at Lille and named thereafter, *M. bovis* BCG, is the world's most used vaccine. More than three billion doses have been administered in the vaccine's 85 year history (*32*). It is the only anti-tuberculosis vaccine that is in use on large scale and has a WHO approval. BCG was generated from a virulent isolate of *M. bovis* from a heifer with mastitis, which was collected by Nocard. The strain that was eventually to become BCG was sub-cultured every three weeks from 1908 to 1918, a total of 230 continuous passages on potato slices soaked in ox bile and glycerol. During in vitro cultivation genetic alterations accumulated, resulting in loss of virulence factors and therefore attenuation occurred. After this in-vitro attenuation, Calmette and Guérin demonstrated stable attenuation in several laboratory animals and cattle. The new vaccine was first used in a human newborn on July 18, 1921 by Weil-Halle, a colleague of Calmette and Guérin in France (*33*, *34*). The baby was at risk to develop tuberculosis and eventually die of tuberculosis because of contact to infected housemates. Starting in 1924 worldwide use and rapid global distribution of primary seed lots was noted. Scepticism concerning the safety and efficacy of BCG vaccine, and the *Lübeck incident* in which 72 of 240 children vaccinated with BCG died as a result of a vaccine batch contamination with virulent *M. tuberculosis* which was cultured in the same incubator as the vaccine (*35*), delayed the acceptance of BCG. A series of controlled trials begun in the 1930s (*36*). Despite inconsistent results from the trials, WHO encouraged widespread dissemination of BCG vaccines, starting in the 1950s (*37*). Since 1924, many different laboratories have maintained BCG, using different culture methods. As a result, the BCG substrains used today are not identical. The bacilli resulting from the original attenuation have never been cloned. The original strain of BCG has been lost and has been replaced by a variant while being serially transferred on artificial culture media at the Pasteur Institute (*38*). Both phenotypic and genotypic differences between substrains have been described (*39-41*). Since BCG substrains vary in protein expression, lipid composition, and behaviour in laboratory animal models and humans, an evidence of genetic instability has been supposed (*42*). Thus the term BCG does not describe a genetic entity but refers to a heterogeneous set of substrains. The BCG vaccine family has undergone major genetic rearrangements resulting in deletion and duplication of segments of its chromosome. More than

129 open reading frames clustered in 16 regions of difference (RD) are absent from the genome of BCG substrain Pasteur as compared to the *M. tuberculosis* reference strain H37Rv (*43*).

There is general agreement that the vaccine is effective against the disseminated forms, miliary tuberculosis and tuberculous meningitis in children (*32*) but the protective efficacy for BCG vaccines against adult pulmonary tuberculosis ranges between 0% in the Southern United States and Chingleput, Southern India versus 80% in the United Kingdom. Protection varies with geographic region, exposure to environmental mycobacteria, nutritional status, intensity of infecting dose, differences in the infectivity of *M. tuberculosis* strains, genetic differences in the population, methodological study design, and vaccine strain variations (*44-46*). The debate continues on how to interpret the results of the meta-analysis of the published field-trials and on the effectiveness of today's BCG vaccinations (*47*). Reliable correlates of protection should be established with respect to novel improved BCG-based vaccine candidates (Table 1).

## 1.3 RecA function and structure

### 1.3.1 Relevance of *recA* in genetic stability of vaccine strains

Genetic alterations may effect attenuation of life vaccines, making them either more harmful to the host and contacts or overattenuate vaccines and thereby lower the protective efficacy (*48*). The history of BCG shows only rare major adverse effects in the world's most used vaccine. Genetic alterations have never led to a reversion to the level of pathogenicity noted with wild-type *M. bovis*. Most of the observed adverse effects were due to immunological disabilities in the vaccinated or wrong application technique (intramuscular injection instead of intracutaneous injection). The estimated case fatality in the process of vaccination is very low (0.19 per million). Therefore BCG is among the safest vaccines. Unfortunately, a decrease of protective efficacy was noted over the years of application. This could either mean that the process of attenuation, which was induced by Calmette and Guérin through culture conditions to isolate the vaccine, was still ongoing or that the natural pathogen *M. tuberculosis* was evolving and changing its immunological behaviour in a way that escaped protection in vaccinated individuals. Because protection was decreasing with higher passage number of vaccine strains, an ongoing genetic evolution of the vaccine strains was supposed (*42*). Freeze-drying (lyophilization), which was introduced in the late 1950s, slowed down *in vitro* evolution as it eliminated the necessity of continuous subcultivation. In 1966, the WHO Expert Committee on Biological Standardisation adopted a series of recommendations for the production of BCG vaccine (*49*). These recommendations stated that the vaccine should be freeze-dried, and that the vaccine strain should be maintained by a seed-lot system whereby the number of passages of the vaccine lot should not exceed twelve passages. Most laboratories soon adopted this method of maintenance and it decreased the likelihood of selecting variants in later BCG vaccine lots. An expert committee recently reviewed the recommendations and now states not to exceed even lower passage counts from a primary lot (*50*). Ho et al. recommended PCR assays to monitor genetic alterations in vaccine production (*51*).

In *M. bovis* BCG – as in enteric bacteria – RecA is a key element of homologous recombination and a central regulator of the error prone DNA repair mechanism (SOS response). The synthesis of the RecA protein is inducible by DNA damage and RecA initiates the SOS response by forming filaments at the site of DNA damage. It facilitates autocatalytic cleavage of the SOS-repressor protein LexA (*52*). A deletion of *recA* leads to attenuation of some intracellular pathogens. E.g. in *Salmonella typhimurium*, deleterious mutations of *recA* have little effect on *in vitro* growth but significantly attenuate virulence in susceptible mice (*53*). In contrast BCG

*recA* mutants show a wild-type persistence in a mouse TB challenge model (*5*). In general, live vaccines should possess a *recA*$^-$ phenotype. *recA*$^-$ mutant strains are genetically more stable than their *recA*$^+$ counterparts, reducing concerns about major genetic changes and rearrangements resulting in an altered phenotype (*2*). Even though RecA acts as an inducible positive regulator of interspecies gene exchange in general (*54*), mycobacteria exhibit strong barriers to this mechanism. None the less a mutant *recA* allele averts reversion of virulence attenuation by interspecies gene transfer (e.g. with *M. tuberculosis*). Mycobacteria lack a classical mismatch repair system (*55*). Variation due to single nucleotide polymorphisms is exceptionally rare (*56*) and only minimally accounts for genome evolution. Genomic evolution happens either through insertion sequence (IS) triggered events, slipped-strand mispairing during replication, or recombination leaving a patched genome with insertions, deletions and gene duplications (*57, 58*). For instance, *M. bovis* BCG substrain Pasteur 1173-P2 has been analyzed by Brosch et al. (*59*). They found a 29-kb tandem duplication (DU1) in its genome, compromising the origin of chromosome replication *oriC* and another duplication of 36 kbs (DU2) corresponding to sequence positions between 3 590 900 and 3 690 124 in *M. tuberculosis* H37Rv (after occurrence of an internal deletion of 63 kbs, Δ*int*). These two major rearrangement in the genome of this reference vaccine strain can be assigned to recombination of short DNA homologies which might have occurred as a result of RecA activity as outlined by Edlund et al. for *E. coli* RecA activity (*60*). A *recA* knock-out mutation in BCG reduces further genomic evolution to rare point mutations and IS triggered events. Even though mutant *recA*$^-$ strains are more sensitive to DNA damaging agents such as UV-light and alkalylating substances, Δ*recA* BCG strains exhibit wild-type persistence and protection in a mouse vaccination model (*5, 31*).

### 1.3.2 Structure of mycobacterial *recA* genes

There is a major difference between the *recA* locus of *Mycobacterium tuberculosis* and its paradigmatical *E. coli* homologue. The *recA* gene in the pathogenic mycobacteria *M. tuberculosis*, *M. bovis* and *M. leprae* contains an in-frame insertion sequence which has no homology to the *recA* sequences found in other bacteria (*61*). The mentioned intervening sequence was termed intein. It is transcribed and translated as a part of RecA but spliced out posttranslationally (*62*). The intein is absent in *Mycobacterium smegmatis* (*63*). The intervening sequence harbours an open reading frame that encodes for a $Mn^{2+}$-ATP-dependent double strand specific endonuclease PI-MtuI (*64, 65*). In the spectrum of the genetically analysed human pathogenic bacteria, inteins are unique to mycobacteria.

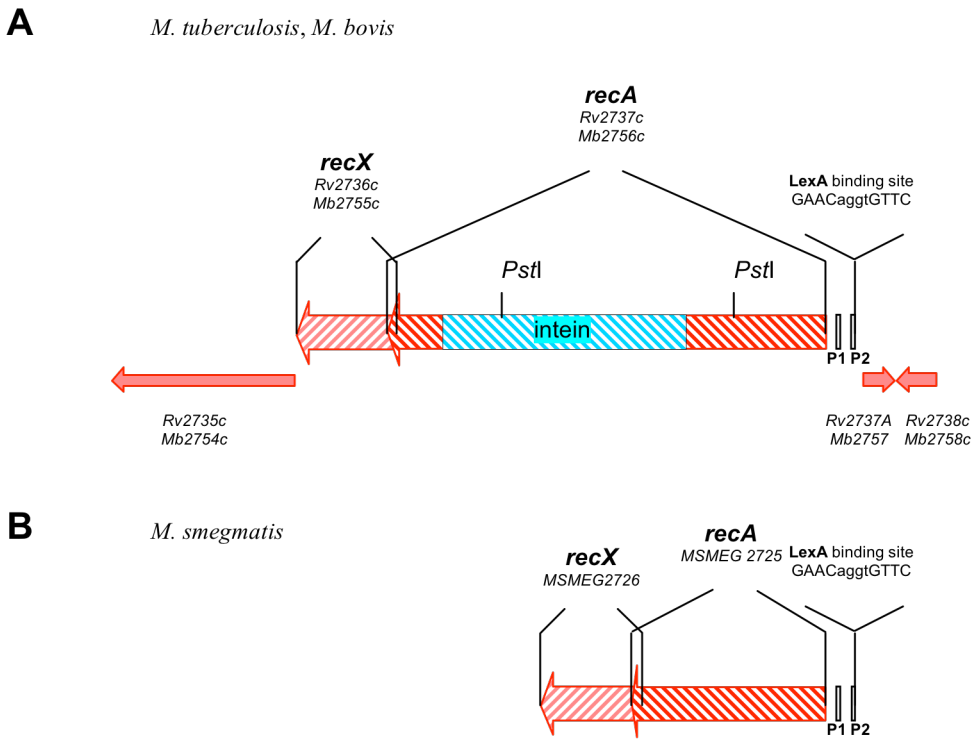



**Figure 2.** Organization of the *recA* locus in *M. tuberculosis*/*M. bovis* **(A)** and *M. smegmatis* **(B)**. In all three organisms, *recA* is cotranscribed with *recX*, which encodes for a putative regulatory protein. In contrast to *M. tuberculosis* and *M. bovis*, *recA* of *M. smegmatis* does not contain an intein (protein-splicing element). The two promoters upstream of *recA* are indicated with P1 and P2 (more distal from start codon of *recA* than P1).

In mycobacteria, the sequence encoding for *recA* lies on the complementary strand. In *M. tuberculosis* and *M. bovis*, the *recA* genes share 100% sequence identity. In *M. tuberculosis*, *M. bovis*, *M. leprae*, and *M. smegmatis recA* is cotranscribed with *recX*, a potential regulator of *recA*. The *recX* open reading frame overlaps *recA*, which makes the mycobacterial *recA* locus unusual. Although *recX* appears in many bacterial species, it rarely forms a *recA*-*recX* operon as in the mentioned mycobacteria. In *M. tuberculosis* and *M. bovis*, this operon is expressed from two promoters termed P1 and P2 upstream of *recA*, including a promoter with a LexA binding site. The two transcription start sites for *recA* have been identified by primer extension (*66*). One of the promoters, P1, is DNA damage inducible independently of LexA and RecA while the other, P2, is regulated by LexA in the classical way (*67*); i.e. the transcription of *recA* is repressed by the binding of LexA to an SOS box and is inducible by DNA damage as in other bacteria. The binding site for LexA is located between the putative -10 and -35 sequences for P2. The P1 promoter sequence are similar to the -10 and -35 sequences in *E. coli* $\sigma^{70}$ (*68*).

### 1.3.3 Structure and function of mycobacterial *RecA* protein

In *M. bovis* (as in *M. tuberculosis*) two important proteins of DNA repair and homologous recombination RecA and DnaB, are interrupted by inteins that must be excised by protein splicing before the translation products are functional (*69*). Overall the structure of spliced *M. tuberculosis* MtRecA is very similar to *E. coli* EcRecA, consistent with a high amino acid sequence similarity of 62% over the 352 amino acid polypeptide chain of the mature MtRecA (*70*). The active 38 kDa form of MtRecA with 352 amino acid residues develops after splicing of the intein region from an 85 kDa inactive precursor protein (*61*). The splicing site at amino acid

residues 253 and 254 (numbering refers to mature form) lies in the major central M domain of the active protein and is well conserved in comparison to other species with the exception of $Cys_{252}$ that forms an ester intermediate with $Cys_{253}$ by nucleophilic attack of the sulfhydryl group thereby starting the splicing process (*71*). The MtRecA protein features three domains; a 30 amino acid residue N-terminal domain consisting of a long α-helix followed by the M domain (amino acids 31-269) which is followed by a 59 amino acid residue C-terminal domain. The M domain, which is made up of twisted eight stranded β-sheets flanked by four α-helixes, harbours all features that are vital for the function of the RecA protein: a P-loop contains a triphosphate hydrolase fold, and two loops L1 and L2 are implicated in the binding of single-stranded DNA (ssDNA).

MtRecA display the hallmarks of EcRecA domains used to promote homologous recombination as outlined by the Holliday (*72, 73*) and subsequent models: binding to single-stranded DNA, ssDNA dependent ATP hydrolysis, formation of D-loops, homologous pairing of single-stranded DNA with duplex DNA, and strand exchange can be demonstrated for MtRecA (*74*). However, there are striking differences to EcRecA in the slower rate of ATP-hydrolysis by MtRecA and cofactor-dependence like the single strand binding protein (SSB) that is absolutely needed for the process in the case of MtRecA and other mycobacterial RecA proteins (*75*). In contrast to the RecA protein of *M. smegmatis* which has a maximal efficiency for strand exchange at neutral pH, MtRecA shows maximal efficiency of strand exchange in the alkaline pH range (*76*). Early gene disruption studies with short (*77*) and long (*78*) linear DNA fragments and suicide plasmids (*79*) suggested that the recombination machinery in *M. tuberculosis* and *M. bovis* in general is inefficient and prone to illegitimate recombination. However, more detailed investigations demonstrate that *M. tuberculosis* is proficient for homologous recombination (*63, 76*). In *M. tuberculosis*, an unusually long lag period has been observed in the appearance of RecA after DNA damage. Western blot analysis of the apparent RecA protein before and after induction of DNA damage has demonstrated only mature (i.e. spliced) RecA protein thereby suggesting that the splicing process is not rate limiting in expression of functional RecA (*80*). Splicing of MtRecA has been achieved in a heterologous expression system, and expression of the *M. tuberculosis recA* locus can complement *M. smegmatis recA* for homologous recombination and DNA repair (*63*).

### 1.3.4 Strategies for the generation of a gene knock-out in mycobacteria

The first report on allelic exchange in mycobacteria was a *pyrF*::*aph* mutant of *M. smegmatis* for which the target gene was inactivated by introduction of a suicide plasmid and subsequent selection for the mutant strain based on the kanamycin resistance of single- and double-crossover mutants (*81*). This strategy has proved useful for other loci in *M. smegmatis*, *M. tuberculosis*, and *M. bovis* (including BCG). Actually, a *recA* deletion mutant in BCG was generated using this strategy too (*5*). The introduction of an antibiotic resistance marker is unpreferable in life vaccines (*6*). Early attempts to achieve gene knock-outs in *M. tuberculosis* and *M. bovis* BCG were based on linear DNA fragments carrying the inactivated allele. The idea behind this delivery system was to increase recombinogenicity of the invading DNA and hence facilitate the recovery of allelic exchange mutants. The respective studies met large difficulties, which were attributed to peculiarities of slow growing mycobacteria, particularly their RecA. Many of the isolated mutants incorporated the invading fragments by illegitimate recombination. A *leuD* auxotrophic mutant could be isolated using very large stretches of homologous genomic DNA flanking the allele to be deleted. Because the *recA* loci of *M. tuberculosis* and *M. bovis* contain an intein there were speculations that these species were defective in homologous recombination (*77*). In the

following years, several workgroups demonstrated that homologous recombination is readily achievable in *M. tuberculosis*. The irrelevance of the intein with respect to *RecA* function was finally revealed and confirmed by Frischkorn et al. in 1998 (*82*) and Papavinasasundaram et al. 1998 (*63*). The next major advance came with easy usable selectable and counterselectable markers. Single gene knock-out by homologous recombination involves crossover events on either side of the desired mutation; mutant identification is simplified by availability of methods to distinguish allelic exchange (double-crossover) from partial merodiploid products of a single-crossover event and from the background resistant mutants. Two step strategies using positive i.e. kanamycin resistance (*aph*) and negative selection i.e. streptomycin sensitivity (*rpsL*) enabeled knock-out generation in slow growing mycobacteria to so far unachieved values. Another two step strategy was described by Pelicic et al. (*83*): A suicide vector construct with an inactivated target allele, flanked by kanamycin and/or hygromycin resistance cassettes as selectable markers, and a sucrose sensitivity conferring cassette (*sacB* levansucrase encoding gene from *Bacillus subtilis*) as counterselectable marker. This strategy results in truly isogenic strains that only differ in the target gene and it does not leave any marker-scar in the mycobacterial genome after a knock-out generation procedure. The incorporation of two positive selectable markers allows choosing the marker in the process of generating the respective mutant according to the performance of the selection. The *B. subtilis* levansucrase was already examined in 1966 by Dedonder (*84*). In *B. subtilis*, levansucrase (sucrose: 2,6-β-D-fructan 6-β-D-fructosyltransferase; EC 2.4.1.10) is a secreted enzyme that catalyzes hydrolysis of sucrose and synthesis of levans (high-molecular-weight fructose polymers). The *sacB* gene was first used successfully as a counterselective marker for the generation of unmarked mutations in *Escherichia coli*, *Erwinia chrysanthemi*, and *Legionella pneumophila*, demonstrating that expression of *sacB* in the presence of sucrose is lethal in gram-negative bacteria (*85-87*). Later on, expression of *sacB* in the gram-positive *Corynebacterium glutamicum* showed a cytotoxic effect in presence of sucrose (*88*). Investigation of *sacB* expression in mycobacteria revealed that it has a toxic effect on cells of *M. smegmatis*, *M. bovis* BCG and *M. tuberculosis* at 2% (in the fast growing species) respectively at 10% (in the slow-growing) sucrose concentration in the culture medium making it a useful tool for counterselective procedures in mycobacteria too (*7, 89*). There are several hypothesises explaining the molecular basis of levansucrase toxicity in hosts other than *B. subtilis*. In gram-negative bacteria the toxicity might be due to accumulation of levans, which bulk the periplasm due to their high molecular weight or due to transfer of fructose residues to inappropriate acceptor molecules thereby compromising the function of certain metabolic pathways. Because the gram-positive *C. glutamicum* owns a membrane-like outer structure mostly consisting of mycolic acids, lethality was suggested to correlate with the gram-negative model. The lethal effect in mycobacteria may be explained in the same way.

Other strategies for the generation of unmarked mutations in mycobacteria like specialized transduction, which allows the generation of unmarked mutants as well, are valuable alternatives for the strategy exploited in this study.

## 1.4 Aims of the study

1) Selection and characterisation of representative *M. bovis* BCG vaccine substrains that have (a) a continuing historical documentation with respect to passage number, morphological features, date of lyophilization, and protective efficacy in case control studies. The BCG substrains to be selected (b) should have a WHO, FDA or government issued approval for human use.
2) Cloning of a *recA* targeting vector: The primary goal of the study is to generate unmarked *recA* gene knock-out mutants; i.e. inactivating *recA* without leaving a residual antibiotic resistance or auxotrophic selectable or counterselectable marker after the genetic manipulation in well characterized vaccine substrains.
3) Transformation of BCG substrains with targeting vector: Culture conditions should represent little selective pressure towards certain phenotypes and the passage count to generate the knock-out mutants should be as low as possibly attainable to prevent further genetic rearrangements.
4) Identification of *recA* single-crossover intermediates.
5) Counterselection of *recA* single-crossover transformants.
6) Identification and characterization of unmarked *recA* mutants.

# 2 Material and Methods

## 2.1 Bacterial strains

### 2.1.1 *Mycobacterium bovis* BCG vaccine strains

The following strains were selected for the knock-out generation:

*M. bovis* BCG, substrain **Pasteur** (1173P2-B), dated 1961, CIP # 105050, TMC 1011, further descriptor 1117 on vial), obtained through Institut Pasteur, Paris, France. This strain is the current vaccine strain as referred by the Institut Pasteur and has been sequenced recently (*41*). It has the broadest documentation (passage number, lyophilization date) and was widely used in large field trials (*46*). It was derived in 1961 from single colony out of 30 colonies examined that yielded cultures, which corresponded – mostly in terms of pigmentations – to the original description of culture morphology by Calmette. This specific CFU was selected for freeze-drying, termed 1173P and the second batch is now the primary seed lot of BCG 'Paris' or 'French'. Recent studies have shown that this particular substrain underwent two major rearrangements with tandem duplications in the genome with respect to the *M. bovis* AF2122/97 sequence (*41, 59*).

*M. bovis* BCG, substrain **Tice** (Chicago), OncoTICE®, dated 1973, lot 177415, exp. date. 2005-05-24, obtained through Organon Ltd, Seattle, USA. The Tice substrain was developed at the University Illinois, Chicago, IL, USA (S. Rosenthal, R. Dubos, and A. Jespersen): In 1933 Dr. Rosenthal was sent to Paris by Dr. Tice to obtain a seed-lot of BCG and learn vaccine production techniques. The substrain he received later became known as BCG Tice-Rosenthal and is called BCG Tice today. Rosenthal cultured at least six substrains (H, K, E, L, LH, and BL) derived from the initial strain during cultivation in the U.S.; all of these were used for vaccine production. In October 1952, Rosenthal mixed three culture flasks of the BL variant with one flask of the new standard strain obtained from Institut Pasteur in 1951. This mixed culture was called BLP. It was lyophilized in 1952 and since then only batches originating from this lot have been used for vaccine production (*90*). Behr et al. described only one pattern of genomic deletions in the Tice substrain (*91*). Horwitz et al. used BCG Tice for the generation of a recombinant *M. bovis* BCG vaccine which expresses a *M. tuberculosis* 30-kDa major secretory protein (*25*). The unmodified strain has shown a slight tendency towards higher protective efficacy and tuberculin conversion rate in recent trials (*48*). BCG Tice has an FDA approval (*92*) as a antineoplastic agent in the treatment of primary or relapsing flat urothelial cell carcinoma in situ (Tis) of the urinary bladder, and as an adjuvant therapy after transurethral resection (TUR) of a primary bladder tumor or relapsing superficial urothelial cell carcinoma of the bladder stage Ta (the cancer is just in the innermost layer of the bladder lining; grade 1, 2 or 3) or T1 (the cancer has started to grow into the connective tissue beneath the bladder lining; grade 1, 2, or 3).

*M. bovis* BCG, substrain **Frappier** (Montreal), primary lot, dated 1973, lot number 1376141, obtained through the American Type Culture Collection (ATCC # 35735), Rockville, MD, USA. The Frappier strain held by ATCC was initially transferred from the Institut Pasteur to Frappier's Institute in Montreal in 1937, was integrated in the Trudeau Mycobacterial Culture Collection and finally went to the ATCC. This strain showed per-

sistently high protective efficacy even after high passage numbers (*48*). It was used in the large Saskatchewan field trial in 1933 (*93*), where it showed a protective efficacy of 82% in a follow-up period of 15 years. These data were extensively peer-reviewed (*44, 94*) and another follow-up study (*95*) confirmed this efficacy data.

*M. bovis* BCG, substrain **Russia** (Moscow), ATCC # 35740, dated 1924, vial from Medgamal (Медгамал), descriptor "Vakzina tuberkueznja (BCG) suchaja", in Cyrillic letters: Вакцина Туберкуленая (БЦЖ) сухая, lot number S. 547-1104 K. 1491, exp. date 12.2006, obtained through Torgovyi Dom Allergen (Allergen Trading House), Moscow, Russia. This was the first documented daughter strain distributed by Institut Pasteur in 1924 (*96*) going directly to Russia. It is referred as an 'early strain' (*43*) in regard to its genetic characteristics (RDs, insertional elements, antigen expression pattern).

*M. bovis* BCG, substrain **Denmark** (1331, Glaxo®, Mérieux®), obtained through the Statens Serum Institute, Copenhagen, Denmark. This substrain is nowadays accredited in the states of the European Union as official BCG vaccine. Several commercial producers offer this substrain in vaccine vials (Aventis, BioMérieux). Vaccine containing this substrain is WHO pre-qualified for sale to UN agencies including UNICEF (*97*). According to Oettinger et al. (*90*) Statens Serum Institut was provided with a BCG lot by Calmette in 1927. This first Danish substrain was believed to be to reactogenic and in November 1931, another BCG strain (transfer 423) was obtained from Institut Pasteur. In 1960 batch 1331 was freeze-dried and became the primary seed-lot of BCG Denmark (*90*).

*M. bovis* BCG, substrain **Phipps** (Philadelphia), ATCC # 35744, dated 1928, vial dated Jan. 9.87, obtained through the American Type Culture Collection, Rockville, MD, USA where it arrived from Institut Pasteur, via J. Aronson, A. Dannenberg, Jr. and the Trudeau Mycobacterial Culture Collection. It has shown only a slight decrease in protective efficacy in trials with increasing passage numbers (*48*).

Even though the *M. bovis* BCG, substrain **Japan** is prequalified for sale to UN agencies and in large use, we didn't implement the Japan substrain, which is also referred as Tokyo 172, because molecular comparisons of the current Japan vaccine substrain and asservational samples from ATCC have shown that this substrain owns more than one molecular type (IS6110 -2/*mpt64*$^{+}$ and IS6110 -1/*mpt64*$^{-}$) (*43*). The current vaccine strain is described as mixed culture of two subpopulations differing in RD16 regions (*98*). As we intended to generate a clone, the presence of more than one genotype is undesirable.

All vaccine strains were delivered in freeze-dried state and revitalized according to the prescriptions of the delivering manufacturer (see 2.2.3.2).

### 2.1.2 *Escherichia coli* strains

***E. coli* XL1-Blue MRF´**, Genotype: Δ(*mcrA*)183, Δ(*mcrCB-hsdSMR-mrr*)173, *endA*1, *supE*44, *thi*-1, *recA*1, *gyrA*96, *relA*1, lac [F´ proAB lacIqZΔM15Tn10 (Tetr)]. XL1-Blue MRF´ cells are tetracycline resistant. The XL1-Blue MRF´ (Minus Restriction) strain is a restriction minus (*McrA*$^{-}$, *McrCB*$^{-}$, *McrF*$^{-}$, *Mrr*$^{-}$, *HsdR*$^{-}$) derivative of Stratagene's XL1-Blue strain. XL1-Blue MRF´ cells are deficient in all known restriction systems

[Δ*(mcrA)183*, Δ*(mcrCB-hsdSMR-mrr)173*], and are endonuclease (*endA*), and recombination (*recA*) deficient. The *hsdR* mutation prevents the cleavage of cloned DNA by the *Eco*K endonuclease system, and the *recA* mutation helps ensure insert stability. The *endA1* mutation improves the quality of miniprep DNA. XL1-Blue MRF´ cells contain the *lacI*q*ZΔM15* gene on the F´ episome, allowing blue/white screening for recombinant plasmids. The XL1-Blue strain was used as standard plasmid construction vehicle in this study unless otherwise indicated. The cells were obtained from Stratagene, La Jolla, CA, USA.

***E. coli* DH5α**, Genotype: $F^-$, φ80Δ*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17($rk^-$, $mk^+$), *phoA*, *supE*44, $\lambda^-$, *thi*-1, *gyrA*96, *relA*1. This strain shows high insert stability due to its *relA*1 mutation. It can be used for blue/white screening and tolerates large plasmids due to its *deoR* mutation. High plasmid yield can be obtained due to an *endA*1 mutation. This strain was originally used for chemo-competent cells but has a high transformation efficiency being used in a protocol for electro-competent cells too (*99*). Because an *Xba*I digestion of a plasmid preparation from this strain is possible, Dam-dependent methylation is considered to be absent in this strain. It was used for plasmid propagation with subsequent Dam susceptible restriction endonuclease digestion. This strain was initially obtained through Invitrogen, Bale, Switzerland and the methylation deficient mutant strain was obtained from Maria Magdalena Senn (unpublished).

## 2.2 Culture techniques

### 2.2.1 Growth media

LB (Luria Bertani) liquid medium:
- Bacto™ Yeast extract (BD, Sparks, MD, USA), 5 g
- Bacto™ Tryptone (BD), 10 g
- NaCl, 10 g
- Distilled water ad 1000 mL, adjustment of pH to 7.5 with 4 % NaOH

LB agar medium:
- LB medium, 1000 mL
- Granulated agar (Difco, Bale, Switzerland), 15 g

Middlebrook 7H9 OADC:
- Middlebrook 7H9 (Difco), 4.7 g
- Glycerol (Sigma, Buchs, Switzerland), 2 mL
- BBL™ Middlebrook OADC enrichment (BD), 100 mL
- 10% Tween-80, 5 mL
- Distilled water, ad 1000 mL

Middlebrook 7H10 OADC agar medium:
- Middlebrook 7H10 agar base (Difco), 19 g
- Glycerol (Sigma), 5 mL
- BBL™ Middlebrook OADC enrichment (BD), 100 mL
- Distilled water, ad 1000 mL

Middlebrook 7H10 OADC, 2 %/5%/10% sucrose agar media:

- Middlebrook 7H10 agar base (Difco), 19 g
- Glycerol (Sigma), 5 mL
- BBL™ Middlebrook OADC enrichment (BD), 100 mL
- 20 g (2%)/50 g (5%)/100 g (10%) cell media tested sucrose (Roth, Reinach, Switzerland)
- Distilled water, ad 1000 mL

CDC 5% Sheep blood agar:

- Pancreatic Digest of Casein (BD), 15 g
- Sodium Chloride (Sigma), 20 g
- Granulated agar (Difco), 20 g
- Yeast Extract (BD), 5 g
- Hemin (BD), 0.0005 g
- Vitamin K1 (from pharmacy), 0.01 g
- L-Cysteine (Sigma), 0.4 g
- Sheep blood, defibrinated (BD), 5%
- Distilled water, ad 1000 mL

All media were autoclaved. OADC enrichment was added under a laminar flow hood after the autoclaved media had cooled down to a temperature of 50 °C. When needed sucrose was added at the same time from a 50% stock solution, which was sterile filtered through a 250 µL Stericup™ Filter Unit (Millipore, Volketswil, Switzerland).

### 2.2.2 Storage media

Storage medium for *E. coli* strains:

- LB liquid medium, 100 mL
- Glycerol (Sigma), 100 mL

For storage cultures were harvested to a volume of 1 mL, 0.5 mL of storage medium was added, and the sample was immediately stored at -80 °C.

2 x Storage medium for *M. bovis* BCG (*100*):

- Bacto™ Tryptone (BD, Sparks, MD, USA), 6g
- Glycerol, 20 mL
- Distilled water, 80 mL

### 2.2.3 Cultivation

#### 2.2.3.1 *E. coli*

Cultures were grown on LB agar plates or in liquid LB medium at 37 °C unless otherwise indicated. The cultures were harvested after 14 to 20 hours. Liquid cultures were shaken with 200 rpm.

According to selection criteria media were supplemented with the following antibiotics:

- Ampicillin (Sigma-Aldrich Chemie, Buchs, Switzerland): 100 μg/mL.
- Kanamycin (Sigma-Aldrich Chemie, Buchs, Switzerland): 50 μg/mL.
- Hygromycin B (Scientific Inc., San Diego, CA, USA): 100 μg/mL.

#### 2.2.3.2 *Mycobacterium bovis* BCG

All vaccine strains arrived in freeze-dried (lyophilized) state and were reconstituted according to the manuals of the respective manufactures. Briefly, ampoules were opened under laminar flow. 1 mL of liquid Middlebrook 7H9 medium was added and the bacteria were resuspended in medium. The resulting suspension was added to 10 mL of Middlebrook 7H9 medium and initially incubated at 37 °C for seven days. The culture bottles were weakly shaken every day to prevent cell aggregation. When logarithmic growth was noted by adequately increasing $OD_{600}$, Middlebrook 7H10 plates were inoculated and incubated for three weeks at 37 °C. Additionally sheep blood agar plates were inoculated with 100 μL from the initial 7H9 culture to exclude the presence of contaminants. All passage steps, i.e. inoculation from the previous to the next culture medium were noted with strain name, media type, culture duration, passage number, date of inoculation and specific remarks on phenotypic growth characteristics. All culture techniques were performed in a laminar flow hood dedicated to the vaccine study. All solid media culture plates were sealed with UV sterilized parafilm to prevent fungal contamination.

Selective Middlebrook media were supplemented with the following antibiotics as needed:

- Kanamycin (Sigma-Aldrich Chemie, Buchs, Switzerland): 25 μg/mL.
- Hygromycin B (Scientific Inc., San Diego, CA, USA): 50 μg/mL.

For the counterselection of transformed mycobacteria sucrose was added in the following percentages (weight/volume): 2%, 5% and 10%.

For DNA damage induction experiments *M. bovis* BCG substrains and *recA* mutants were grown in Middlebrook 7H9 medium in the presence of 10% (vol/vol) OADC in motionless tissue culture flasks in a 37 °C incubator. To generate DNA damage in SOS response induction experiments as described by Papavinasasundaram et al. (*80*), mitomycin C (0.2 μg $ml^{-1}$) was added to 20 mL of growing cultures (at an $OD_{600}$ of 0.6) and incubated for 24 h. Thereafter bacteria were harvested by centrifugation and prepared for cell lysis and Western blot analysis (see 2.15).

Remarks concerning the group assignment of *M. bovis* BCG vaccine strains in the classification of human pathogenic organisms: The current official classification list of the Swiss Agency for the Environment, Forest and Landscape assigns unmodified, wild-type *M. bovis* to the group of biosafety level 3 organisms but BCG vaccine strains are assigned to the biosafety level 2 group (*101*). All BCG substrains in this study could be

handled in biosafety level 2 laboratories because all six substrains had a solid documentation proofing their use as vaccine strains in humans.

## 2.3 Minimal inhibitory concentration assay in *E. coli*

The degree of resistance towards hygromycin B of *E. coli* DH5α after transformation with hygromycin resistance conferring plasmids was determined on 96-well microtiter plates (Greiner, Frickenhausen, Germany) using the following MIC assay: Three different single colonies of DH5α cells were collected from LB plates one day after transformation with a plasmid containing the antibiotic resistance gene to be tested and 50 mL LB precultures were inoculated with the three colonies. After 14 hours $OD_{600}$ was determined and all cultures were diluted to an equal $OD_{600}$ of 0.02. 25 μL of diluted culture was then mixed with 25 μL of antibiotic solution containing a decreasing amount of antibiotic along a row of a microtiter plate. Final antibiotic culture concentrations from 500 μg $mL^{-1}$ to 1.95 μg $mL^{-1}$ with decay of factor two from well to well were chosen. After an incubation period of additional 14 hours, the lowest concentration at which no visible growth could be observed was considered as minimal inhibitory concentration. The derived MIC values were compared to MIC values of non-transformant parental DH5α.

## 2.4 Generation of electrocompetent cells and transformation

### *2.4.1 E. coli*

*E. coli* XL1-blue or DH5α were streaked on LB agar from frozen stock and incubated over night. A single colony was picked and 15 mL of liquid LB was inoculated. 1 L Erlenmeyer culture flasks were filled with 500 mL of LB and 5 mL from the starter culture was added after an initial incubation of 12 to 14 hours. The 500 ml cultures were incubated for 3 to 5 hours and at an $OD_{600}$ of 0.4 - 0.6 the cultures were split on 50 mL falcon tubes. Those were cooled on ice water during 15 minutes and thereafter centrifuged at 4 °C with 5000 x g for 20 minutes. For the subsequent procedures the cells were constantly held at 4 °C. After centrifugation, the supernatant medium was discarded and the cell pellet was resuspended ice-cold 1 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.0-7.2 in half of the initial volume. The content of the falcon tubes was pooled to half of the originating number and again centrifuged with 5000 x g for 20 minutes. The washing step with HEPES and centrifugation was repeated once with half of the volume and further pooling. After discarding the supernatant the cells were resuspended in a total volume of 150 mL cold 10% Glycerol (volume/volume), centrifuged with 5000 x g for 20 minutes, the supernatant was discarded, the cells were pooled to one tube and 50 mL of 10 % glycerol was added. Finally, after an additional centrifugation step the supernatant was discarded and 2.5 mL 10% glycerol were added. The cells were aliquoted in 40 μL, frozen in liquid nitrogen and stored at -80 °C until further use.

Electroporation was performed with a Gene Pulser XCell™ electroporation system (Bio-Rad Laboratories, Hercules, CA, USA) in 0.1 cm electroporation cuvettes. 40 μL of electrocompetent cells were thawed slowly on ice and mixed with 1-2 ng of supercoiled plasmid DNA or 2 μL ligation sample. The transformation was achieved with a single pulse at 1.8 kV, 200 Ω and 25 μF followed by immediate addition of 1 mL LB me-

dium. The transformants were incubated at 37 °C, 200 rpm for 1 hour and then selected on LB plates containing the appropriate antibiotics.

### 2.4.2 *Mycobacterium bovis* BCG

The applied technique for the generation of electrocompetent *M. bovis* BCG was described by Sander et al. (*5*). The mentioned protocol was modified. Cells were not grown in rolling culture flasks but in tissue culture flasks, which were not in motion.

Briefly, after initial reconstitution in Middlebrook 7H9 medium during seven days *M. bovis* BCG vaccine strains were streaked to solid 7H10 medium to obtain single colonies. After approximately three weeks a single colony was picked and inoculated in 10 mL of 7H9 in a 25 mL tissue culture flask. The cells were incubated for 10 days at 37 °C and the settled cells were resuspended twice a day by motion. Growth was measured by quantifying the optical density at 600 nm. 1 mL of this preculture was inoculated in 200 mL of 7H9 and cultured in tissue culture flasks. Bacteria were cultured until an $OD_{600}$ of 0.7 to 1.0 was reached (approximately after 10 to 14 days). Once the culture had reached the desired density, 0.1 volumes of 2 M glycine (final concentration of 1.5% weight/volume) was added. The incubation was continued for 20 to 24 hours. All following steps were performed at room temperature: The culture was split into 50 mL falcon tubes and the cells were collected by centrifugation at 5000 x g for 30 minutes. The supernatant medium was discarded and the cells were diligently resuspended in an equal volume of 10% glycerol (i.e. 50 mL per centrifugation tube). The centrifugation at 5000 x g was repeated and the cells were resuspended in approximately 40 mL of 10% glycerol per tube. The centrifugation was repeated by further decreasing the total volume. The pellet of four centrifugation tubes were pooled into two, 40 mL of 10% glycerol added to each of them. After an additional centrifugation step the content of the two remaining tubes was pooled and 40 mL of 10% glycerol was added. Finally, after a last centrifugation step, 4 mL ($1/50^{th}$ of the initial culture volume) of 10% glycerol was added. For each electroporation experiment 400 µL of the final 4 mL cell suspension was taken. In an attempt to increase the number of transformants up to 5 µg of supercoiled plasmid DNA was used in a single transformation experiment (the limiting factor for the amount of vector DNA was the decreasing electrical resistance in the transformation cuvette and linked to this the risk of light bow formation and vaporization of the sample). Where indicated, 100 mJ $cm^{-2}$ UV irradiation was applied to the vector DNA on a UV Gene Linker 1800 (Stratagene, La Jolla, CA, USA) in order to assess whether DNA damage increases gene replacement (*102*). The DNA and competent cells were gently mixed and left aside at room temperature for 5 to 30 minutes. The cells were electroporated in a 4 mm electroporation cuvette with a single pulse and exponential voltage decay using the following settings on a Gene Pulser XCell™ electroporation system (Bio-Rad Laboratories, Hercules, CA, USA): 2.5 kV, 25 µF, 1000Ω. An integrative vector pMV361-*hyg* (*103*) was included as control for the competency of the cells. Immediately after a single electroporation pulse the cells were diluted in 4.6 mL of 7H9 medium and incubated for 24 hours at 37 °C. 100 µL of the liquid culture was then plated on 7H10 containing the appropriate antibiotic (hygromycin B or kanamycin). In addition, 1 mL of the culture was concentrated by centrifugation in a microfuge at 2700 x g, resuspended in 100 µL of saline (0.9%) and plated to the appropriate antibiotic 7H10 agar. The plates were sealed with parafilm and incubated for 14 to 28 d. Single colonies were picked, amplified, and analysed by PCR for genetic characterization of the transformants.

### 2.4.3 Determination of transformation efficiency

To calculate the efficiency of transformation a 400 µL vial of every batch was transformed with 1 µg scDNA of the integrative pMV361-*hyg* plasmid (*103*) from a stock solution [1000 ng $\mu L^{-1}$]. Transformation efficiency per ng of integrative plasmid was calculated as follows:

(a) Calculation of DNA in the transformation sample:

**Formula 1** $$c_{DNA} = \frac{m_{DNA}}{v}$$

$c_{DNA}$ is the concentration of the plasmid DNA in the transformation sample, $m_{DNA}$ is the mass of the added control plasmid, v is volume of the transformation sample. For this study 1000 ng of pMV361-*hyg* was used as a control. The volume of the transformation sample was approx. 400 µL. This results in a $c_{DNA}$ of 2.5 ng $\mu L^{-1}$.

(b) Calculation of the concentration of the control plasmid in the diluted culture after transformation:

**Formula 2** $$c_{DNA,diluted} = \frac{c_{DNA}}{d} = \frac{m_{DNA}}{vd}$$

$c_{DNA,diluted}$ is the concentration of the control plasmid in the diluted culture, d is the dilutional coefficient. In this example 4.6 mL of 7H9 medium was added to the 400 µL sample (d = 12.5, $c_{DNA,diluted}$ = 0.2 ng $\mu L^{-1}$).

(c) Concentration of plasmid DNA on the control solid medium plate was calculated:

**Formula 3** $$m_{DNA,plate} = c_{DNA,diluted} \cdot f \cdot v_{diluted} = \frac{m_{DNA} f v_{diluted}}{vd}$$

$m_{DNA,plate}$ is the concentration of the plasmid on the solid medium plate, f is the fraction of the initial culture after transformation that is used to inoculate the control plate, and $v_{diluted}$ is the total volume of the diluted culture. For this experiment 1 mL from the diluted culture was centrifuged, cells were harvested and plated on 7H10-Hyg or 7H10-Kan ($v_{diluted}$ = 1000 µL, f = 0.2, $m_{DNA,plate}$ = 200 ng per plate).

(d) Transformation efficiency (ε) per ng of control plasmid DNA was calculated:

**Formula 4** $$\varepsilon = \frac{n}{m_{DNA,plate}} = \frac{nvd}{m_{DNA} f v_{diluted}}$$

n is the number of colony forming units per plate.

## 2.5 Macroscopic examination of cultures

After initial revitalisation in liquid Middlebrook 7H9 medium, all vaccine strains were cultured on Middlebrook 7H10 agar and colony morphology was examined and photographically documented with a Canon D1 digital camera and 105 mm macro lens (Canon, Dietikon, Switzerland) on day 21 of culture on solid medium. Under appropriate conditions of growth colonies showing fine wrinkling (rugosity) of their surface, which is characteristic of certain BCG substrains, can be distinguished from colonies with a smoother non-rugose morphology that is characteristic of some other BCG substrains.

## 2.6 Microscopic examination of cultures

The initial liquid 7H9-cultures of the vaccine strains were examined microscopically: A Ziehl-Neelsen staining and an Auramine-Rhodamine fluorescence staining was performed. The samples were inspected under a standard light microscope at 400 x magnification and fluorescence microscopy for the Auramine-Rhodamine stained samples was done on a Zeiss Axioskop microscope (Carl Zeiss, Feldbach, Switzerland) using green excitation light ($\lambda$=554 nm) result in emission of red light ($\lambda$=627 nm). Bacterial morphology, absence of contaminants and staining behaviour was noted and documented with a Canon G1 digital camera.

## 2.7 Molecular identification

The current control methods recommended by the WHO (*104*) and the European Pharmacopoeia Commission (*105*) for the confirmation of identity of BCG vaccines involve microscopic examination, demonstrating their acid-fastness as outlined in 2.6, and macroscopic examination of the characteristic appearance of colonies growing on solid medium (as mentioned in 2.5). This methodology is insufficient to differentiate between BCG and related pathogenic mycobacteria and it doesn't allow the separation of BCG substrains. Other means of classification rely on biochemical and immunological techniques. Early BCG substrains, derived prior to 1926 secrete MPB70 antigens (*106*), produce methoxymycolates (*107, 108*) and contain two copies of the insertion element IS986 (*109*). Restriction fragment lengths have been used with low reproducibility to differentiate BCG substrains (*110*).

Recent publications by Marcel Behr and co-workers applying a microarray technique have demonstrated that substrains of *M. bovis* BCG own a total of 16 specific chromosomal sequence deletions greater than 80 bps (Table 2) in comparison to the published *M. tuberculosis* H37Rv genome (*43, 48, 91, 111-115*). However, the close relationship of BCG to the other members of the *M. tuberculosis* complex, which share > 99.9% identity at the DNA level (*59*), requires very specific molecular tools to distinguish various BCG substrains basing on these deletions. Strain specific deletions can be used to differentiate *M. bovis* BCG from *M. tuberculosis* and are also applicable to determinate BCG substrains. For this study several known regions of difference (RD) were examined with PCR assays in a way to exclude batch contamination and assure identity of the used strains.

In different substrain not only deletions greater than 80 bps have been published but also single nucleotide polymorphisms have been documented, e.g. at locus *Mb3700* (*Rv3676*) which encodes for a putative transcriptional regulatory protein, of the CRP/FNR family (*116*). A comparative sequence analysis for this locus was done too in all included vaccine strains and *M. tuberculosis* H37Rv.

In one case a species identification was necessary because the specific culture showed an exceptionally slow growth: Genus identification was confirmed by sequencing of the 16S rRNA gene exactly as described in (*117*); the species was determined by characterisation of the genome with respect to RD1.

**Table 2.** Regions of deletion (RD and nRD) from the BCG family of vaccine strains listed according to their location in the *M. tuberculosis* H37Rv genome (in *italics*: according to the *M. bovis* AF2122/97 genome). Genome base pair numbers are indicated according to the Geno-List genome browser at the Pasteur Institute (*118*). Deletion attributes include the BCGs from which they are deleted, start and end point in the H37Rv/*AF2122* genome, length in base pairs (bps), and the open reading frames (ORFs) affected by the deletion. Parts of the table have been published by Mostowy et al. (*115*).

| Deleted sequence | Deleted from | Start | End | Length (bps) | Affected ORFs |
|---|---|---|---|---|---|
| RD8 | BCG Frappier and Connaught | 378269<br>*379299* | 381696<br>*382726* | 3428 | *Rv0309* to *Rv0312* |
| nRD18 | BCG Pasteur, Phipps, Frappier, Connaught, Tice | 1332920<br>*1334170* | 1334466<br>*1335715* | 1547 | *Rv1189* to *Rv1191* |
| RD14 | BCG Pasteur | 1998225<br>*1988663* | 2007297<br>*1997753* | 9073 | *Rv1765c* to *Rv1773c*<br>[*Rv1766* to *Rv1773c* (*115*)] |
| RDDenmark/Glaxo | BCG Denmark and Glaxo | 2052590<br>*2043044* | 2053316<br>*2043768* | 726 | *Rv1810* to *Rv1811* |
| RD2 | BCG Pasteur, Phipps, Frappier, Connaught, Tice, Denmark, Glaxo, Prague | 2221057<br>*2199971* | 2231845<br>*2210759* | 10788 | *Rv1978* to *Rv1988* |
| RD16 | BCG Moreau | 3817365<br>*3769702* | 3824973<br>*3777601* | 7608 | *Rv3400* to *Rv3405c* |
| RDFrappier | BCG Frappier | 3914224<br>*3859847* | 3916195<br>*3861818* | 1971 | *Rv3495c* to *Rv3497c* |
| RDRussia | BCG Russia | 4140085<br>*4077433* | 4141688<br>*4079036* | 1603 | *Rv3697c* to *Rv3698* |
| RD1 | All BCG substrains | 4350262<br>*4286582* | 4359720<br>*4296085* | 9458 | *Rv3871* to *Rv3879c* |
| Total BCG family | | | | 46202 | |

### 2.7.1 Specific PCR for regions of difference and CRP polymorphism

All PCR assays were performed on either a Tpersonal or a Tgradient Thermocycler (Biometra, Göttingen, Germany), which are essentially comparable devices with heatable lids (heated to 105 °C for the full term of the reaction); no mineral oil had to be used. PCR reactions were undertaken in 200 µL thin walled PCR tubes (Eppendorf, Bale Switzerland) in a reaction volume of 50 µL. *Taq* polymerase was used unless otherwise indicated (kit from Roche Diagnostics, Bale, Switzerland). Primers were ordered from Microsynth, Belgach, Switzerland. All primers are listed in Table 4. The primers were compared with the published sequence of *M. bovis* AF2122/97 using BLASTN (DNA vs. DNA) software on the BoviList and TubercuList server (see 2.15.1) to exclude mispriming of the oligonucleotides. The PCR analyses were performed on all included BCG strains and *M. tuberculosis* H37Rv.

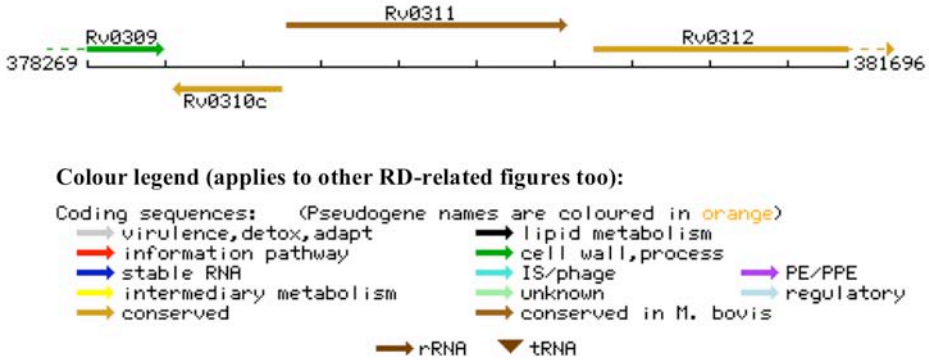
#### 2.7.1.1 RD8

The following multi primer assay PCR was set up to determine whether RD8 was deleted or present in a specific BCG substrain: Primers RD8-flank.F (10mM, 1µL), RD8-flank.R (10mM, 1µL) and RD8-Rv0309.int.R (10mM, 5 µL) were mixed with 10 µL of a DNA mickle preparation (see 2.8.4.1). 23 µL of PCR master mix consisting of 4.0 µL dNTP (2.5 mM each), 5.0 µL 10x *Taq* Gold Puffer (Roche Diagnostics, Bale, Switzerland), 0.5 µL Taq (Roche Diagnostics, Bale, Switzerland), and 13.5 µL distilled water was added. After composition the reaction was immediately started with the listed settings. The expected product lengths were: 292 bps, which

indicates absence of RD8 sequence and 286 bps, which indicates its presence. The amplificates were further differentiated by sequencing using the RD8-flank-F primer.

| Step # | Repetitions | Duration | Temperature | Purpose |
|---|---|---|---|---|
| 1 | 1 x | 5 min | 95 °C | initial denaturation |
| 2 | 10 x | 30 s | 95 °C | denaturation |
| 3 | | 30 s | 59 °C | primer annealing |
| 4 | | 50 s | 72 °C | elongation |
| 5 | 25 x | 30 s | 95 °C | denaturation |
| 6 | | 30 s | 55 °C | primer annealing |
| 7 | | 50 s | 72 °C | elongation |
| 8 | 1 x | 7 min | 72 °C | final elongation |
| 9 | | pause | 4 °C | |

**Figure 3.** PCR program for RD8-amplification using *Taq* polymerase.



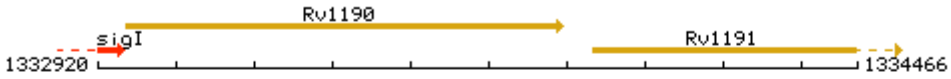

```
Mb   gtgggaaccg gcggttcgac ggccaagatg gatgtctacc aacgcaccgc cgccggctgg      -200
Mt   gtgggaaccg gcggttcgac ggccaagatg gatgtctacc aacgcaccgc cgccggctgg

     cagccgctca agaccggtat caccacccat atcggttcgg cgggcatggc gccggaagcc      -140
     cagccgctca agaccggtat caccacccat atcggttcgg cgggcatggc gccggaagcc

     aagagcggat atccggccac tccgatgggg gtttacagcc tggactccgc ttttggcacc      - 80
     aagagcggat atccggccac tccgatgggg gtttacagcc tggactccgc ttttggcacc

     gcgccgaatc ccggtggcgg Start RD8                          End RD8       - 20
     gcgccgaatc ccggtggcgg gttg..          3428 bps            ..gacg

     tcatggtcat cgaccatgag gccttgaccg atcgcgagtt ggcctggtcg cagaccgact      + 60
     tcatggtcat cgaccatgag gccttgaccg atcgcgagtt ggcctggtcg cagaccgact

     tcccaagcga agctccggcg cgtttcgagg gcgactcgta taacgaaggc ggcccctgct      +120
     tcccaagcga agctccggcg cgtttcgagg gcgactcgta taacgaaggc ggcccctgct

     ggtcgatgcg tctgaacgcg gtcgagcccc ccaaaggacc agcgtggcgg cgaatccggg      +180
     ggtcgatgcg tctgaacgcg gtcgagcccc ccaaaggacc agcgtggcgg cgaatccggg

     tgtcgcagtt gctcatcggg                                                 +200
     tgtcgcagtt gctcatcggg
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 4.** Region of difference RD8, start 378269 (*379299*), end 381696 (382226), length 3428 bps. Drawing of affected ORFs and surrounding nucleotide sequence (flanking primers underlined) according to the published sequences and base pair numbers (*118*).

### 2.7.1.2 nRD18

For investigation of nRD18, PCR reactions with the flanking primer nRD18-flank.F, nRD18-flank.R (both of them at a concentration of 10 µM, 1 µL) and 10 µL DNA mickle preparation of the respective strain were used. The PCR reaction was set up as previously described for RD8 with an initial denaturation for 5 minutes at 95 °C followed by 35 cycles of 95 °C for 30 seconds, annealing at 55 °C for 30 seconds, elongation at 72 °C for 2 minutes. The products were extended at 72 °C for 7 minutes and stored at 4 °C until agarose gel electrophoretic analysis. A non-deleted nRD18 should result in a product length of 1796 bps whereas a deleted should give a product of 249 bps.



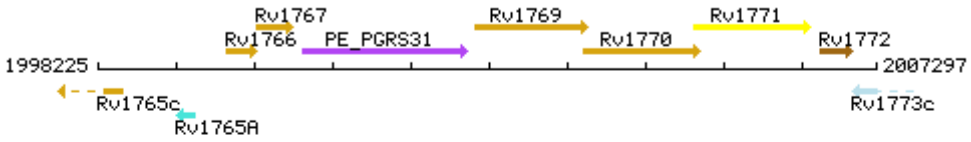

```
Mb    caaaggcgtt gtcgtcgtgg gcgcggatcg ggttggcccg accatcctgc gccactggag      -200
Mt    caaaggcgtt gtcgtcgtgg gcgcggatcg ggttggcccg accatcctgc gccactggag

      tcaccccgcc accgtcctgg tagcccagcc ggtgtgcggt caaccggcgg tgctggcctt      -140
      tcaccccgcc accgtcctgg tagcccagcc ggtgtgcggt caaccggcgg tgctggcctt

      tgtcaaccga gcgcttgccg gcgtgttggc cctgtcgatc gaggccggca agatcacaaa      - 80
      tgtcaaccga gcgcttgccg gcgtgttggc cctgtcgatc gaggccggca agatcacaaa

      aatccatgtc ttagtgcagc Start nRD18                      End nRD18      - 20
      aatccatgtc ttagtgcagc cttc..          1547 bps           ..tcgc

      gtcccctgcc ctcgacctgt atggcctgct ctcaccgcgg gccgaccggc atttgctggc        60
      gtcccctgcc ctcgacctgt atggcctgct ctcaccgcgg gccgaccggc atttgctggc

      tcggcccgag ttcaaggcga tgttcctcga cgatctgctc aacggtagtc gcaagcagct       120
      tcggcccgag ttcaaggcga tgttcctcga cgatctgctc aacggtagtc gcaagcagct

      cgctgcgccg ttcgccgatg tcatcgcctt tgcccgcgac tggggattcc ggctggacga       180
      cgctgcgccg ttcgccgatg tcatcgcctt tgcccgcgac tggggattcc ggctggacga

      ggtgaaagtc cccgtccgct                                                  200
      ggtgaaagtc cccgtccgct.
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 5.** Region of difference nRD18, start 332920 (1334170), end 1334466 (1335715), length 1547 bps. Drawing of affected ORFs and surrounding nucleotide sequence (flanking primers underlined) according to the published sequences and base pair numbers (*118*).

### 2.7.1.3 RD14

To uncover locus RD14 two different PCR had to be performed: For a deleted RD14, a PCR with two flanking primers RD14-flank.F and RD14-flank.R, amplifying a product length of 518 bps, was indicative. The respective settings were as mentioned for nRD18 whereby the elongation time could be shortened to 40 seconds and the annealing temperature had to be lowered to 46 °C. The second PCR for a non-deleted RD14 used two internal primers (RD14-Rv1769.int.F and RD14-Rv1769.int.R) and besides an annealing temperature of 48 °C unchanged thermocycler settings. The resulting amplificates were analyzed in a 2% Agarose gel electrophoresis.



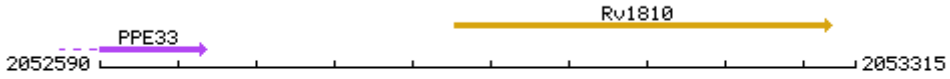

```
Mb    aactgctcat cggcgctcct ggtgatcgag ggccgcggtt ccggccgaaa atccggttcg      -200
Mt    aactgctcat cggcgctcct ggtgatcgag ggccgcggtt ccggccgaaa atccggttcg

      ggttcgggtc gcggttccaa cttgagcgcg gtccgcaact gattcaccgt ggcaaccccg      -140
      ggttcgggtc gcggttccaa cttgagcgcg gtccgcagct gattcaccgt ggcaacgccg

      gccaactgcg cataatgcgc atccgaaccc tcacccgccc gccccgcgat caccccaacc      - 80
      gccaactgcg cataatgcgc atccgaaccc tcacccgccc gccccgcgat caccccaacc

      tgatccaacg acaaccgccc Start RD14                        End RD14      - 20
      tgatccaacg acaaccgccc ctcc..           9073 bps           ..tcga

      gagcagccag cgccgctttt ccattcgcgg ttccgttcaa cgggaagcgg agcccgacgg        60
      gagcagccag cgccgctttt ccattcgcgg ttccgttcaa cgggaagcgg agcccgacgg

      ctgagaccgc acgcagccgg taagacgatt cgatctggtc gacaaaccac attcgctggc       120
      ctgagaccgc acgcagccgg taagacgatt cgatctggtc gacaaaccac attcgctggc

      cgcgcagtac cgacaggtcg accgtttcgc cgtcggtcgc gcgggcaact cgctcgacgg       180
      cgcgcagtac cgacaggtcg accgtttcgc cgtcggtcgc gcgggcaact cgctcgacgg

      tcggccggaa cgccgcggct                                                  200
      tcggccggaa cgccgcggct
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 6.** Region of difference RD14, start 1998225 (*1988663*), end 2007297 (*1997753*), length 9073 bps.

### 2.7.1.4 RDDenmark/Glaxo

The specific PCR used two flanking primers, RDDenmark-flank.F and RDDenmark-flank.R. An absent RDDenmark/Glaxo led to a product length of 234 bps versus 960 bps for a present RDDenmark/Glaxo. The PCR settings included the standard 35-cycle amplification with an annealing temperature of 55 °C and an elongation time of 1 minute. PCR products were analyzed on a 2% agarose gel.



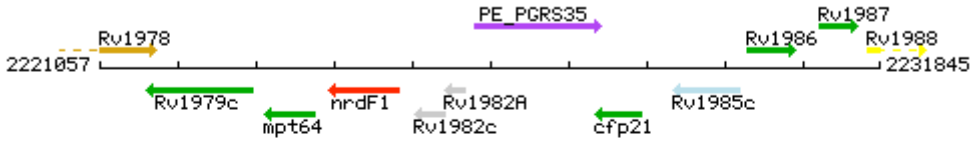

```
Mb    ctctggggag catgaccgga ggtgccctcg gtgccgccgc ccccgccatc tacacgggca      -200
Mt    ctctggggag catgaccgga ggtgccctcg gtgccgccgc ccccgccatc tacacgggca

      gtggcgcccg ggcccgcgcc aatgggggaa cgcccagcgc tgagccggtc aagctggagg      -140
      gtggcgcccg ggcccgcgcc aatgggggaa cgcccagcgc tgagccggtc aagctggagg

      ctgtcatcgc gcagctacaa aagcaaccgg acgcagtgcg acactggaat gtcgataagg      - 80
      ctgtcatcgc gcagctacaa aagcaaccgg acgcagtgcg acactggaat gtcgataagg

      ccgatcttga tggcctgctg Start RDDenmark/Glaxo   End RDDenmark/Glaxo      - 20
      ccgatcttga tggcctgctg gatc              726 bps             gccg

      tggtgcgcgg tggtgtggtg cgtccgaggg ggttgcgatg cacccggttc gcctaggctc        60
      tggtgcgcgg tggtgtggtg cgtccgaggg ggttgcgatg cacccggttc gcctaggctc

      aaactgctgt taacctgcgc gtggttggct gccgtggccg tcttgcgatc gggaaggact       120
      aaactgctgt taacctgcgc gtggttggct gccgtggccg tcttgcgatc gggaaggact

      cggcgtcatg caaacgctga ctgtcgccga tttcgctctc cggctggccg tcggagtggg       180
      cggcgtcatg caaacgctga ctgtcgccga tttcgctctc cggctggccg tcggagtggg

      ttgcggggcc attatcgggc                                                  200
      ttgcggggcc attatcgggc
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 7.** Region of difference RDDenmark/Glaxo, start 2052590 (*2043043*), end 2053316 (*2043768*), length 726 bps.

### 2.7.1.5 RD2

To characterise RD2 the following PCR multiprimer assay was performed: A 50 µl reaction mixture with the Triple Master PCR System (Vaudaux-Eppendorf, Bale, Switzerland) was prepared; 1 µL of two flanking primers (RD2-flank.F and RD2-flank.R) at a concentration of 10 mM, 1 µL of a 50 mM internal primer (RD2-Rv1979-int.R), 5 µL High Fidelity buffer, 0.5 µL Triple Master PCR enzyme blend, 4 µL of dNTP (each deoxynucleoside triphosphate at a concentration of 2.5 mM), 10 µL of a mickle DNA preparation and distilled water ad 50 µL were prepared in a 500 µL PCR tube. After denaturation at 95°C for 3 minutes, the reaction mixture was incubated for 40 cycles at 95 °C for 25 seconds and 15 seconds at 54 °C, followed by 80 seconds at 72 °C, and final extension for 7 minutes at 72 °C. The following product lengths were expected: For an absent RD2 1043 bps, for a present RD2: 1884 bps.



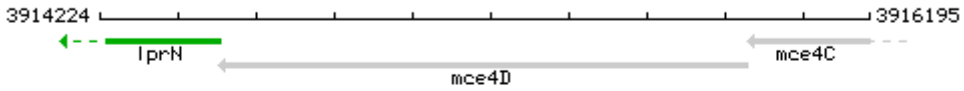

```
Mb    agcggtctca tagcattgcg aacacgggac gtcgagaggg gaagagctgc catgggtgag      -200
Mt    agcggtctca tagcattgcg aacacgggac gtcgagaggg gaagagctgc catgggtgag

      gcgaacatcc gcgagcaggc gatcgccacg atgccacggg gtggccccga cgcgtcttgg      -140
      gcgaacatcc gcgagcaggc gatcgccacg atgccacggg gtggccccga cgcgtcttgg

      ctggatcgtc gattccagac cgacgcactg gagtacctcg accgcgacga tgtgcccgat      - 80
      ctggatcgtc gattccagac cgacgcactg gagtacctcg accgcgacga tgtgcccgat

      gaggtcaaac agaagatcat Start RD2                          End RD2      - 20
      gaggtcaaac agaagatcat tgcc..          10788 bps           ..gcgc

      gggtggtcgc cgtggagttg cacccgcgac gagtcggtgt cctccgcgag cgattccctg        60
      gggtggtcgc cgtggagttg cacccgcgac gagtcggtgt cctccgcgag cgattccctg

      gcattaccgt ggtgcacgcg gacgccgcct cgatccggtt gcccggccgg ccgttccggg       120
      gcattaccgt ggtgcacgcg gacgccgcct cgatccggtt gcccggccgg ccgttccggg

      ttgtggcgaa cccgccgtac gggatttcgt cccgcctgct gcggacgctg ctggcaccca       180
      ttgtggcgaa cccgccgtac gggatttcgt cccgcctgct gcggacgctg ctggcaccca§

      acagcgggct tgtcgcggcc                                                  200
      acagcgggct tgtcgcggcc
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 8.** Region of difference RD2, start 2221057 (*2199971*), end 2231845 (*2210759*), length 10788 bps.

### 2.7.1.6 RDFrappier

RDFrappier was examined with a multiprimer PCR as follows: 1 µL RDFrappier-fl.F (10 pM), 1 µL RDFrappier-fl.R (10 pM), 1 µL RDFrappier-Rv3495c-int.R (50µM), 4.0 µL dNTP (2.5 mM each), 5.0 µL 10x Taq Gold Puffer (Roche Diagnostics, Bale, Switzerland), 0.5 µL Taq (Roche Diagnostics, Bale, Switzerland), 5.0 µL template (DNA mickle preparation), and distilled water ad 50 µL were assorted. The thermocycler settings included denaturation at 95 °C for 5 minutes, 35 cycles of 95 °C for 30 seconds, 50 °C for 30 seconds, and 72 °C for 20 seconds. Final extension was performed at 72 °C for 3 minutes. The expected product lengths were 177 bps for a *M. tuberculosis* like locus and 229 bps for the deleted locus.




```
Mb    caatgatctc ttccaacgca ccgacgttac ccttgttgac caccaccccc agcgccgaga     - 200
Mt    caatgatctc ttccaacgca ccgacgttac ccttgttgac caccaccccc agcgccgaga

      aaacctcctc ggtggtgggg aatcggtcgg tgttggcctc ggtgattctc gagccgtcaa     - 140
      aaacctcctc ggtggtgggg aatcggtcgg tgttggcctc ggtgattctc gagccgtcaa

      ccaacctccc ggtcggcggg cggtccgtcg gtggcgccag ctctacatgt aacgaaccca      - 80
      ccaacctccc ggtcggcggg cggtccgtcg gtggcgccag ctctacatgt aacgaaccca

      gcagcgaggt ctgggagacc Start RDFrappier             End RDFrappier      - 20
      gcagcgaggt ctgggagacc ttcg            1971 bps             aaag

      ccggaaatct gcgctgcaac atcgtcgatc ccagagatca acgcgctgag cgcggcccgc        60
      ccggaaatct gcgctgcaac atcgtcgatc ccagagatca acgcgctgag cgcggcccgc

      cgggcatcga gcgccgcaaa caactggttg ccgtcctcga ccagcttgtt gacctgttcg       120
      cgggcatcga gcgccgcaaa caactggttg ccgtcctcga ccagcttgtt gacctgttcg

      gcgcgttcgg acaacaccga tgtcaccgac ttggcgtgcg ccagcaggcc ttgcagcgct       180
      gcgcgttcgg acaacaccga tgtcaccgac ttggcgtgcg ccagcaggcc ttgcagcgct

      tcgtcgcgac gattcagggc                                                  200
      tcgtcgcgac gattcagggc
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 9.** Region of difference RDFrappier, start 3914224 (*3859847*), end 3916195 (*3861818*), length 1971 bps.

### 2.7.1.7 RDRussia

To examine the locus RDRussia two primers flanking the prospective deletion were used (RDRussia-flank.F and RDRussia-flank.R). As both primers have high annealing temperatures and the amplificates have a high GC-content of more than 65% for both amplificates the GC-Rich PCR System (Roche Diagnostics, Bale, Switzerland) was used for this reaction (for composition of PCR master mix see 2.7.1.9). The PCR reaction was performed with an initial denaturation at 95 °C for 3 minutes, 35 cycles of 95 °C for 30 seconds, followed by 30 seconds at 58 °C, and 2 minutes at 72 °C. After 35 cycles a final extending step at 72 °C was applied. Due to high primer melting temperatures GC-bond dissolving agents that have to be used to perform this PCR. More favourable primer pairs to examine this region could not be determined because of a persistently high GC-content of the genomic sequence of RDRussia.

For a present H37Rv homologue sequence at RDRussia locus, a 1956-bp amplificate was expected, for an absent a fragment length of 353 bps was expected.


```
Mb    cgtgatggcg catgagcgcg acgaggtggg cgtcgggaac cagattgcct cgcggcttga     - 200
Mt    cgtgatggcg catgagcgcg acgaggtggg cgtcgggaac cagattgcct cgcggcttga
      ccgggtcggc tactcgccga tagacgggcc agaatccgtt ggcctcgccg acctgccgca     - 140
      ccgggtcggc tactcgccga tagacgggcc agaatccgtt ggcctcgccg acctgccgca
      cgtgcggtcg tgaggtgaat tgctcgatgt tttcgacggc gacctcaggc gccagcggcg      - 80
      cgtgcggtcg tgaggtgaat tgctcgatgt tttcgacggc gacctcaggc gccagcggcg
      cacccaacaa cgtcggatgg Start of RDRussia           End of RDRussia      - 20
      cacccaacaa cgtcggatgg gtga             1603 bps              cgaa
      tttcctggca gccgccaggt acggcatgga cgcccagttg gattggccgg gcttgggcga        60
      tttcctggca gccgccaggt acggcatgga cgcccagttg gattggccgg gcttgggcga
      ggtgacgacg cgggagttgg tgttgggcac gttgttgcca atggcacacg agggactgcg       120
      ggtgacgacg cgggagttgg tgttgggcac gttgttgcca atggcacacg agggactgcg
      gcggtggggt gtcgacgcgg aggtacgcga ccggttcctg ggtgtcatcg gcggtcgcgc       180
      gcggtggggt gtcgacgcgg aggtacgcga ccggttcctg ggtgtcatcg gcggtcgcgc
      ccagaccggc cgcaacggcg                                                  200
      ccagaccggc cgcaacggcg
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 10.** Region of difference RDRussia, start 4140085 (*4077433*), end 4141688 (*4079036*), length 1603 bps.

### 2.7.1.8 RD1

RD1, which has been shown to be absent in all BCG strains was examined using a multiplex PCR technique as described by Talbot et al. (*119*). Briefly, a PCR containing primers ET1 and ET3, which are complementary to regions flanking RD1, and primer ET2 (complementary to DNA within RD1 deletion) was prepared. In strains without RD1, primers ET1 and ET3 bind and amplify a 200-bp region. In strains with RD1, ET2 and ET3 yield a 150-bp product (Figure 10). The PCR mix contained AmpliTaq Gold PCR buffer, 200 mM each deoxynucleoside triphosphate, 1.25 U of *Taq* DNA polymerase (AmpliTaq Gold; Roche, Bale, Switzerland), 5 pmol each of primers ET1 and ET3, and 25 pmol of ET2. The mixture was denatured for 3 min at 95 °C and cycled 40 times to 94 °C for 30 s and 65 °C for 1 min, followed by a final 10-min extension at 72 °C. The published protocol was modified, namely by simplifying the composition of the PCR reaction: PCR was performed without separate addition of KCl, gelatine and Tris-HCl. Instead of those the 10 x AmpliTaq Gold buffer as supplied by Roche was used. PCR amplificate band size was estimated on a 2% agarose gel by comparison to a 1500-bp DNA ladder (Fermentas, Hannover, MD, USA).

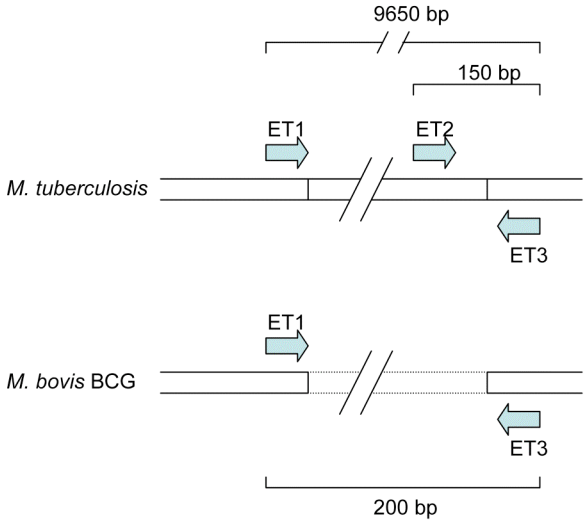
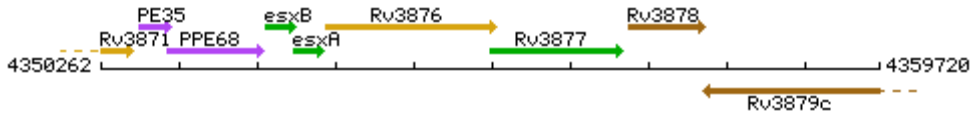



**Figure 11.** Multiplex PCR design. Wild-type *M. tuberculosis* complex DNA is represented by the parallel lines in the top diagram. The shaded region contains the 9.5-kb sequence RD1. RD1 is deleted in *M. bovis* BCG DNA, represented by the dashed lines in the lower diagram. The PCR primers ET1, ET2, and ET3 are shown as arrows oriented in the direction of amplification. This figure was adapted from (*119*), copyright 1997, American Society for Microbiology.




```
Mb    cctgctggac gcggtgccgg acacccatct gctgggcgcc ggcgcgatca accgcaacag     - 200
Mt    cctgctggac gcggtgccgg acacccatct gctgggcgcc ggcgcgatca accgcaacag

      cgcgtcgcta gacgaggccg ttcaagcact ggcggtcaac ctgaagaagc ggttgccgcc     - 140
      cgcgtcgcta gacgaggccg ttcaagcact ggcggtcaac ctgaagaagc ggttgccgcc

      gaccgacctg acgacggcgc agctacgctc gcgttcgtgg tggagcggat ttgacgtcgt      - 80
      gaccgacctg acgacggcgc agctacgctc gcgttcgtgg tggagcggat ttgacgtcgt

      gcttctggtc gacgattggc Start of RD1                    End of RD1      - 20
      gcttctggtc gacgattggc acat            9458 bps            ttcc

      acccagccgc ccggatccag catctgtctg gcatagctgc ccgtcggcct ggtaatactc        60
      acccagccgc ccggatccag catctgtctg gcatagctgc ccgtcggcct ggtaatactc

      atcccctact gccctcccca aaccgccaga tcgcctcgcg gatcaccgtc cggttggcct       120
      atcccctact gccctcccca aaccgccaga tcgcctcgcg gatcaccgtc cggttggcct

      ccggcatttc acgccggctc ggccgctgga tccaccccgc gccggtattc gcagtaaccc       180
      ccggcatttc acgccggctc ggccgctgga tccaccccgc gccggtattc gcagtaaccc

      gttgaatccg cgcgcatgat                                                  200
      gttgaatccg cgcgcatgat
```

Mb *M. bovis* AF2122/97
Mt *M. tuberculosis* H37Rv

**Figure 12.** Region of difference RD1, start 4350262 (*4286582*), end 4359720 (*4296085*), length 9458 bps. Primers ET1 (forward) and ET3 (reverse) are underlined.

### 2.7.1.9 Mb3700/Rv3676 (cAMP receptor protein, CRP)

The amplification and sequencing of the Rv3676 region was performed as described by Spreadbury et al. (*116*) with minor modifications. The published primer sequences in their study are part of the *Rv3676* gene and don't allow sequencing of the full gene length directly from the primary amplificate. To amplify *Mb3700*/*Rv3676* from genomic DNA two flanking primers Rv3676_for and Rv3676_rev (Figure 13) derived from the published sequence of *M. bovis* AF 2122/97 (*120*) were ordered. The expected amplificate had a length of 882 bps and a GC-content of 64%. Approximately 10 ng of genomic DNA were used as a template and amplified using the two flanking primers and the GC-Rich PCR System (Roche Diagnostics, Bale, Switzerland) in a 50 µL reaction with a composition according to the instructions of the manufacturer. No GC-Rich resolution buffer was added. The samples were incubated using a touchdown PCR protocol: Initial denaturation at 95 °C for 3 minutes was followed by 10 cycles of 95 °C for 30 seconds, 61 °C for 30 seconds and 72 °C for 50 seconds which was followed by 25 equal cycles with a lowered annealing temperature of 58 °C in stead of 61 °C. Final product extension was performed at 72 °C for 7 minutes.

The PCR products were purified by agarose gel extraction and used for direct DNA sequencing with the amplification primers mentioned above. The resulting sequence files were assembled and compared using SeqMan™ II (DNASTAR, Madison, WI, USA) sequence alignment software and sequence differences were noted.

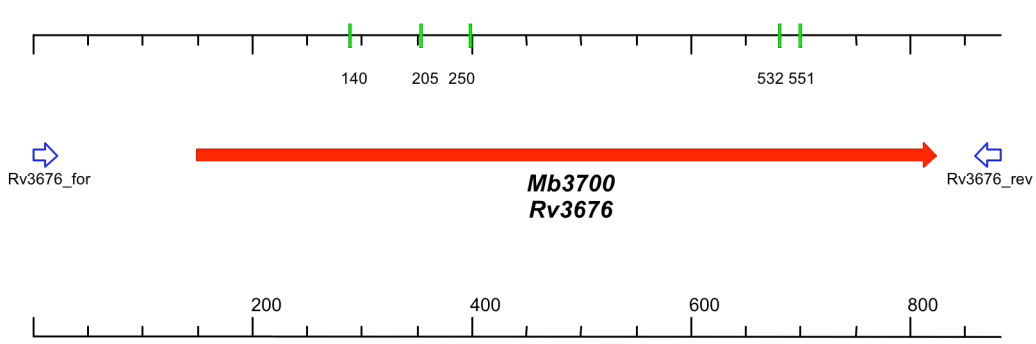



**Figure 13.** *Rv3676*/*Mb3700* and position of the flanking primers Rv3676_for and Rv3676_rev. Published point mutations (*116*) are marked with the corresponding base pair number (counting from the start of *Rv3676*). The amplificate length is 882 bps.

#### *2.7.1.10* Amplification and sequencing of *recA*

For sequence comparison, *recA* was amplified using *Pfu* DNA polymerase (Promega). A large-scale genomic DNA preparation (2.8.4.2) from a Middlebrook 7H10 plate culture served as template. 50 ng of genomic DNA was added to a master mix containing dNTPs (final concentration 200 µM each), recA3_start and recA3_stop primers (final concentration 0.8 µM each) and 10 x PCR buffer with $MgSO_4$ (supplied with the polymerase kit), and 1.25 u *Pfu* DNA polymerase. Cycling conditions were an initial 1 min 30 s lasting denaturation step at 95 °C, followed by 35 amplificatory cycles including denaturation at 95 °C for 30 s, primer annealing at 59 °C for 30 s, and elongation at 72 °C for 4 min (corresponding to an amplification product of 2562 bps). Final elongation at 72 °C was performed during 5 min. The PCR tube was then kept at 4 °C until fragment analysis by agarose gel electrophoresis and fragment excision was performed. 200 ng of the amplificated DNA were used for subsequent sequencing (see 2.10) using *recA* binding primers (Table 4).

## 2.8 DNA methods

### 2.8.1 Plasmid extraction from *E. coli*

For small scale plasmid isolation, 3 to 5 mL of LB containing the appropriate antibiotic was inoculated with a single bacterial colony from solid medium and incubated over night at 37 °C and constant shaking at 200 rpm. The cells from 1.5 mL culture were harvested by centrifugation at 10000 x g in a microfuge during 5 minutes. The following isolation steps were performed with a NucleoSpin® Plasmid Kit (Macherey Nagel, Düren Germany) according to the user's manual of the producer. This kit uses the alkaline SDS lysis protocol and adsorbs the plasmid DNA on a column with treated silica membrane from where it can be eluted with water or TE buffer.

For medium scale plasmid preparations a single bacterial colony from an overnight culture was inoculated to 150 mL of LB supplemented with the appropriate antibiotic. The culture was incubated for 10 to 14 hours at 37 °C and 200 rpm. Cells were collected by centrifugation for 15 minutes at 5000 x g. The plasmid preparation was performed with Promega's PureYield™ Plasmid Midiprep System (Catalys AG, Wallisellen, Switzerland) according to the instructions of the producer. This kit uses the alkaline SDS lysis protocol too and the plasmid DNA is finally eluted in a volume of 300 µL. For this study DNA was generally eluted with water to prevent enzyme inhibition by buffer contents (especially EDTA) in subsequent restriction assays.

### 2.8.2 Agarose gel electrophoresis and analysis

According to separatory needs (DNA lengths) 0.8% to 2% Agarose gels were produced from highly pure Agarose and TAE buffer pH 8.3 and immerged in TAE running buffer. DNA samples were mixed with 6 x loading dye (Fermentas) and filled in the gel pouches. DNA length separation was achieved with field strength of 5 V/cm for 1 to 2 hours. Gels were transferred to a 0.0002% ethidium bromide bath for 15 minutes and washed in deionised water during 5 minutes. Afterwards gels could be analysed under UV light (λ=312 nm).

### 2.8.3 DNA purification

#### 2.8.3.1 Ethanol precipitation

Ethanol precipitation was used to remove salts from DNA solutions and to concentrate DNA samples. To precipitate DNA in water solutions 2.5 M sodium acetate solution was added to a final concentration of 0.3 M. 2.5 volumes of 100% ethanol was added and the solution was cooled at -80 °C for an hour or at -20 °C over night. DNA was collected by centrifugation at 16000 x g in a microcentrifuge. The pelletted DNA was then washed twice with 70% ethanol (-20 °C) and dried at room temperature under a vacuum of 0.25 mbar in a Speedvac Concentrator (Inotech AG, Wohlen, Switzerland). The pellet was solved in an appropriate volume (50 - 200 µL, according to DNA amount) of distilled water or TE buffer after purification.

#### 2.8.3.2 Phenol extraction

Phenol extraction was used to purify chromosomal DNA and to remove enzymes after enzymatic manipulations of DNA. To extract DNA, an equal volume of phenol/chloroform/isoamyl alcohol 25/24/1 mixture was added to a DNA solution. The resulting mixture was centrifuged at 16000 x g for 20 minutes and the aequos phase was pipetted in a new microcentrifuge tube. The DNA was then precipitated using ethanol precipitation described above.

#### 2.8.3.3 Gel extraction

Gel extraction was applied to isolate a specific fragment from a DNA sample. Therefore a digested DNA sample was loaded on an Agarose gel and separated electrophoretically as described above. After separation the gel was stained with ethidium bromide and analyzed under UV light. The gel piece containing the desired DNA fragment was cut out under direct UV light visualization (λ=312 nm, minimal UV exposure time). The DNA was isolated from the gel piece using a QIAquick gel extraction kit (Qiagen, Bale, Switzerland) according to the manufacturer's user's manual.

## 2.8.4 Isolation of chromosomal DNA from *Mycobacterium bovis* BCG

### 2.8.4.1 Minimal scale DNA preparation

For PCR screening of transformants the following DNA preparation method was applied (*117*): A small colony from a slant was picked and transferred into a 1.5 mL screw cap microcentrifuge tube with 50 µL 0.9% NaCl. One or two 1 mm acid washed sterile glass beads (Sigma-Aldrich Chemie, Buchs, Switzerland) were added and the tube was vortexed shortly to separate the cells. Alternatively the colony was suspended using a volume of 400 µL normal saline and a 26 gauche needle by several strokes through the needle. 10 µL were removed for further culture. The remaining cell suspension was heated for 10 min at 80 °C to inactivate the mycobacteria. The tube was centrifugated at maximum speed (approximately 17000 x g) in a microcentrifuge for 10 minutes. The supernatant and glass beads were discarded. 100 µL of TE buffer and a loopful of acid washed glass beads (100 µm diameter) were added to the cell pellet and the tubes were placed on mickle tissue disintegrator at maximal speed for 2 minutes. Then the sample was centrifugated at maximum speed in a microcentrifuge for 5 minutes. The supernatant was transferred in a new tube, 5 µL aliquots were used for PCR.

### 2.8.4.2 Large-scale DNA preparation

For Southern blot analyses and sequencing DNA was prepared from a three week culture on Middlebrook 7H10. The bacteria were harvested from half a Petri dish and suspended in 340 µL of TE buffer. The bacteria were heat inactivated at 80 °C during 20 minutes. 4 µL 10% Tween-80 and 20 µL lysozyme (40 mg/mL, Roche, Bale, Switzerland) were added. The sample was incubated at 37 °C for two hours. Afterwards 40 µL 10% SDS and 20 µL proteinase K (2 mg/mL, Roche, Bale, Switzerland) were gently mixed to the sample. The preparation was incubated at 60 °C for one hour. DNA was extracted by adding 400 µL phenol/chloroform/isoamyl alcohol 25/24/1 and centrifugation with 16000 x g at 4 °C. Thereby the DNA was solved in the aequous phase which was transferred in a new tube. 8 µL 5 M NaCl and 1 mL of ethanol were added to precipitate the DNA. The sample was cooled to -80 °C for one hour and centrifugated with 16000 x g for 20 minutes. The supernatant was carefully removed with a vacuum pipette. The DNA pellet was washed twice with 500 µL cold 70% Ethanol, dried under a vacuum of 200 mbar in a Speedvac Concentrator (Inotech AG, Wohlen, Switzerland) and depending on the visual size of the pellet 50 to 200 µL distilled water were used to solve the pellet. DNA solutions were stored at 4 °C for short terms (1 week) and at -20 °C for longer periods.

## 2.8.5 Determination of DNA concentration

Nucleic acids were quantified by measuring absorbance at 260 nm in an Ultrospec® 2100 Pro or GeneQuant™ ultraviolet/visible light spectrometer (Amersham Bioscience, Otlefingen, Switzerland). A DNA solution with an optical density of 1.0 at 260 nm has a concentration of 50 µg/mL in a 10 mm path length quartz cell. Linear regression was used to calculate concentration in a 100 µL volume in a 1:50 diluted sample. For determination of DNA purity, the $A_{260/280}$ and $A_{260/230}$ coefficients were determined photometrically. An $A_{260/280} < 1.8$ indicated contamination of the DNA preparation with protein or aromatic substances such as phenol; while an $A_{260/230} > 2.0$ indicated possible contamination with RNA.

## 2.8.6 Southern blot analysis

### 2.8.6.1 Blotting

For Southern blot analysis, 1 µg of genomic DNA was digested over night with an appropriate restriction endonuclease and separated by gel electrophoresis in a 1.2% agarose gel. The gel was immerged in 0.25 M HCl for 10 minutes and then the bath solution was exchanged for a denaturation solution (1.5 M NaCl, 0.5 M NaOH) and shaken for 15 min. For another 15 minutes, the gel was incubated in neutralization solution (0.3 M NaCl, 0.03 M sodium citrate). Finally the gel was transferred to a bath of 2 x SSC (1:10 dilution of 20 x SSC stock: 30 mM sodium citrate, 3 M NaCl, pH 7.0) and incubated for 15 minutes. The DNA was blotted from the gel to a Hybond-N membrane (Roche, Bale, Switzerland) using a Vacu-Blot System (Biometra, Göttingen, Germany) with 20 x SSC as blotting buffer and a vacuum of 500 mbar for 90 minutes. After blotting, the Hybond-N membrane was quickly washed in 2 x SSC and the DNA was bound to the membrane by crosslinking with 1200 mJ $cm^{-2}$ UV radiation on each side in an UV Gene Linker 1800 (Stratagene, La Jolla, CA, USA).

### 2.8.6.2 Hybridization

Prehybridization was performed by incubation at 42 °C for one hour in hybridization tubes containing 10 mL of hybridization solution (0.7 g SDS, 5 mL of formamide, 2.5 mL of 20 x SSC, 0.5 mL of 1 M $NaPO_4$, 10 mg N-laurosylsarcosine, 2 mL 10 x blocking solution as described in 2.8.6.3). Hybridization was performed at 42 °C over night by addition of 100 ng digoxigenin (DIG) labelled probe to the hybridization solution. The probe was labelled with a DIG-High Prime DNA Labelling Kit (Roche, Bale, Switzerland) according to the instructions of the manufacturer. The approximate concentration of the DNA probe stock was determined by 5 serial logarithmic dilutions of the probe and comparison of the signal strength on photographic film with a dilution series of labelled DNA with known concentration (supplied in kit) in a CDP-Star assay mentioned under paragraph 2.8.6.3.

### 2.8.6.3 Detection

After decanting of the hybridization solution, the membrane was washed twice with 2 x SSC / 0.1% SDS solution for 5 minutes at room temperature and twice with 0.1 x SSC / 0.1% SDS solution for 15 minutes at 68 °C. Subsequently, the membrane was shortly equilibrated in 50 mL washing buffer. The washing buffer was compounded from maleic acid buffer (1L maleic acid buffer: 11.6 g maleic acid, 8.7 NaCl, 7.8 g NaOH, pH 7.5) by addition of 0.6% (volume/volume) 10% Tween-80. After washing, the binding capacity of the membrane was blocked by incubation at room temperature for 30 minutes with 50 mL 1 x blocking solution (10 x blocking solution: 25 g skim milk powder, maleic acid ad 250 mL). To detect DIG labelled DNA fragments, the blocking solution was replaced by another 50 mL of blocking solution containing 2.5 µL of Anti-digoxigen-alkaline-phosphatase Fab Fragments (Roche, Bale, Switzerland). After 30 min of incubation the membrane was washed twice for 15 minutes with washing buffer (same composition as mentioned above). The membrane was pored off, and 50 mL of detection buffer was added (0.1 M Tris-HCl, 0.1 M NaCl, and pH 9.5). After 3 minutes the detection buffer was replaced by 50 mL of CDP-Star (ultra-sensitive chloro-substituted 1,2-dioxetane chemiluminescent substrate for alkaline phosphatase, Roche, Bale, Switzerland) diluted 1:100 in detection buffer and incubated at room temperature and light protection for 5 minutes. The membrane was pored off and welded in a polyethylene bag. The chemiluminescent membrane was photographed with BioMax Light film (Eastman Kodak

Company, Rennes, Switzerland) with different exposure times in the range from 2 to 60 minutes (depending on the relative signal intensity).

### 2.8.7 Specific Southern analysis for the examination of the *recA* locus

Genomic DNA from BCG transformants was prepared as previously described and digested with *Sma*I. A 412 bps *Nhe*I/*Not*I fragment from plasmid pBS-KSII(-)-*recA* (H37Rv base pairs 3051677 to 3052088; according to TubercuList) was labelled with digoxigenin and used as probe. The probe is situated at the 5' end of the *recA* gene. After digestion with *Sma*I and hybridization with the mentioned probe, wild-type *recA* allele was expected to result in a labelled 2.0 kbs fragment. Single crossover transformants should either show a 2.0 kbs wild-type fragment and an additional fragment at 9.6 kbs demonstrating plasmid integration 5' of the *recA* gene or a signal at 4.0 kbs and at 5.6 kbs for integration 3' downstream of the *recA* gene. The *recA* knock-out mutant should result in a single signal at 5.9 kbs (Figure 47).

## 2.9 Enzymatic manipulations of DNA

### 2.9.1 Endonucleolytic restriction digestion of plasmid DNA

For analytical restriction digestion a minimal quantity of 200 ng supercoiled plasmid DNA or chromosomal DNA was mixed with the restriction enzymes mentioned in Table 3.

For preparational reactions 0.5 to 1 µg DNA was digested with 10 units of the respective endonucleolytic restriction enzyme and the specific 10 x concentrated buffer that was recommended by the enzyme manufacturer was added in correct dilution. The sample was incubated for at least two hours at the appropriate temperature (usually 37 °C unless especially noted). Digestions for cloning experiments and Southern blot analyses were performed over night to allow almost complete digestion.

**Table 3.** List of restriction enzymes used in this study.

| Name (synonyms) | recognition sequence (^ marks cut) | Supplier |
|---|---|---|
| *Apa*I | 5'-G G G C C^C-3'<br>3'-C^C C G G G-5' | Fermentas, Nunningen, Switzerland |
| *Bcl*I | 5'-T^G A T C A-3'<br>3'-A C T A G^T-5' | Fermentas |
| *Bgl*II | 5'-A^G A T C T-3'<br>3'-T C T A G^A-5' | Fermentas |
| *Bpu*1102I | 5'-G C^T N A G C-3'<br>3'-C G A N T^C G-5' | Fermentas |
| *Eco*RV (*Eco*32I) | 5'-G A T^A T C-3'<br>3'-C T A^T A G-5' | Fermentas |
| *Hind*III | 5'-A^A G C T T-3'<br>3'-T T C G A^A-5' | Fermentas |
| *Mun*I (*Mfe*I) | 5'-C^A A T T G-3'<br>3'-G T T A A^C-5' | Fermentas |
| *Nhe*I | 5'-G^C T A G C-3'<br>3'-C G A T C^G-5' | Fermentas |
| *Not*I | 5'-G C^G G C C G C-3'<br>3'-C G C C G G^C G-5' | Fermentas |
| *Nsi*I (*Mph*1103I) | 5'-A T G C A^T-3'<br>3'-T^A C G T A-5' | Fermentas |
| *Pst*I | 5'-C T G C A^G-3'<br>3'-G^A C G T C-5' | Amersham Biosciences, Otelfingen, Switzerland |
| *Sac*I | 5'-G A G C T^C-3'<br>3'-C^T C G A G-5' | Fermentas |
| *Sap*I (*Lgu*I) | 5'-G C T C T T C $(N)_1$^-3'<br>3'-C G A G A A G $(N)_4$^-5' | NEB, Ipswich, MA, USA |
| *Sca*I | 5'-A G T^A C T-3'<br>3'-T C A^T G A-5' | Amersham Biosciences |
| *Sex*AI | 5'-A^C C N G G T-3'<br>3'-T G G N C C^A-5' | NEB |
| *Sma*I | 5'-C C C^G G G-3'<br>3'-G G G^C C C-5' | Fermentas, NEB |
| *Sna*BI (*Eco*105I) | 5'-T A C^G T A-3'<br>3'-A T G^C A T-5' | Fermentas |
| *Spe*I (*Bcu*I) | 5'-A^C T A G T-3'<br>3'-T G A T C^A-5' | Fermentas |
| *Sph*I (*Pae*I) | 5'-G C A T G^C-3'<br>3'-C^G T A C G-5' | Fermentas |
| *Van*91I (*Pfl*MI) | 5'-C C A N N N N^N T G G-3'<br>3'-G G T N^N N N N A C C-5' | Fermentas |
| *Xba*I | 5'-T^C T A G A-3'<br>3'-A G A T C^T-5' | Fermentas |

### 2.9.1.1 Dephosphorylation of DNA

In cloning experiments the vector plasmid DNA was usually dephosphorylated at its 5'-ends. This was done by adding 6 units (in 6 µl) of SAP (shrimp alkaline phosphatase, Roche, Bale, Switzerland) and 4 µL of 10 x buffer (also from Roche, comes in same kit) to up to 10 µg of digested, open circle vector DNA in a total reaction volume of 40 µL. The mixture was incubated for two hours at 37 °C. The enzyme was inactivated by heating to 65 °C for 15 minutes. The DNA was purified by gel extraction or phenol extraction for further use.

### 2.9.1.2 Blunting of DNA fragments

In order to produce blunt ends DNA was treated with recombinant T4 DNA polymerase (Roche, Bale, Switzerland) in presence of an overextend of dNTPs. T4 polymerase is a DNA dependent DNA polymerase that combines 5'-3' polymerase with 3'-5 exonuclease activity and lacks 5'-3' exonuclease activity. In a total reaction volume of 50 µL purified DNA (up to 5 µg) was added to 1 unit (in 1 µL) of T4 DNA polymerase, the provided 10 x reaction buffer and dNTPs were added to a final concentration of 0.4 mM each. The reaction was incubated for exactly 15 minutes at 37 °C and stopped through 20 minutes lasting thermal inactivation at 80 °C.

### 2.9.1.3 Ligation

For ligation reactions DNA was purified as mentioned above and different molar ratios of vector and insert were prepared. A molar ratio of 1:3 (vector: insert) was included in the ligation experiment where exact concentration of the DNA samples could be determined (using absorption photometry). Quantities to be used were calculated using Formula 5 indicated below. Ligation reactions were performed in a total volume of 10 or 20 µL with 1 unit (in 1 µL) T4 DNA ligase (Roche, Bale, Switzerland) and the supplied 10 x ligation buffer.

**Formula 5.**
$$m_{insert} = \frac{m_{vector} \cdot l_{insert} \cdot r}{l_{vector}}$$

$m_{insert}$: mass of insert to be used in ng, $m_{vector}$: mass of vector in ng, $l_{insert}$: length of insert in base pairs, $l_{vector}$: length of vector in base pairs, $r$: molar ratio vector : insert

## 2.10 DNA Sequencing

DNA sequencing was done on the basis of a semi automated protocol using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Forster City, CA, USA). The kit uses Sanger's PCR chain-termination sequencing method modified with fluorescence marked dideoxynucleotides (ddNTPs) as PCR terminator (*121*) in a defined molar ratio to unmarked deoxynucleotides (dNTPs). The primary reaction is a single primer PCR to generate labelled amplificate fragments which end on a fluorescence marked ddNTP. Sequencing PCR started from a previously determined ssDNA primer (see Table 4 for complete list of primers used in this study). In a thin walled 0.5 mL PCR tube 600 to 1,000 ng of template DNA was given to 4 µL BigDye® reaction premix (ddNTPs, dNTPs and polymerase), 2 µL 5 x BigDye® buffer, the appropriate sequencing primer at a concentration of 1 pmol $\mu L^{-1}$ and deionised water to a final reaction volume of 20 µL were compounded. The following cycle PCR program was executed on a Tpersonal Thermocycler (Biometra, Göttingen, Germany): The closing lid of the device was heated to 105 °C for the full term of the reaction. Initially the sample was heated to 94 °C for 5 minutes to completely separate the template DNA strands. This was followed by 25 cycles of 96 °C for 30 seconds (strand separation), 55 °C (or lower with lower primer annealing temperatures, especially noted on primer Table 4) for 15 seconds and 60 °C for 4 minutes (elongation). Post reactive the PCR samples were purified using the PERFORMA® DTR Gel Filtration Cartridges (Edge Biosystems, Gaithersburg, MD; USA) with the recommended protocol. 5 µL eluate from the purification column was diluted with 15 µL distilled water. The so diluted amplificates were electrophoretically separated and the fluorescence signal of the terminal ddNTPs was determined using an ABI Prism 310 or 3100 sequencer (Applied Biosystems). The resulting elec-

tropherogram was analyzed using sequence analysis software ABI Prism 10 Collection in its version 3.4.1 (ABI Prism 310) and 3.7 (ABI Prism 3100).

## 2.11 PCR primers

PCR primers for direct sequencing and PCR amplification DNA were ordered in lyophilized state from Microsynth, Belgach, Switzerland and stored at -20 °C.

Oligonucleotide properties were calculated using Oligo Calculator, version 3.07 (Research Computing, Northwestern University Medical School, Chicago, IL, USA). This software also allowed excluding primers with self-complementarities and a built-in BLASTN query (see 2.15) allowed excluding primers with high probability of mispriming (i.e. several annealing sites).

### 2.11.1 Basic melting temperature calculations

The two standard approximation calculations were used to calculate approximate melting temperatures when selecting primers for PCR reactions. For sequences less than 14 nucleotides Formula 6 was applied:

**Formula 6.** $T_m = 2(wA + xT) + 4(yG + zC)$

where w, x, y, z are the number of the bases A, T, G, C in the sequence, respectively.

For sequences longer than 13 nucleotides, the equation used is Formula 7:

**Formula 7.** $T_m = 64.9 + \frac{41(yG + zC - 16.4)}{wA + xT + yG + zC}$

**Assumptions:**

Both equations assume that the annealing occurs under the standard conditions of 50 nM primer, 50 mM $Na^+$, and pH 7.0.

### 2.11.2 Salt adjusted melting temperature ($T_m$) calculations

A variation on two standard approximation calculations was used:

**Formula 8.** $T_m = 100.5 + \frac{41(yG + zC)}{wA + xT + yG + zC} - \frac{820}{wA + xT + yG + zC} + 16.6 \log_{10}[Na+]$

## 2.11.3 Comprehensive list of primers used in this study

**Table 4.** PCR primers used in this study.

| Name | 5'-3' sequence | Length | $T_m$ (°C)[¶] | GC-content (%) |
|---|---|---|---|---|
| 244[*] | CCC ACT GCT GCC TCC CGT AG | 20-mer | 60 | 70 |
| 259[*] | TTT CAC GAA CAA CGC GAC AA | 20-mer | 50 | 45 |
| 264[*] | TGC ACA CAG GCC ACA AGG GA | 20-mer | 56 | 60 |
| 271[*] | CTT AAC ACA TCG AAG TCG AAC | 21-mer | 50 | 43 |
| 285[*] | GAG AGT TTG ATC CTG GCT CAG | 21-mer | 54 | 52 |
| aph5' | CGT TTC CCG TTG AAT ATG GCT C | 22-mer | 55 | 50 |
| aph3' | CTC CTT CAT TAC AGA AAC GG | 20-mer | 50 | 45 |
| ET1 | AAG CGG TTG CCG CCG ACC GAC C | 22-mer | 64 | 73 |
| ET2 | CTG GCT ATA TTC CTG GGC CCG G | 22-mer | 60 | 64 |
| ET3 | GAG GCG ATC TGG CGG TTT GGG G | 22-mer | 62 | 68 |
| hyg_start | TCG ACG ACG CGG GGG AC | 19-mer | 57 | 76 |
| hyg_stop | TCG CAG CAG CGG GCT TCG | 18-mer | 57 | 72 |
| nRD18-flank.F | GCG GTG CTG GCC TTT GTC A | 19-mer | 55 | 63 |
| nRD18-flank.R | GCG GGC AAA GGC GAT GAC | 18-mer | 55 | 67 |
| P1 | CCC AGT CAC GAC GTT GTA AAA C | 22-mer | 55 | 50 |
| P2 | GCA GAC GTC CCC GAT CAA AG | 20-mer | 56 | 60 |
| P3 | CAT CAA ATA AAA CGA AAG GCT CAG | 24-mer | 52 | 38 |
| RD2-flank.F | CTC GAC CGC GAC GAT GTG C | 19-mer | 58 | 68 |
| RD2-flank.R | CCT CGT TGT CAC CGC GTA TG | 20-mer | 56 | 60 |
| RD2-Rv1979-int.R | ATC GGC ATC TAT GTC GGT GT | 20-mer | 52 | 50 |
| RD2-Rv1979.int.F | TAT AGC TCT CGG CAG GTT CC | 20-mer | 54 | 55 |
| RD8-flank.F | CTG GCA GCC GCT CAA GAC C | 19-mer | 58 | 68 |
| RD8-flank.R | GGC TCG ACC GCG TTC AGA C | 19-mer | 58 | 68 |
| RD8-Rv0309.int.R | GTA CTG CGG GAT TTG CAG GTT C | 22-mer | 57 | 55 |
| RD14-flankR | GGG CTG GTT AGT GTC GAT TC | 20-mer | 54 | 55 |
| RD14-flankF | TTG ATT CGC CAA CAA CTG AA | 20-mer | 48 | 40 |
| RD14-Rv1769.int.F | GTG GAG CAC CTT GAC CTG AT | 20-mer | 54 | 55 |
| RD14-Rv1769.int.R | CGT CGA ATA CGA GTC GAA CA | 20-mer | 52 | 50 |
| RDDenmark-flank.F | CTA CAA AAG CAA CCG GAC GCA GTG | 24-mer | 59 | 54 |
| RDDenmark-flank.R | GGC CAG CCG GAG AGC GAA ATC | 21-mer | 60 | 67 |
| RDFrappier-fl.F | GAT TCT CGA GCC GTC AAC CAA CC | 23-mer | 59 | 57 |
| RDFrappier-fl.R | CGA ACG CGC CGA ACA G | 19-mer | 51 | 69 |
| RDFrappier-Rv3495 | ACG ACC CGA CGG ATC CTT TTA T | 22-mer | 55 | 50 |
| RDRussia-flank.F | GGG CGT CGG GAA CCA GAT TG | 20-mer | 58 | 65 |
| RDRussia-flank.R | CGC GAC CGC CGA TGA CAC | 18-mer | 57 | 72 |
| recA3_start | CAT GCA TAT GAC GCA GAC CCC CGA TCG | 27-mer | 64 | 59 |
| recA3_stop | GCA CTA GTG CCC GCG CCT GCT CTT C | 25-mer | 66 | 68 |
| recAPstI_for | GAA CAG GCA CTC GAG ATC | 18-mer | 50 | 56 |
| recAPstI_rev | CTA AGA CCT GCT TCA TTC C | 19-mer | 55 | 47 |
| recA_965_983 | CCG ATC ACA AGG TGC TGA C | 19-mer | 60 | 58 |
| recA_599_618 | TCG GAG TGA TGT TCG GGT CG | 20-mer | 63 | 60 |
| recA_1307_1326 | CCG GTA TCT ACG GCA AGC TC | 20-mer | 63 | 60 |
| recA_1687_1705 | GCG CTT ATC CAG GCG ATT C | 19-mer | 60 | 58 |
| recA_1972_1990 | TCC GTG ATC CGA GAA GTG C | 19-mer | 60 | 58 |
| Rv3676_for | GGC CAG CCC AAT CCG CAG TTT AG | 23-mer | 61 | 61 |
| Rv3676_rev | CTG TGT CTG CTG GGC GTG AGG AAG | 24-mer | 63 | 63 |
| SP6[**] | ATT TAG GTG ACA CTA TAG | 18-mer | 41 | 33 |
| T3[**] | AAT TAA CCC TCA CTA AAG GG | 20-mer | 48 | 40 |
| T7[**] | GTA ATA CGA CTC ACT ATA GGG C | 22-mer | 53 | 45 |

[¶] Calculated according to Formula 7
[*] 16S rRNA primers (*117*)
[**] vector primers

## 2.12 Plasmids

### 2.12.1 Plasmids used in this study

**Table 5.** Plasmids used in this study

| Name, *Accession*[¶] | Features (selectable/counterselectable markers, size) | Source or Reference |
|---|---|---|
| pBluescript II KS(-), commonly pBS *X52329* | *bla* (ampicillin$_r$), 2961 bps | Stratagene, La Jolla, CA USA (*122*) |
| pMCS5 | *bla*, 3081 bps | MoBiTec, Göttingen, Germany (*123*) |
| pGEM®-7-Zf(+) *X65310* | *bla*, 2997 bps | Promega, Madison, WI, USA |
| pUC4k *X06404* | *bla*, *aph* (kanamycin$_r$), 3914 bps | Calgene, David, CA, USA |
| pBR322 *J01749* | *bla*, *tet* (tetracyclin$_r$), 4361 bps | (*124*) |
| pEJ126 | *bla*, *recA*, 17608 bps | (*61*) |
| pIJ903 | *bla*, *thio* (thiostreptone$_r$), *hyg* | (*125*) |
| pBS-*recA*, #819 | *bla*, *recA* (anticlockwise), 8189 bps | (*5*) |
| pBS-*recA*, #820 | *bla*, *recA* (anticlockwise, identical with #819) | (*5*) |
| pBS-*recA*::*aph*, #820 | *bla*, *recA*::*aph* (clockwise), 8174 bps | (*5*) |
| pBS-*recA*::3xstop | *bla*, *recA*::3xstop, 6956 bps | this study |
| pMCS5-*hyg* | *bla*, *hyg* (hygromycin$_r$), 4838 bps | this study |
| pMCS5-*hyg*(short) | *bla*, *hyg* (short *hyg* cassette from PCR), 4417 bps | this study |
| pMV361-*hyg* | *hyg*, 8694 bps | (*103*) |
| pGEM7-*hsp60-sacB* | *bla*, *sacB* (sucrose$_s$), 5494 bps | Sven Hobbie, (*126*) |
| pGEM7-*hsp60-sacB-aph* | *bla*, *aph*, *sacB*, 6734 bps | this study |
| pGEM7-*hsp60-sacB-aph-hyg* | *bla*, *aph*, *hyg*, *sacB*, 8066 bps | this study |
| pGEM7-*hsp60-sacB-aph-hyg*(short) | *bla*, *aph*, *hyg* (short), *sacB*, 7986 bps | this study |
| pGEM7-*recA*::3xstop-*hsp60-sacB-aph-hyg* | *bla*, *aph*, *hyg*, *sacB*, *recA*::3xtop, 12066 bps | this study |

[¶] GenBank/EMBL accession number (only where available)

### 2.12.2 Construction of the *M. bovis* BCG ΔrecA::3xstop suicide plasmid

The *recA* gene of *M. tuberculosis* was originally cloned and sequenced by Davis et al. in (*62*): A 13.2-kb *Eco*RI *recA* fragment was cloned in the *Eco*RI site of pBR322 and the resulting vectors were termed pEJ126 and pEJ127 (opposite orientation) respectively.

For the generation of suicide vector pGEM7-*hsp60-sacB-recA*::3xstop-*hyg-aph* the following cloning steps were performed: A 5.2-kb *Apa*I fragment from plasmid pEJ126 (*62*) containing *M. tuberculosis recA* was subcloned into the *Pst*I site of plasmid pBluescript-KSII(-) after blunting of both fragments with T4 polymerase, resulting in plasmid pBS-*recA* as described by Sander et al. (*5*).

From pBS-*recA* a 1.3-kb internal *Pst*I fragment (*recA* bps 504-1758 in protein residues 169-586) was replaced by a 22-mer triple translation stop insert with an internal *Mun*I restriction site and *Pst*I overhanging ends (Figure 17). This 22-mer was generated by annealing of dimeric oligonucleotide primers with 5'-phosphorylated ends. Briefly: two dimeric 5'-phosphorylated oligonucleotides were ordered from Microsynth (Belgach, Switzerland) in a concentration of 100 μM each. The oligonucleotides were diluted in annealing buffer (10 mM Tris-HCl, ph 8.0; 50 mM NaCl; 1 mM EDTA) to a concentration of 10 μM. The two oligonucleotides

were mixed at equal volumes (resulting in a final concentration of 5 μM per oligonucleotide). The mixture was heated to 96 °C for 5 minutes in a PCR machine and cooled down to 25 °C over a period of 8 hours. The annealing product was diluted 1:10 and used for ligation.

An *hsp60-sacB* construct in form of an *Eco*RV/*BamH*I fragment from plasmid pLO2 (*127*) was ligated to the *BamH*I/*Eco*RI site of pMV361 (*103*) to result in pMV361-*hsp60-sacB*, wherefrom an *Xba*I/*Xho*I fragment was ligated in the *Xba*I/*Sac*I site of pBluescript-IIKS(-), to result in pBS-*hsp60-sacB*. The later plasmid was digested with *Apa*I and *Nhe*I and a 2514-bp fragment ligated into pGEM7-Zf(+) digested with *Xba*I and *Apa*I, thereby excising 17 bps from the multiple cloning site to result in pGEM7-*hsp60-sacB* (5494 bps), expressing *sacB* under the control of the mycobacterial *hsp60* promoter. The plasmid pGEM7-*hsp60-sacB* (5494 bps) was obtained from Sven Hobbie (*126*).

A 1.3-kb *aph* cassette isolated as a *Pst*I fragment from plasmid pUC4K was ligated into the *Nsi*I restriction site of pGEM7-*hsp60-sacB*. Then we introduced a 1.3-kb *hyg* fragment (from pMCS5-*hyg* cut with *Bpu1102*I and *Van91*I, treated with T4 polymerase to create blunt ends) containing the hygromycin B resistance conferring hygromycin phosphokinase gene from *S. hygroscopicus*. Cutting pMCS5-*hyg* with *Van91*I removes the first 48 bps of the *hyg* gene promoter (*128*). The resulting plasmid pGEM7-*hsp60-sacB-hyg-aph* was cut with *Eco*RV and *Spe*I. A 4-kb *recA*::3xstop fragment from pBS-*recA*::3xstop (cut with *Eco*RV and *Spe*I) was introduced to result in suicide vector pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph* which was used for transformation of the BCG substrains mentioned above.

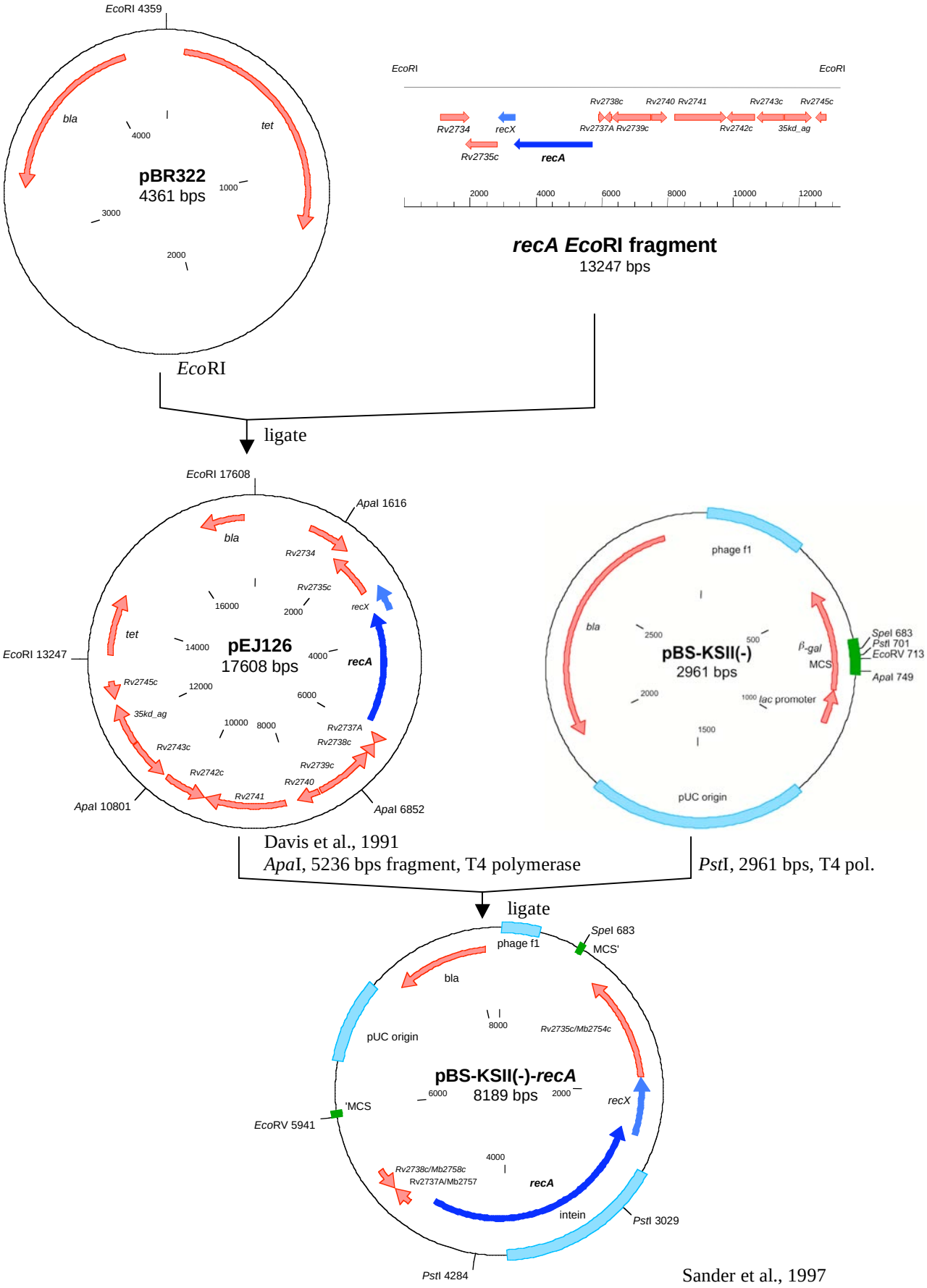



**Figure 14.** Cloning of plasmid pBluescript-*recA* as performed by Sander et al. (*5*)

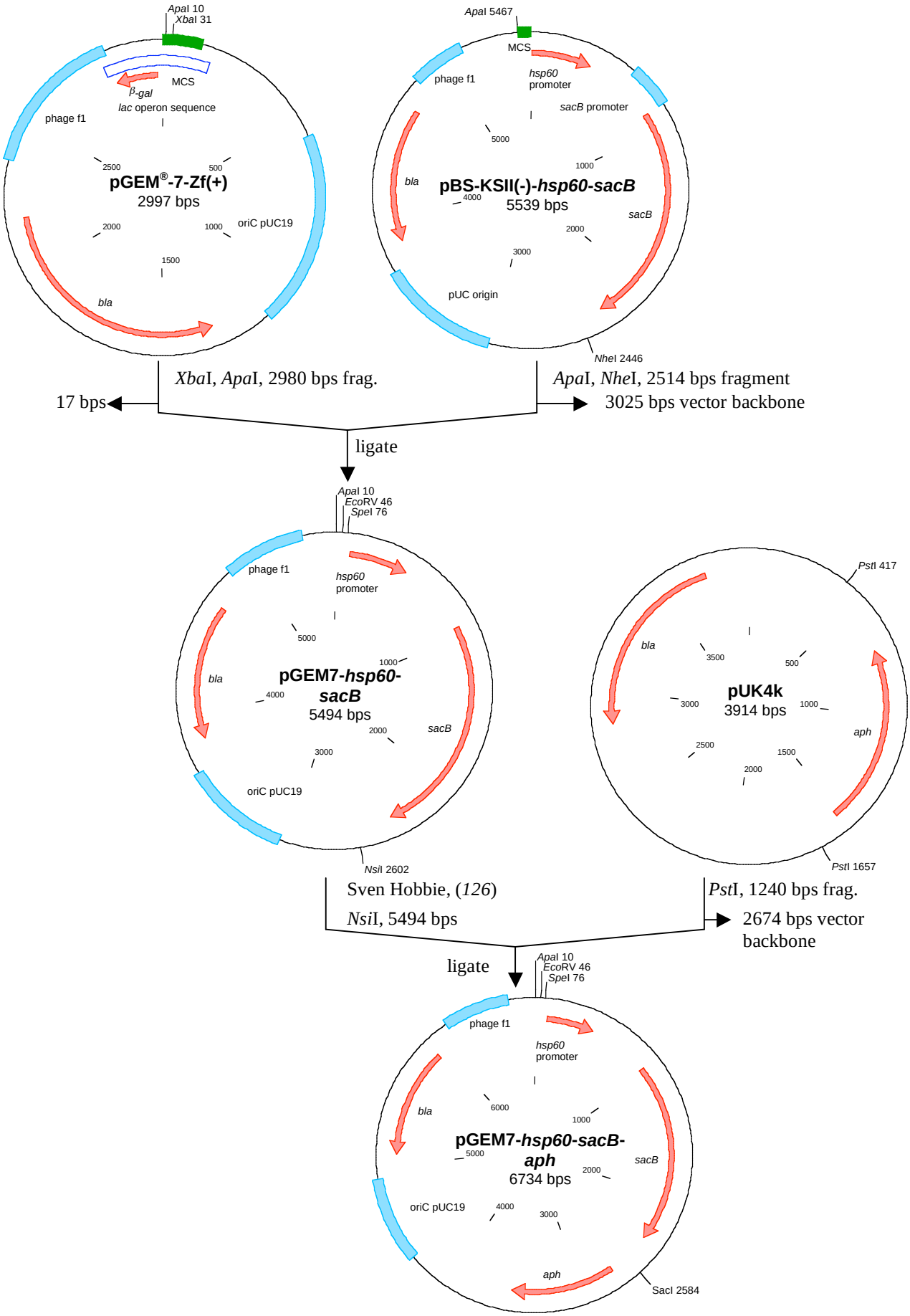



**Figure 15.** Cloning of plasmid pGEM7-*hsp60-sacB-aph*.

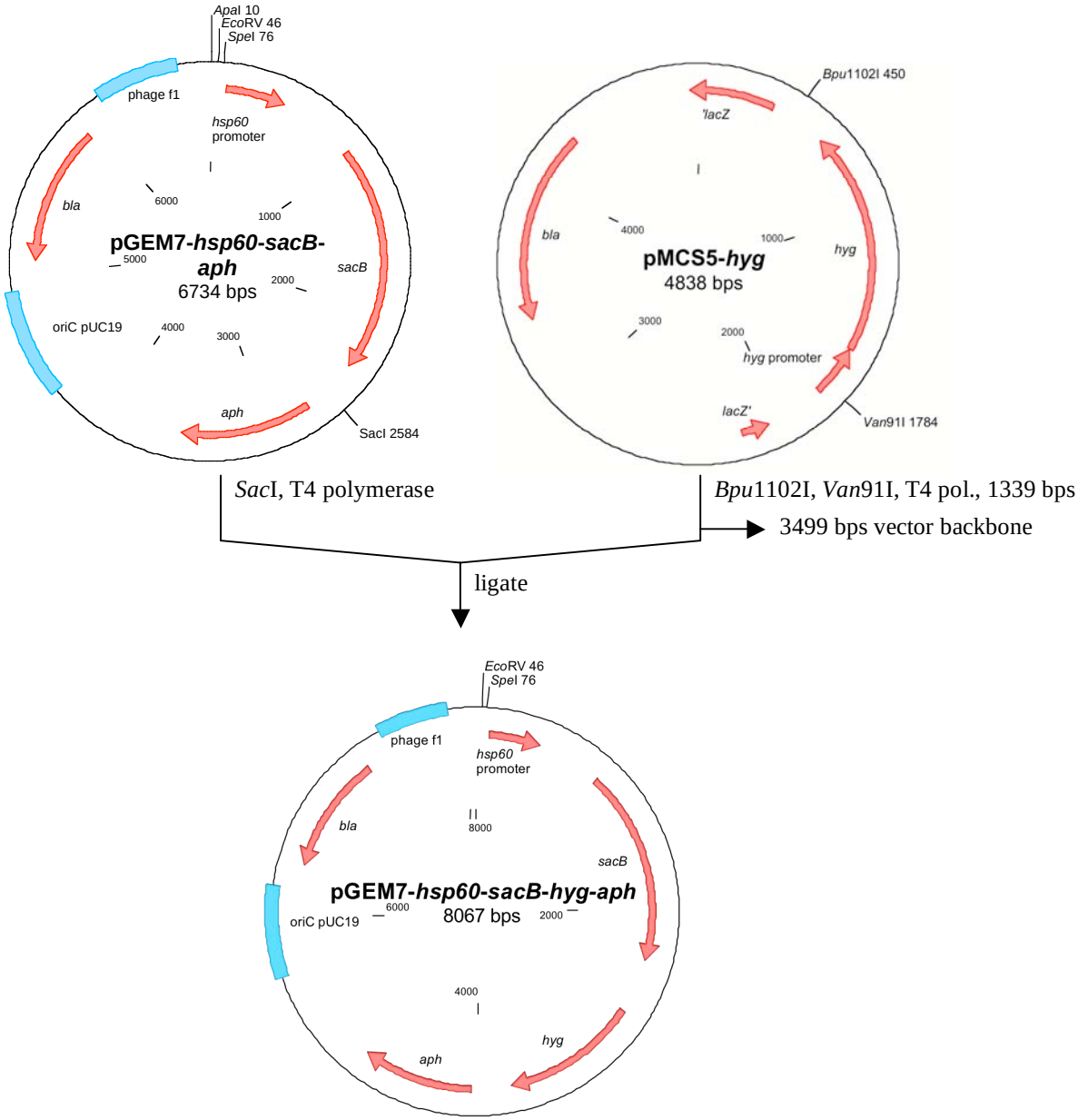



**Figure 16.** Cloning of plasmid pGEM7-*hsp60-sacB-hyg-aph* (empty suicide vector).

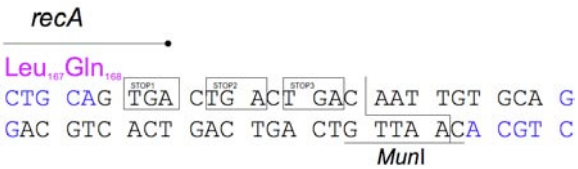



**Figure 17.** 22-bp triple translation stop insert with *Mun*I restriction site and *Pst*I overhanging ends. The synthetic insert contains a stop codon for every reading frame.

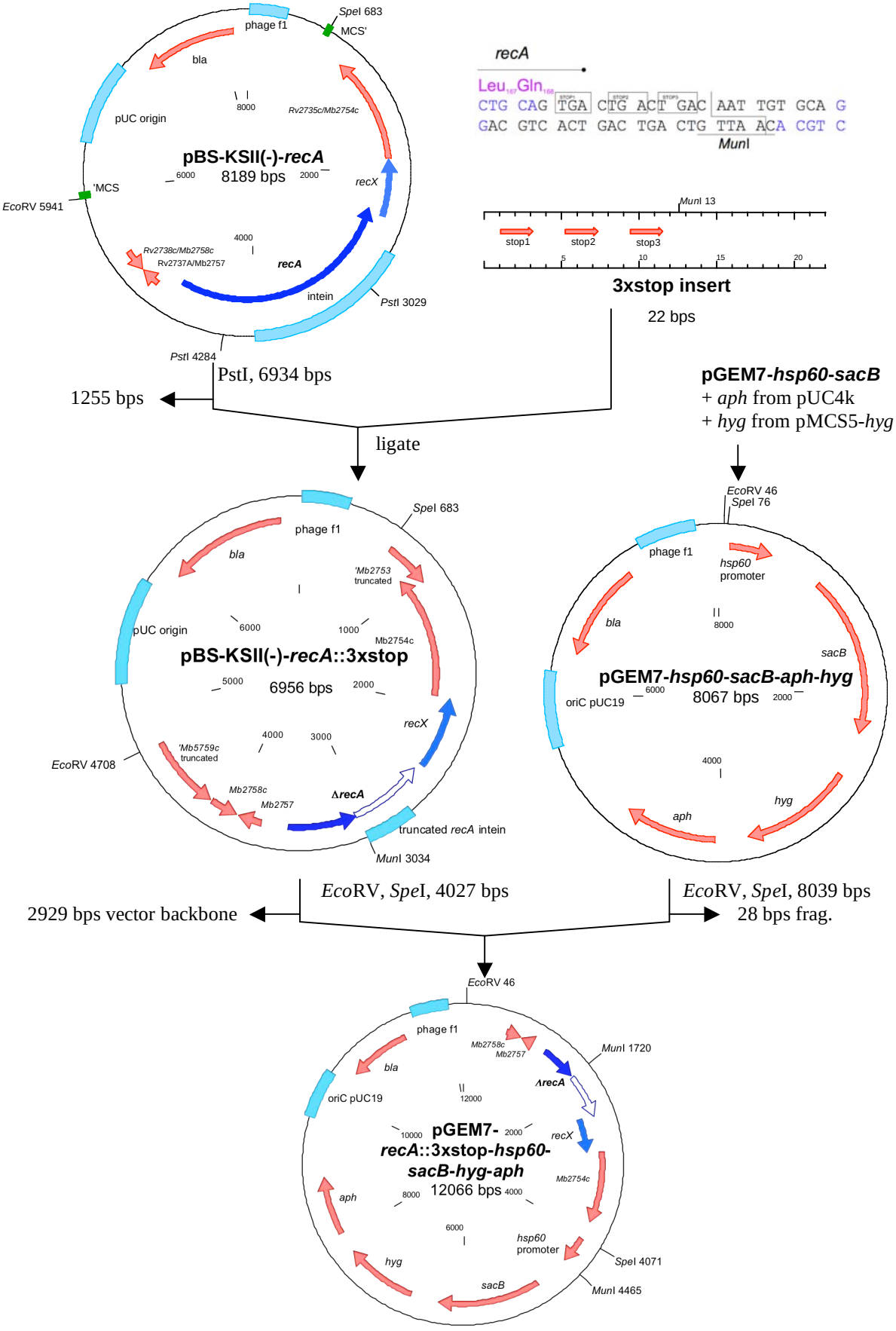



**Figure 18.** Cloning of suicide vector for knock-out generation pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*.

## 2.13 PCR for generation of short hyg cassette

The standard laboratory *hyg* gene cassette from pIJ903 (*125*) has a size of 1.8 kbs which is undesirably long for a suicide knock-out vector. I intended to shorten the hygromycin B resistance cassette to an essential minimum by PCR amplification of *hyg* with its promoter and ligation of the product into a standard high copy vector. *Hyg* gene and promoter sequence was obtained through the NCBI database (*129*), accession numbers: M18863 (*hyg* promoter), X99315 (*hyg* gene) and verified with original literature (*128*). Two flanking primers were ordered (hyg_forward and hyg_reverse, see Table 4).

The template has a GC-content of 77%, which is considerably high: Therefore I used a GC-RICH PCR System (Roche Diagnostics, Bale, Switzerland) for the amplification of the resistance cassette. The amplificate has a size of 1.3 kbs. The PCR product was cleaned by gel extraction and blunted with T4 polymerase. The blunt *hyg* cassette was inserted directly into pGEM7-*hsp60-sacB-aph* into the *Apa*I site. The resulting plasmid was sequenced with the following primers: P1, P2, P3, hyg_forward, hyg_reverse, aph5, and aph3.

## 2.14 Protein methods

### 2.14.1 Preparation of cell-free extracts

For the determination of RecA expression after DNA damage induction with mitomycin C, BCG cell-culture lysates were analysed by Western blot. Untreated and mitomycin-C-treated bacteria were harvested by centrifugation, washed with PBS buffer (pH 7.2) and resuspended in 300 µL PBS containing 0.5% protease inhibitor cocktail (Sigma). Bacteria were lysed in screw-cap tubes in a water immersion (ice-chilled) sonicator (Elma, Singen, Germany), using maximum duty-cycle during 1 h 20 min with a cooling intervall every 20 min. The crude lysate was clarified by spinning at 1 000 x g during 10 min.

### 2.14.2 Protein quantification

The total protein content of cell-free culture lysates was determined according to Bradford et al. (*130*). Quick Start Bradford Protein Assay (Bio-Rad) was used according to the instructions of the manufacturer. The Bradford dye assay is based on the equilibrium between three forms of Coomassie Blue G dye. Under acidic conditions, the dye is predominantly in the doubly protonated red cationic form ($A_{max}$= 470 nm). However, when the dye binds to protein, it is converted to a stable unprotonated blue form ($A_{max}$= 595 nm).

500 µL of 1:500 diluted cell-free lysates were mixed with 500 µL of a working reagent consisting of 1 vol Coomassie Brilliant blue G dye stock (100 mg Coomassie Blue G in 50 mL methanol, 100 mL 85% $H_3PO_4$, 50 mL $H_2O$) and 4 vol $H_2O$. Absorbance at 595 nm was measured and protein content of the samples was calculated by regression based on a standard curve obtained from a bovine serum albumine (BSA) standard dilution series.

### 2.14.3 SDS gel electrophoresis

Cell-free extracts corresponding to 50 µg total protein (quantification via Bradford assay) were used for SDS gel electrophoresis and subsequent Western blot analyses. Samples were mixed with one sample volume of sample buffer. A 13.3% polyacrylamide gel with a 4% stack gel was used to separate proteins. As a loading control, polyacrylamide gels were stained with Coomassie Blue G staining solution (Bio-Rad).

### 2.14.4 Western blot analysis

Following SDS-PAGE in a 13.3% polyacrylamide, gel proteins were blotted to a PVDF membrane using semi-dry blotter (Bio-Rad) at constant current (0.3 A) for 30 min. Equal loading of proteins was confirmed by Coomassie staining of an identical gel the efficiency of blotting was controlled by Coomassie staining of the blotted gel. Membranes were incubated with mouse-antibodies raised against mycobacterial RecA (provided by E. O. Davis, 1:1000), and by parallel incubation with an additional primary mouse-antibody raised against a constitutively expressed mycobacterial protein (KatG, provided by S. T. Cole, 1:2500). Secondary antibody was a goat-anti-mouse HRP-conjugated antibody (GE Healthcare) used in 1:2500 dilution. After incubation with the antibodies, membranes were developed in ECL solution and used to expose photographic films.

### 2.14.5 Buffers used for protein analyses

**Table 6.** Protein analysis buffers

| Descriptor | Composition |
|---|---|
| PBS | 220 mM NaCl, 150 mM $Na_2HPO_4$, 130 mM $KH_2PO_4$, pH 7.2 |
| SDS sample buffer | 62.5 mM Tris-base, 25% glycerol (vol/vol), 2% SDS (weight/vol), 0.01% (weight/vol) bromphenol blue, 5% β-mercaptoethanol (vol/vol) |
| Running buffer | 25 mM Tris-base, 0.1% SDS, 1.44% glycine, pH 8.3 |
| TBST | 50 mM Tris-base, 150 mM NaCl, adjusted to pH 7.5 with HCl, 0.05% Tween-20 |
| blocking solution | TBST, 4% milk powder (weight/vol) |
| ECL | Immun-Star™ HRP, Luminol/Enhancer, peroxide buffer (1:1), all Bio-Rad |

## 2.15 Bioinformatics

Bioinformatics were utilized for sequence analysis, alignments and similarity searches.

### 2.15.1 Primary databanks

1. The **GenBank** (*131*) in the USA is under the auspices of the National Centre for Biotechnology (NCBI) and is an official sequence data bank which contains protein and nucleotide sequences from more than 55,000 different organisms. All sequences are identified or tagged with a unique accession number. A GenBank sequence is usually divided into two parts: the annotation, which contains precise and detailed information about the sequence and the sequence itself. The ENTREZ search machine is coupled with the GenBank and allows a specific search based on an accession number, organism, gene, protein or author.
2. The **EMBL** nucleotide sequence database (*132*) is the European equivalent of the GenBank and utilizes the sequence retrieval system (SRS), a search machine similar to the ENTREZ for specialized searches of the database and many other databanks over the web interface. In addition to the text-based SRS and ENTREZ search engines described above, the basic logic alignment search tool (BLAST) was also extensively utilized. BLAST enables comparison of a particular sequence of interest with available databanks, leading to identification of similar sequences or relationships with previously described gene families.
3. **GenoList** (*118*), a database dedicated to the analysis of the genomes of the tubercle bacilli at the Pasteur Institute was used for sequence comparisons. Its purpose is to collate and inte-

grate various aspects of the genomic information from *M. tuberculosis* H37Rv (TubercuList), *M. africanum*, *M. bovis* AF2122/97 (BoviList), *M. bovis* BCG Pasteur 1173P2, *M. canetti*, *M. microti*, and other species. TubercuList provides a complete dataset of DNA and protein sequences derived from the reference strain *M. tuberculosis* H37Rv, linked to the relevant annotations and functional assignments. The data contained in TubercuList originates mainly from the *M. tuberculosis* H37Rv genome sequencing project (*133*), supplemented with data and observations made by researchers at the Unité de Génétique Moléculaire Bactérienne at the Pasteur Institute. The data contained in BoviList originates from the *M. bovis* AF2122/97 genome sequencing project (*120*). The database is searchable using BLAST and FASTA software.

4. *M. smegmatis* $mc^2$ 155 sequences were obtained through **TIGR** from the J. Craig Venter Insitute (Rockville, MD, USA) which is a non-profit center dedicated to deciphering and analyzing genomes (*134*).

The BLAST program (*135*), the TubercuList and BoviList databases were extensively used for sequence analysis and primer alignment studies. The BLASTN algorithm, which compares a nucleic acid query sequence with nucleic acid databanks directly, was used in this work.

## 2.15.2 Sequence analysis and plasmid maps

To align comparable DNA sequences, to identify point mutations DNA in files from automated DNA sequencing, and to perform sequence comparisons with published DNA sequences, sequences were assembled with SeqMan™ II DNA analysis software (DNASTAR, Madison, WI, USA).

Plasmid maps were drawn and annotated using Clone Manager, version 6.0 (Scientific & Educational Software, Cary, NC, USA).

# 3 Results

## 3.1 Generation of a *recA* targeting vector

A targeting vector for generation of unmarked *M. bovis* BCG *recA* mutants was cloned by subsequent modular assembly of the inactivated *recA* allele (*recA*::3xstop), positive (*aph*, *hyg*), and negative (*sacB*) selectable markers (Figs. 14, 15, 16, 17, 18)

### 3.1.1 Vectors pBS-*recA* and pBS-*recA*::*aph*

*M. tuberculosis* H37Rv *recA* gene and flanking DNA which have 100% sequence identity to *M. bovis* AF2122/97 segments was isolated from recombinant *E. coli* XL1-blue carrying pBS-KSII(-)-*recA* (#819, #820) and pBS-KSII(-)-*recA*::*aph* (#821). These vectors were cloning intermediates for generating marked *recA* deletion mutant strains (*5*). Identity of the recombinant plasmid was verified by phenotypic characterization of *E. coli* strains with respect to the resistance markers conferred by the vector backbone, restriction enzyme digestion and sequencing of plasmids. The plasmid containing *E. coli* strains were cultured on LB media with ampicillin and kanamycin which revealed ampicillin resistance for plasmids pBS-KSII(-)-*recA* (#819, 820) and for pBS-*recA*-*aph* (#821). For pBS-KSII(-)-*recA*::*aph* kanamycin resistance was noted. The plasmids were prepared on midi scale and 1 µg supercoiled DNA of vectors #819 and #821 was digested with restriction endonucleases *Eco*RV, *Pst*I, and *Spe*I and fragments were separated by agarose gel electrophoresis (Figure 19).

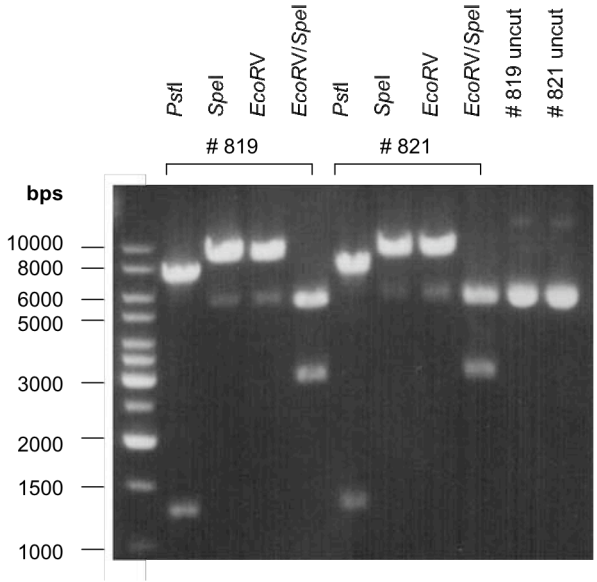



**Figure 19.** Gel electrophoresis (1% agarose gel, 5 $Vcm^{-1}$, 1 h 20 min, ethidium bromide stained) with restriction analysis of vectors pBS-*recA* (#819) and pBS-*recA*::*aph* (#821). The undigested supercoiled vectors run at 6 kbs. *Spe*I and *Eco*RV alone linearize both vectors, resulting in one fragment of > 8 kbs (besides a fragment of the same size as undigested vectors), both enzymes in a double digestion led to two fragments at approximately 6 kbs and 3 kbs for each vector. Digestion with *Pst*I results in two fragments at approximately 1.2 kbs and 7 kbs.

The design of the multiple cloning site (MCS) of the used skeleton vector pBluescript allows sequencing over the multiple cloning site with primers T3 and T7. The obtained sequences revealed that the *recA* insert had the same reading direction as the *bla* ampicillin resistance cassette in vectors pBS-*recA* (#819 and #820) and

reverse direction in pBS-*recA*::*aph* (#821); all constructs exhibited the published wild-type *recA* flanking sequences of *M. tuberculosis* H37Rv and *M. bovis* AF2122/97. Primers recA_PstI_for and recA_PstI_rev were designed to sequence a 1.2-kb internal *Pst*I fragment. In #819 and #820 this *Pst*I fragment presented with the same sequence as in the published *recA* of *M. tuberculosis* H37Rv and *M. bovis* AF2122/97, in #821 the *Pst*I fragment was replaced by an *aph* kanamycin resistance cassette originating from pUC4K. The *aph* gene had the same reading direction as the interrupted *recA* gene. To compute a complete vector map with sequence, the vector sequence of pBluescript-KSII(-) was obtained from GenBank (accession number X52329) and assembled with the insert sequence from direct sequencing, TubercuList queries on the H37Rv genome, and pUC4K *aph* sequence (GenBank accession number X06404). Vectors #819 and #820 have an identical sequence and the same size of 8189 bps (by computation analysis). #821 has a computed size of 8174 bps. Vector pBS-*recA* (#819, Figure 18) was used for further cloning in this study: After digestion of pBS-*recA* with *Pst*I and gel electrophoresis the 7-kb fragment was excised from gel and 5'-dephosphorylated. It was ligated to an artificial 22 bps oligonucleotide generated by annealing of two 5'-phosphrylated, complementary primers (Table 4, Figure 17). The oligonucleotide contains three stop codons, one in each frame on the complementary strand, a *Mun*I restriction recognition site and *Pst*I overhanging ends. After ligation of the fragments and transformation into DH5α a total of 19 CFU were obtained (low transformation efficiency in comparison to > 2500 CFUs after transformation with 1 ng pUC4K for the same batch of electrocompetent cells). Plasmids were harvested from 6 clones, DNA was digested with *Mun*I and *Spe*I and analyzed by agarose gel electrophoresis (Figure 20). Sequencing with primers recA_PstI_for and recA_PstI_rev was performed.

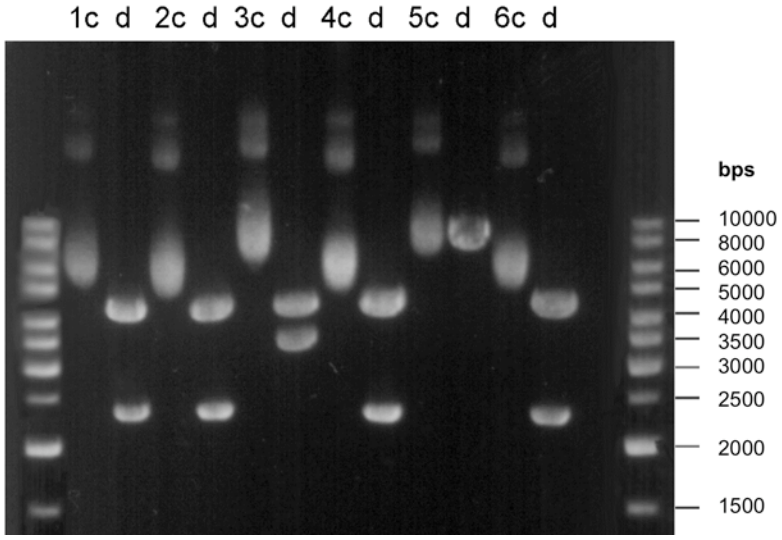



**Figure 20.** 1% agarose gel, ethidium bromide stained. Plasmids isolated from DH5α after transformation with the ligation product pBS-*rec*A::*Pst*I and triple translation stop insert. The insert contains a *Mun*I restriction site which is not present in the vector. Lanes marked with c contain undigested plasmid preparation whereas d marked lanes *Mun*I/*Spe*I digested plasmids. The numbers represent the clone numbers. Clones 1, 2, 4, and 6 show the expected restriction pattern with a 2.3 kbs and 4.6 kbs fragment.

Sequencing revealed the presence of one to four inserted oligonucleotides in the four clones (Figure 21). The orientation of the oligonucleotides was random, as expected. Clone 2 exhibited the desired sequence, i.e. a deleted *recA* allele and a single oligonucleotide with the correct orientation, i.e. the stop codon of the 3x transla-

tion stop insert is in frame with the reading frame of *recA* (Figure 22). The plasmid was named pBS-*recA*::3xstop (Figure 18).

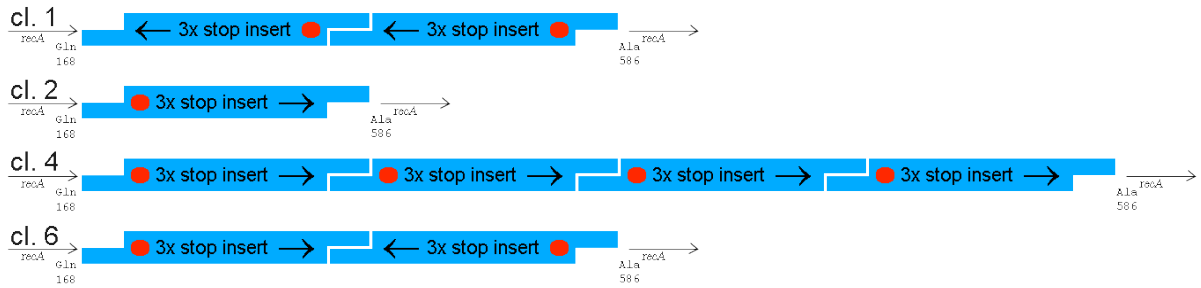



**Figure 21.** Sequencing results of clones 1, 2, 4, and 6 from the ligation experiment pBS-*rec*::*Pst*I + 3x translation stop insert. Clone 2 exhibits the desired sequence.

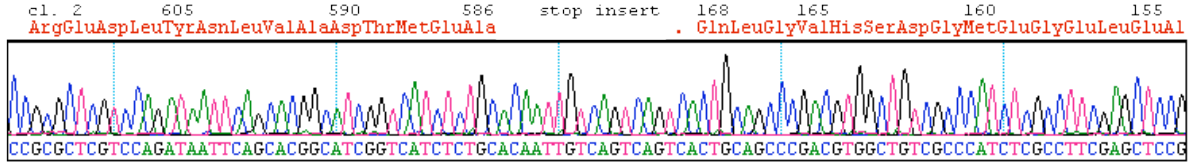



**Figure 22.** Sequence electropherogram of pBS-*recA*::3xstop, cl. 2 sequenced with primer recA_PstI_rev with amino acid translation (numbers counting from $Met_1$, *recA* initiation, reading direction is reverse, corresponding to genomic orientation of *recA*).

### 3.1.2 Vector pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*

*E. coli* XL1-blue were transformed with 2 ng of pGEM7-*hsp60-sacB* from frozen stock plasmid DNA solution, a midi scale plasmid preparation was performed, and the vector was characterised by restriction endonuclease analyses and sequencing. The vector was initially generated by Sven Hobbie (*126*). The MCS of this vector carries the mycobacterial *groEL2* (*hsp60*) promoter from *M. bovis* and further downstream the counterselectable marker gene *sacB* from *B. subtilis*. A vector map including sequence was constructed by sequencing of essential parts of the vector with T7 and SP6 primer and computational assembling with the vector pGEM®-7-Zf(+) sequence and annotations from GenBank (accession number X65310, Figure 15). Comparison of the sequences of the *groEL2* promoter (accession number BX248335) and *sacB* gene (accession number Z99121) from sequencing to the published sequences in GenBank by BLASTN software confirmed identity with the sequence deposited in the database.

A *Pst*I excised kanamycin resistance cassette (*aph*; 1.2 kbs) from pUC4K was introduced downstream of *sacB* into the *Nsi*I restriction site of vector pGEM7-*hsp60-sacB* to result in plasmid pGEM7-*hsp60-sacB-aph* (Figure 23A). The resistance cassette ideally should be integrated in the same reading direction as *sacB* to avoid interference of transcription as indicated in plasmid map pGEM7-*hsp60-sacB-aph*(+). Restriction analysis of the new construct with *Sma*I showed, that *aph* integrated stochastically in both directions as expected. Sequencing of three clones with primers aph3', aph5', and SP6 followed, confirmed the outlined restriction maps, and showed absence of mutations in the *aph* gene. Plasmid pGEM7-*hsp60-sacB-aph*(+) was used for further subcloning.

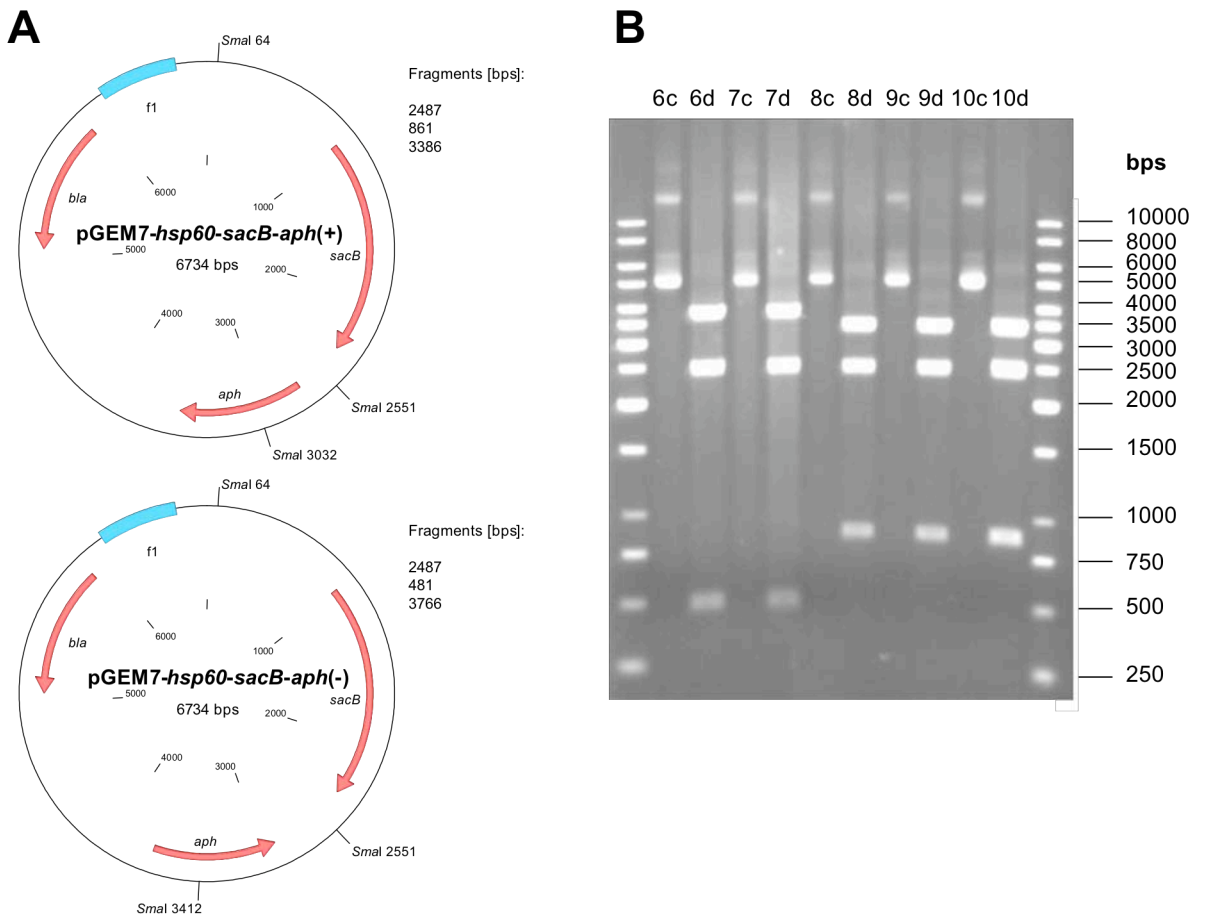



**Figure 23.** Restriction maps **(A)** and restriction digestion with *Sma*I **(B)** of pGEM7-*hsp60-sacB-aph*(+) and (-). After transformation plasmids from 10 clones were prepared and digested with *Sma*I (d, digested) and compared to undigested vector (c, control). The *aph* gene in clones 6 and 7 has the same orientation as *sacB*; in clones 8 to 10 reverse orientation can be noted.

To generate pGEM7-*hsp60-hyg-aph*, a hygromycin B resistance conferring cassette (1.3 kbs) was prepared from vector pMCS5-*hyg* with *Bgl*1102 and *Van*91I and subsequent blunting with T4 polymerase. To cut pMCS5-*hyg* with *Van*91I the vector had to be transformed in the *dcm*$^-$ strain DH5α Δ*dcm* in order to allow normal function of the *Van*91I restriction endonuclease since this enzyme is Dcm-methylation sensitive. The vector pMCS5-*hyg* was generated by cloning of a standard *Streptomyces hygroscopicus hyg* cassette from pMV361-*hyg* into the multiple cloning site of pMCS5 (Martin Bosshard, unpublished). A vector map (Figure 17) was established using the published pMCS5 sequence (*123*) and direct sequencing of the insert region of a plasmid preparation using T7, SP6, hyg_start, and hyg_stop primers. The sequencing showed a single base pair insertional mutation in the *hyg* promoter with respect to the published sequence of the *hyg* promoter (GenBank accession number M18863), which was present in the other laboratory held *hyg* vectors too (in originating pMV361-*hyg* as well as pBS-*hyg* derivate). To investigate normal function and *hyg* expression in *E. coli* a minimal inhibitory concentration (MIC) assay of four clones of DH5α cells containing pMCS5-*hyg* was performed: In all clones an MIC of 125 µg mL$^{-1}$ hygromycin B was determined, confirming a *hyg* resistant phenotype (resistance level of non-transformant strain was 31.25 µg mL$^{-1}$).

In a subsequent cloning experiment vector pGEM7-*hsp60-sacB-aph*(+) was linearized with *Sac*I, blunted with T4 polymerase, 5'-dephsophroylated, and a *Bgl*1102I/*Van*91I 1.3-kb *hyg* fragment originating from pMCS5-*hyg* was ligated into the vector to result in pGEM7-*hsp60-sacB-hyg-aph*. After transformation and selec-

tion on LB-Hyg plates the orientation of the *hyg* gene in the newly generated pGEM7-*hsp60-sacB-hyg-aph* was uncovered by restriction digestion with *Sac*I and *Spe*I. The *hyg* gene had the same orientation as the *sacB* gene in 13 out of 16 clones analyzed, in 1 clone reverse orientation was noted, the remaining 2 clones showed inconclusive restriction patterns. A computed vector map (Figure 25A) was confirmed by direct sequencing of clone 1 (number as in Figure 24) with aph5' and P3 primer. This plasmid was used for further subcloning.

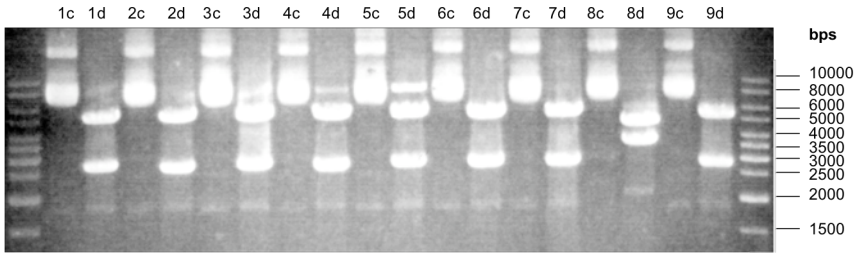



**Figure 24.** 1% agarose gel, ethidium bromide stained. Result of the ligation of pGEM7-*hsp60-sacB-aph* and *hyg* gene from pMCS5-hyg. The isolated plasmid (c, control, undigested) was analyzed in restriction digestion with *Sac*I and *Spe*I (d, digested) resulting in fragments of 2811 bps and 5255 bps for clockwise and 3545 bps and 4521 bps for anticlockwise orientation of the *hyg* insert.

Based on pGEM7-*hsp60-sacB-hyg-aph* (Figure 25, panel A) the final suicide vector for the knock-out experiment was cloned: A 4027-bp fragment was prepared from pBS-*recA*::3xstop using *Eco*RV and *Spe*I. This insert was ligated to pGEM7-*hsp60-sacB-hyg-aph*, which was digested with the same enzymes and 5'-dephosphorylated to result in pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph* (Figure 25, panel B). 22 clones were analyzed in restriction digestion with *Eco*RV and *Spe*I. The obtained fragments corresponded to single insert integration in 20 clones. Two plasmids (cl. 7 and 8) out of those were sequenced with primers P1, P2, P3, SP6, recA_PstI_for, recA_PstI_rev, aph5', and aph3'. They had the calculated sequence. The MIC of hygromycin B of recombinant *E. coli* DH5α containing the vector pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph* (Figure 25, panel B), cl. 7 to be used for the knock-out experiment (see 3.3) was 250 μg $mL^{-1}$ versus 31.25 μg $mL^{-1}$ for the non-transformant parental strain (the experiment was performed in triplicate). Plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph* has all the features required for generating unmarked *recA* deletion mutants.

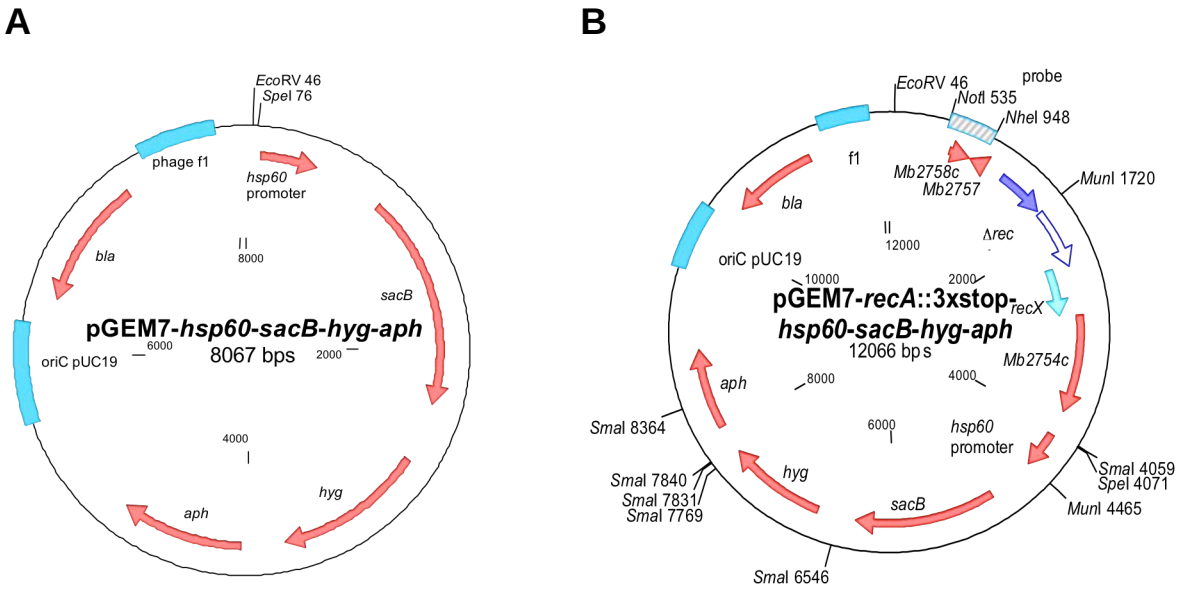



**Figure 25. (A)** Empty suicide vector pGEM7-*hsp60-sacB-hyg-aph* for the generation of unmarked mutants in mycobacteria and **(B)** suicide plasmid for the generation of unmarked *recA* mutants.

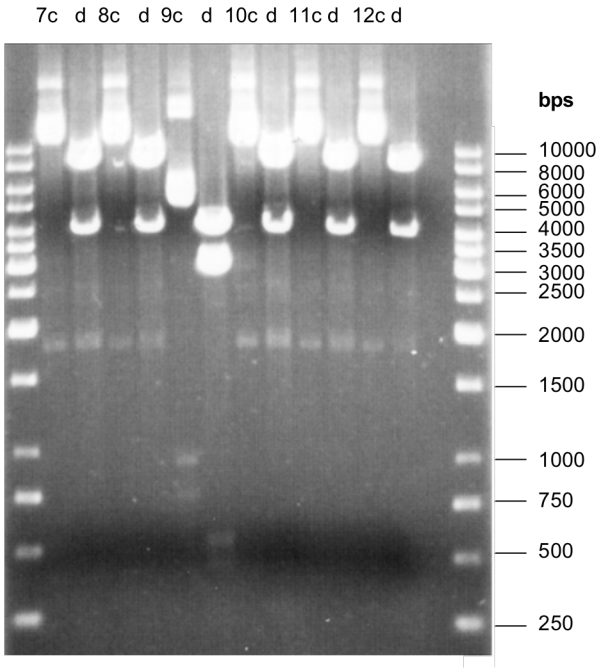



**Figure 26.** 1% agarose gel, ethidium bromide stained. Ligation of pGEM7-*hsp60-sacB-hyg-aph* with 4 kb *recA*::3xstop fragment from pBS-*recA*::3xstop. Control (c) versus digested (d) plasmids (*Eco*RV/*Spe*I), clones 7 to 12. Clone 7 was used for transformation of BCG vaccine strains.

## 3.2 Choice and characterisation of *Mycobacterium bovis* BCG vaccine substrains

### 3.2.1 Choice of *M. bovis* BCG substrains

M. *bovis* BCG was propagated in several laboratories around the world. This resulted in a still ongoing *in vitro* evolution of different substrains. Alterations are mainly due to deletion and duplication of defined regions (*41*). Single nucleotide polymorphisms (i.e. in *crp*, and *sigK*) have also been described (*107, 116*). The various genetic alterations – some presumably depending on RecA-mediated recombination - affect the antigenic, protective, and metabolic properties of BCG. The selection of BCG substrains based on criteria defined in section 1.3: History, passage number, morphological description and date of lyophilization could be obtained for all BCG used in this study. The selection comprises a representative for each branch of the phylogenetic tree displaying BCG *in vitro* evolution (Figure 27).

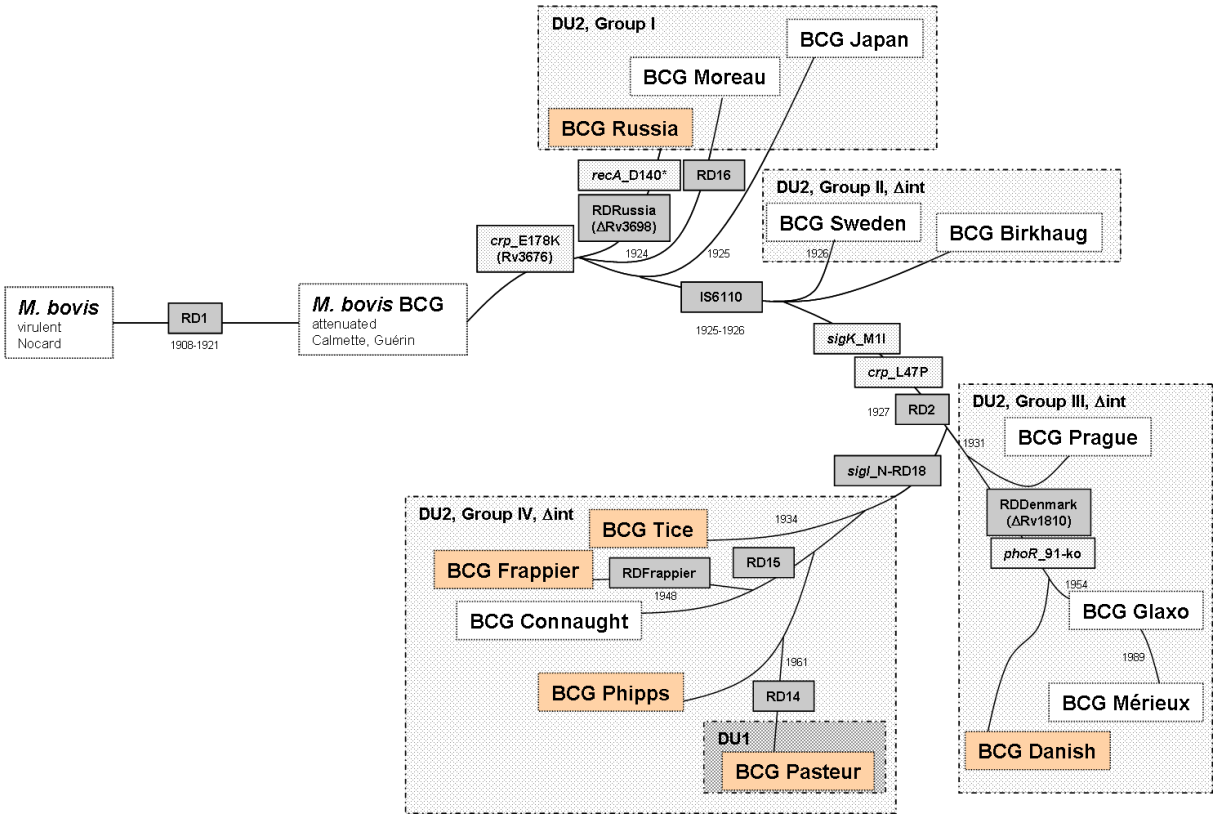



**Figure 27.** Genealogy of BCG vaccine substrains, modified from (*41*) with permission, copyright (2007) National Academy of Sciences, U.S.A., displaying the original virulent ancestor strain *Mycobacterium bovis* (isolated by Nocard in 1908) and the subsequent series of genomic alteration including deletions of regions of difference (RD), single nucleotide polymorphisms, and duplications of genomic regions. The *recA* alteration of BCG Russia described in section 3.4 (*recA*_D140*) is added to the scheme. Substrains that were used in this study are shaded in orange.

### 3.2.2 Morphology and growth

The BCG substrains all arrived in freeze-dried state in intact glass ampoules, which were welded in the case of BCG Frappier, Phipps, Pasteur, and Russia and closed with a rubber seal in the case of BCG Tice and Denmark. The number of CFU per vial was indicated only for BCG Tice (5 x $10^8$). The visible amount of lyophilizate varied greatly: BCG Denmark and Russia ampoules contained very small amounts whereas the amount in the BCG Tice ampoule appeared the most. First of all the specific appearance and molecular features of the six substrains were confirmed.

#### 3.2.2.1 BCG Frappier (ATCC #35735)

From a freeze-dried lot, BCG Frappier was cultured in Middlebrook 7H9 medium in two culture flasks with 10 fold different innocula. In both culture flasks turbidity of media was visible starting on the third culture day. On day seven 100 µL of the culture were streaked on solid 7H10 agar and incubated for 21 days. On the same occasion samples were preserved for staining with Ziehl-Neelsen and Auramine-Rhodamine. Microscopic inspection of the Ziehl-Neelsen preparation showed acid-fast rods in aggregates and single occurrence without contaminants. In Auramine-Rhodamine red fluorescent bacilli of approx 3.5 x 0.5 µm were notable (Figure 29).

After the incubation period uniform culture morphology was noted and on the 21$^{st}$ day of culture a single typical colony was inoculated in 7H9 again. After three weeks of culture 100 µL of liquid culture were transferred to a 7H10 plate, incubated for 21 days, and then photographed (Figure 28). Single colonies were white-

yellowish with marked difference in size between 3 and 9 mm and confluent. The colonies appeared rough with high tendency to the formation fine wrinkling (rugosity) of 1 to 2 mm in height.

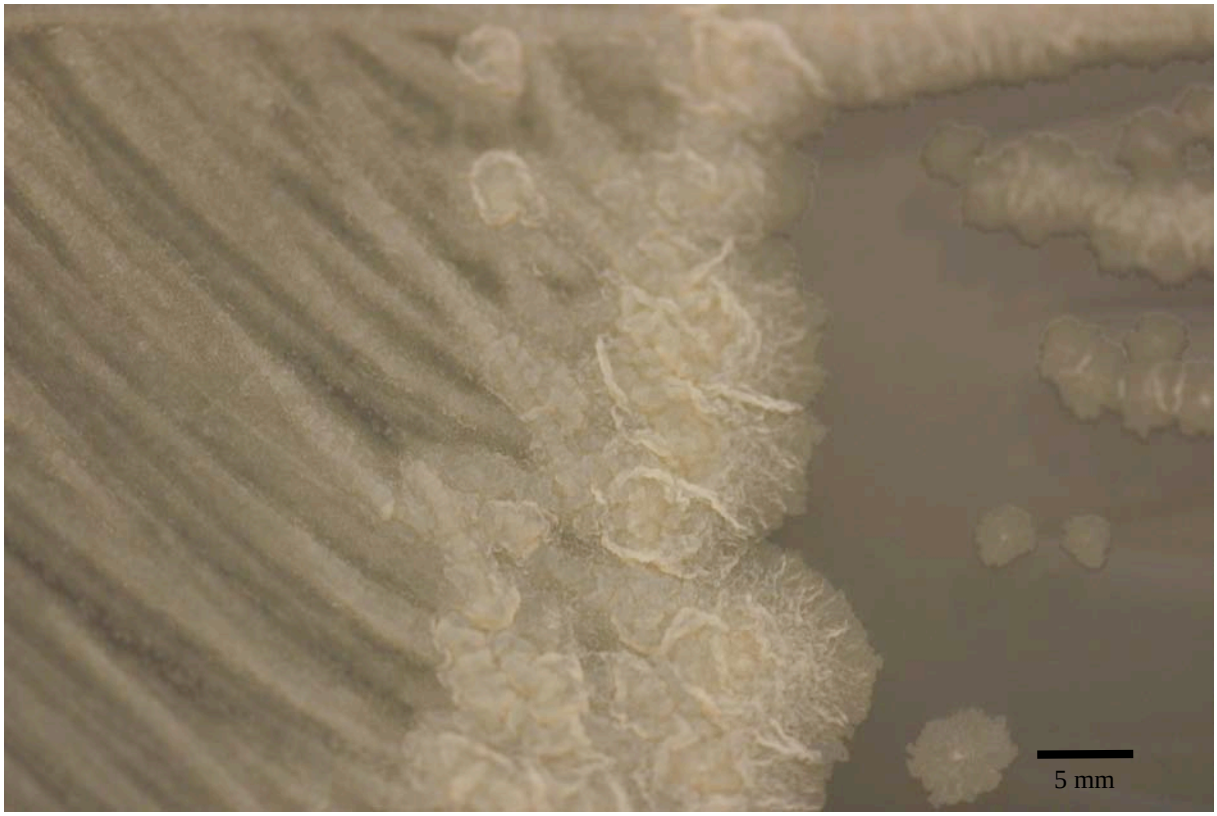



**Figure 28.** BCG Frappier on 7H10 after 21 days of culture (black bar equals 5 mm).

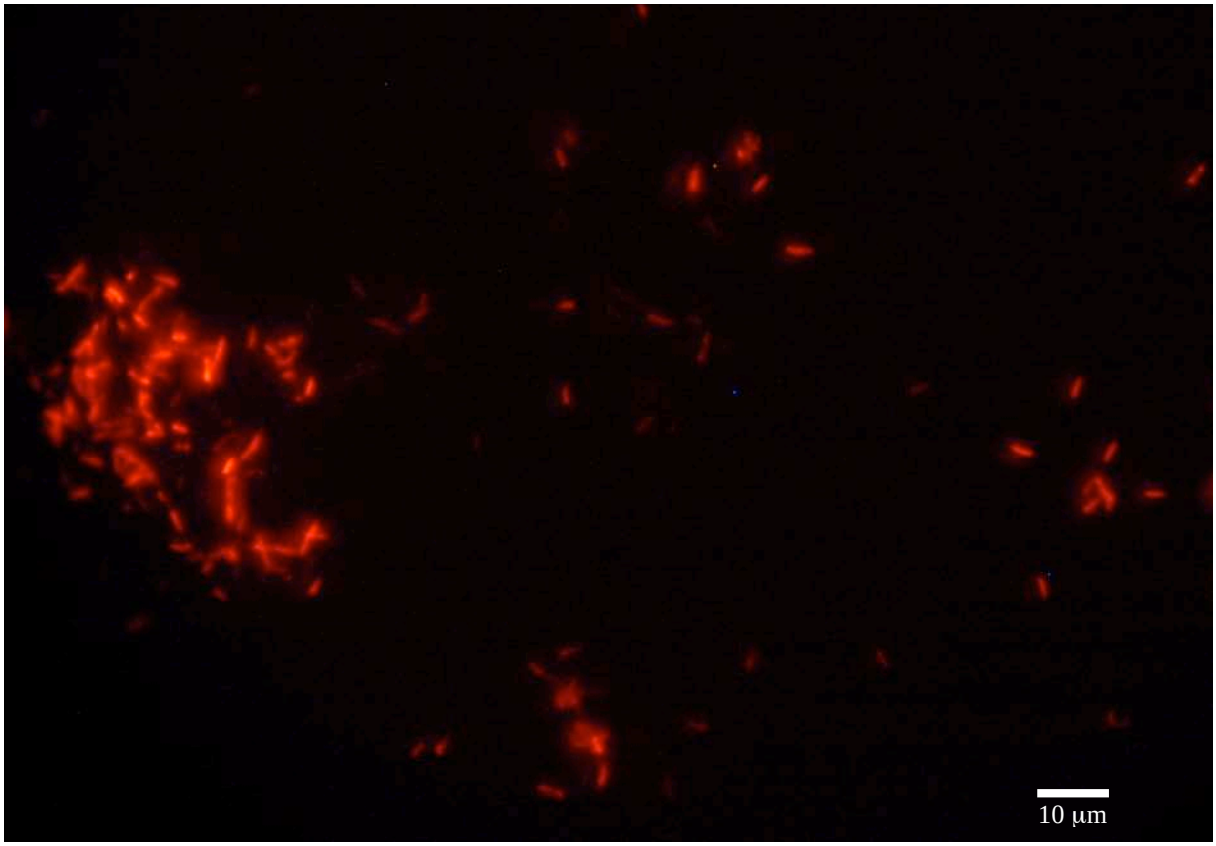



**Figure 29.** Auramine-Rhodamine stained direct preparations from 7H9 culture of BCG Frappier (white line corresponds to 10 μm).

### 3.2.2.2 BCG Phipps (ATCC #35744)

For this strain the same procedure was applied: Cells were revitalised in 7H9, microscopic examination was performed (with the same results), cells were cultured on solid medium, a single colony was transferred to 7H9 and the liquid culture was transferred to 7H10 for macroscopic culture examination. Rugosity could be observed in macroscopic culture morphology (Figure 30), even though in diminuend extent compared to BCG Frappier (Figure 28).

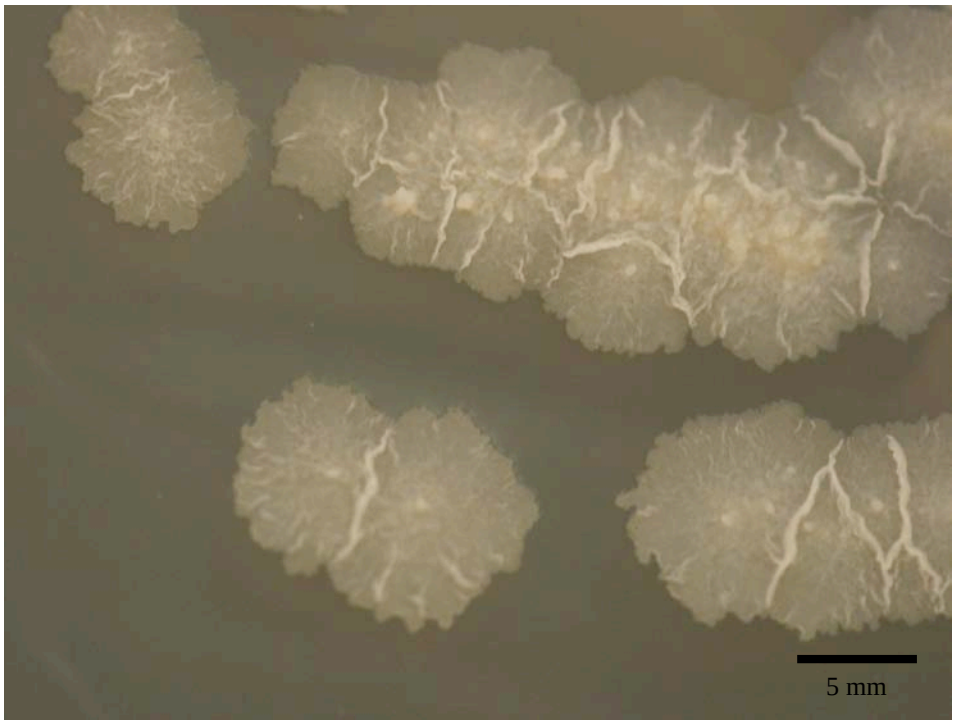



**Figure 30.** Culture of BCG Phipps on 7H10 after 21 days of incubation.

### 3.2.2.3 BCG Tice

BCG Tice was cultured and examined as the other substrains mentioned previously. According to acid-fast staining the culture was pure. Macroscopic culture evaluation again showed less rugosity than in BCG Frappier.

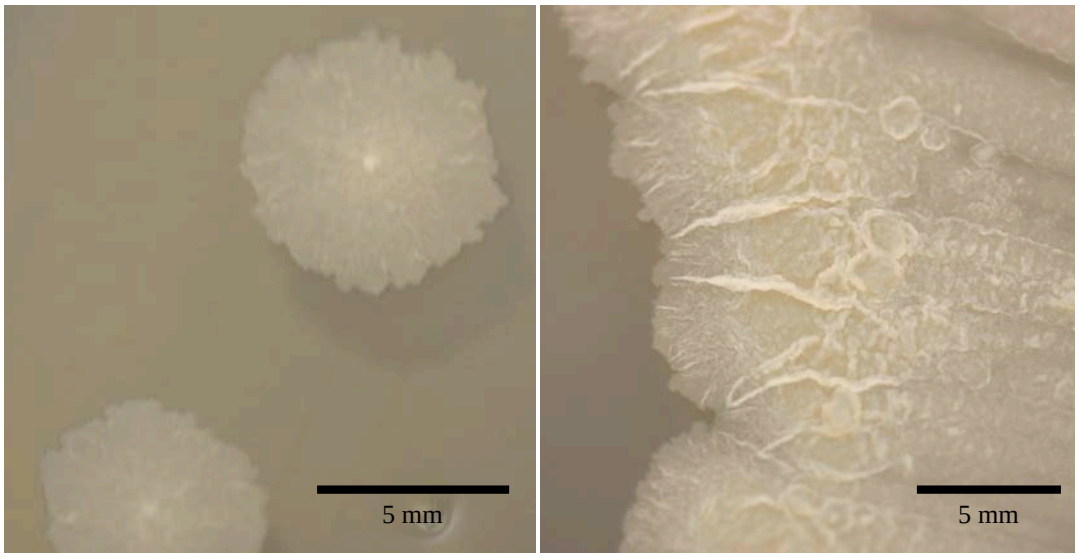



**Figure 31.** 7H10 culture of BCG Tice, photographed on the 21st day of incubation.

### 3.2.2.4 BCG Pasteur

BCG Pasteur was cultured and examined in the same way as the previous substrains. Microscopic examination was indifferent but macroscopic morphology showed a marked difference: There were no line-like rugous elevations but spike-like structures.

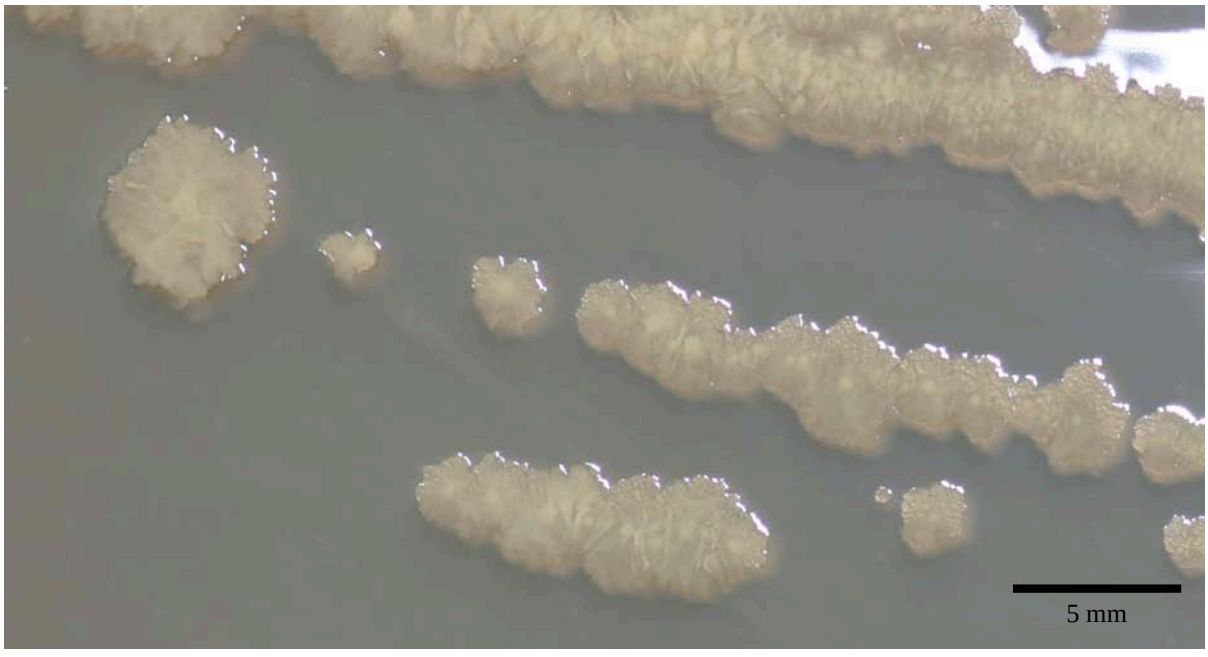



**Figure 32.** 7H10 with culture of BCG Pasteur on 21st day of incubation.

### 3.2.2.5 BCG Denmark 1331

The cultivation and documentation of BCG Denmark was handled identically as for the previous strains. BCG Denmark exhibited strong formation of rugous structures. Single colonies appeared drier than colonies of the other substrains and were harder to pick with a loop.

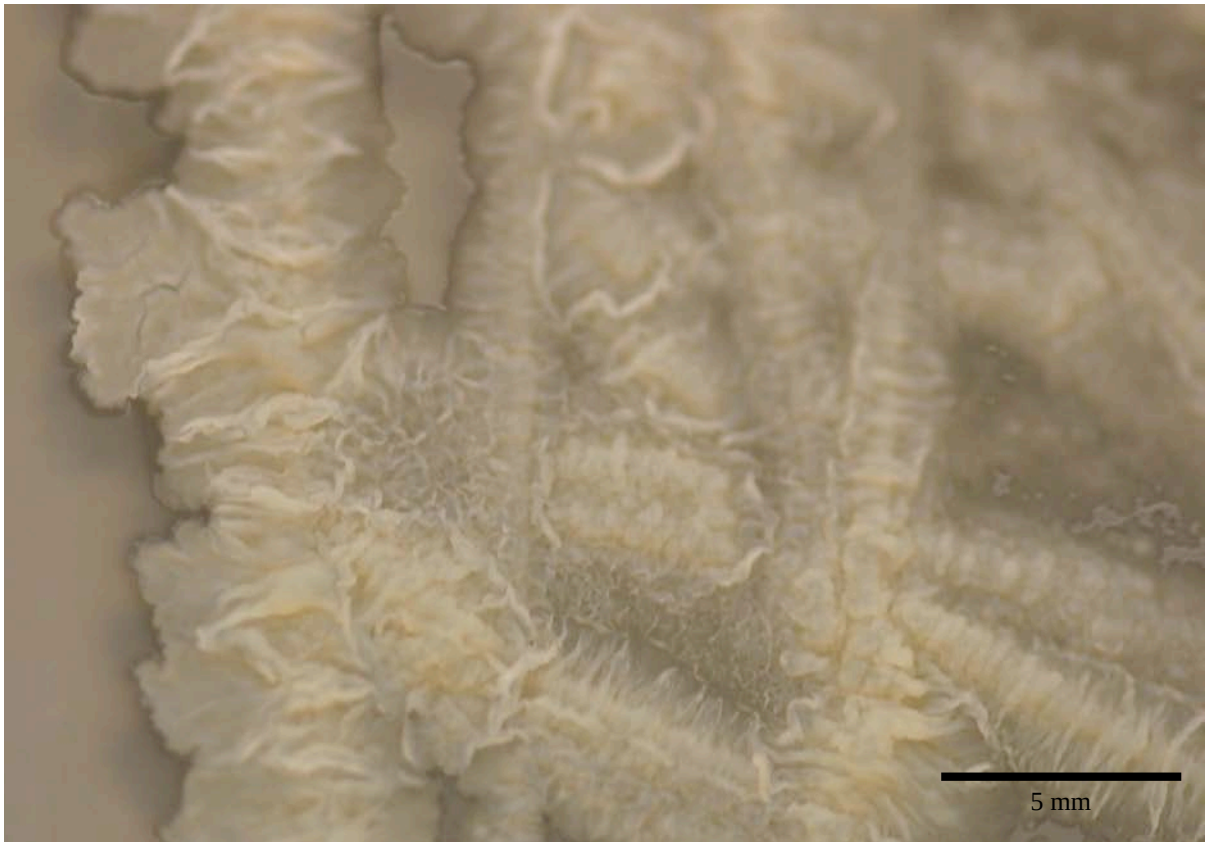



**Figure 33.** BCG Denmark on 7H10 on 21st day of incubation.

### 3.2.2.6 BCG Russia

A remarkable growth lag in comparison to the other strains was noted for the initial culture from a lyophilized batch in a glass ampoule, there was no visible growth until the 7th day of incubation. Even though the medium was inoculated with approximately 20 mg of dried bacteria visible growth was only observed at day 29. 16S-RNA gene PCR and subsequent sequencing confirmed that the medium contained bacteria of the *M. tuberculosis* complex. The sequence of a 1040 bps amplificate, which was generated by application of a protocol described by Kirschner et al. (*136*) using primers #285 and #264 showed 100% sequence identity with an amplificate from BCG Pasteur. An additional 7H9 culture flask was inoculated with the content of another fresh ampoule of BCG Russia. The second batch proved more viable bacteria exhibiting faster growth with visible turbidity on the third day of incubation. Again microscopic examination was indifferent to the previously mentioned substrains. After 21 days of incubation on solid agar the colonies produced relative high elevations (> 2 mm) with undulating forms.

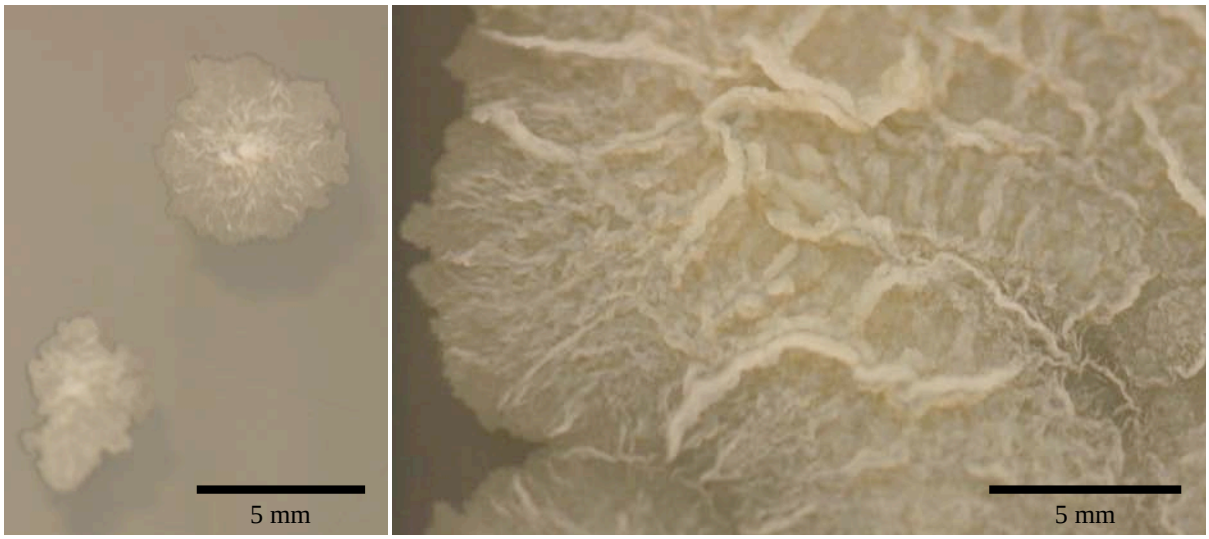



**Figure 34.** 7H10 culture of BCG Russia after 21 days of incubation.

## 3.2.3 Molecular identification of BCG substrains

### 3.2.3.1 Regions of difference

DNA/DNA hybridization techniques, PCR analyses and sequencing have identified several regions of difference (RD) between *M. tuberculosis* and *M. bovis* BCG substrains (*41, 43, 91, 107, 111, 113-115, 137-139*). Some regions of difference are characteristic for all BCG strains (RD1), while others allow the identification of specific substrains. The regions of difference (RD) indicated in Table 2 were analyzed in specific PCR reactions. DNA samples were prepared from the first 7H9 culture of every substrain. *M. tuberculosis* H37Rv served as a positive control and its annotated sequence (*118*) served as a reference.

**Table 7.** Comparative results of RD specification of the analyzed BCG substrains confirming results of Mostowy et al. (*115*).

| BCG substrain | Deleted regions | Number of RDs | ORFs affected | Base pairs deleted |
|---|---|---|---|---|
| *M. tuberculosis* H37Rv | none | 0 | 0 | 0 |
| Frappier ATCC #35735 | RD1, RD2, nRD18, RD8, RDFrappier | 5 | 30 | 30866 |
| Phipps ATCC # 35744 | RD1, RD2, nRD18 | 3 | 23 | 21793 |
| Tice | RD1, RD2, nRD18 | 3 | 23 | 21793 |
| Pasteur 1173B-P2 | RD1, RD2, nRD18, RD14 | 4 | 31 | 30866 |
| Denmark 1331 | RD1, RD2, RDDenmark/Glaxo | 3 | 22 | 20972 |
| Russia | RD1, RDRussia | 2 | 11 | 11061 |

#### *3.2.3.1.1 RD8*

RD8 is a signature sequence of BCG substrain Frappier comprising of a deletion of 3428 bp. Behr et al. describe that BCG Frappier has a deleted RD8 (*43*). They examined three isolates from the Institut Armand-Frappier (IAF), Laval, Canada, without documenting the lot that is preserved at the ATCC. The three reference vials from the former bacterial collection of the IAF were not available for this study. DNA from BCG substrains and *M.*

*tuberculosis* H37Rv served as a template. A multiplex PCR with primers RD8-flank.F, RD8-flank.R, and RD8-Rv0309.int.R generates either of two fragments of similiar size, i.e. 286 bps and 292 bps. The 286-bp fragment results from template amplification with primers RD8-flank.F and RD8-Rv03909.int.R, whereas the 292-bp fragment results from template amplification with primers RD8-flank.F and RD8-flank.R. These fragments cannot be differentiated by agarose gel electrophoresis but may be sequenced.

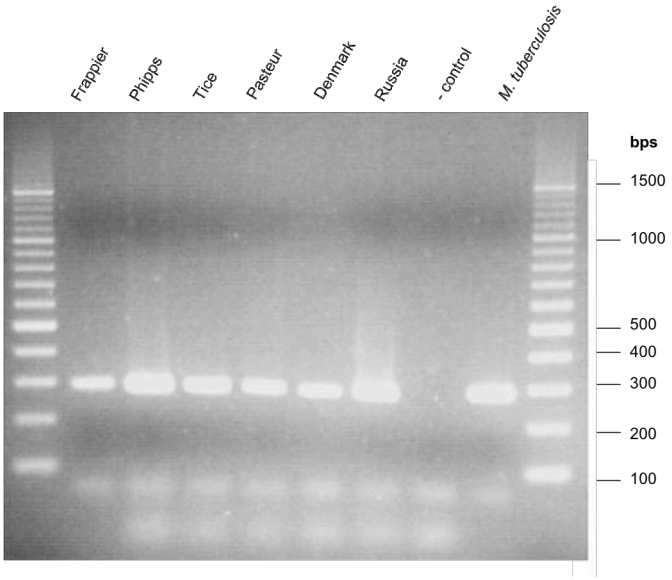



**Figure 35.** 2% agarose gel, ethidium bromide stained. PCR amplificates of multiprimer PCR over RD8. Amplificate length of 286 bps corresponds to present RD8 sequence. For an absent RD8 a 292-bp amplificate was expected.

Fragments of a size of approximately 300 bps were obtained with both *M. tuberculosis* and *M. bovis* BCG (Figure 35). Sequencing using the RD8-flank.F primer in order to confirm the absence of *M. tuberculosis* homologue sequence at RD8 locus in BCG Frappier was necessary because a difference of 6 bps is not visible on a 2% agarose gel. Sequencing proofed that BCG Frappier ATCC #35735 has a deletion of 3428 bps at the RD8 locus (Figure 36, lower panel); sequencing of the other products revealed the presence of the corresponding sequence (Figure 36 upper panel).

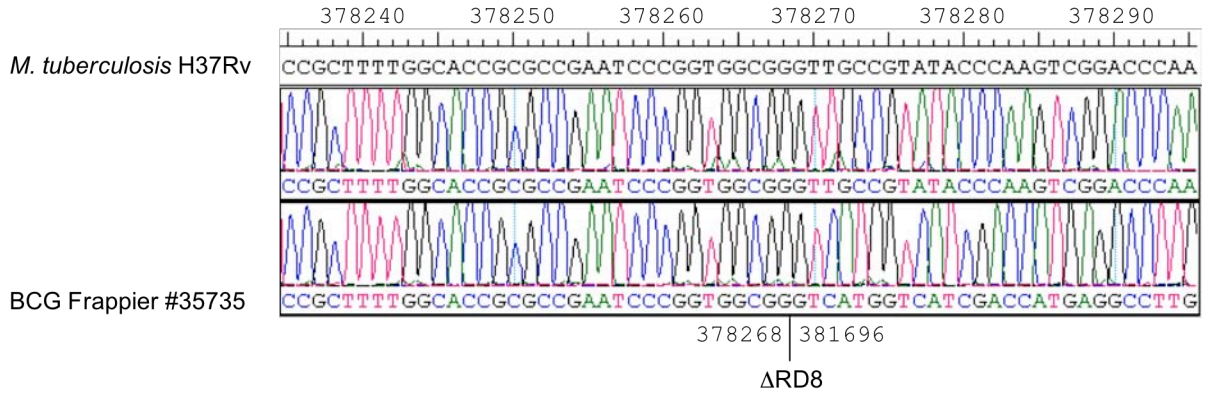



**Figure 36** Excerpt from sequence electropherograms of RD8-PCR amplificates of *M. tuberculosis* H37Rv (identical with sequences of all BCG substrains except BCG Frappier) and *M. bovis* BCG Frappier with *M. tuberculosis* H37Rv enumeration (numbers according to TubercuList).

### *3.2.3.1.2 nRD18*

nRD18, a deletion of 1547 bps, distinguishes BCG substrains Pasteur, Phipps, Frappier, and Tice from substrains Denmark, Russia, and *M. tuberculosis* H37Rv. Due to the small size of the deletion, presence and absence of this region can be demonstrated in a PCR assay containing only two flanking primers (nRD18-flank.F; nRD18-flank.R). These primers amplify a fragment of 249 bps (deletion) and 1796 bps (no deletion). In concordance with the published literature the length of the amplificates demonstrate deletion in nRD18 in BCG Frappier, Phipps, Tice, and Pasteur. BCG Denmark and Russia showed the same amplificate length as *M. tuberculosis* H37Rv, indicating the same sequence for these strains at locus nRD18.

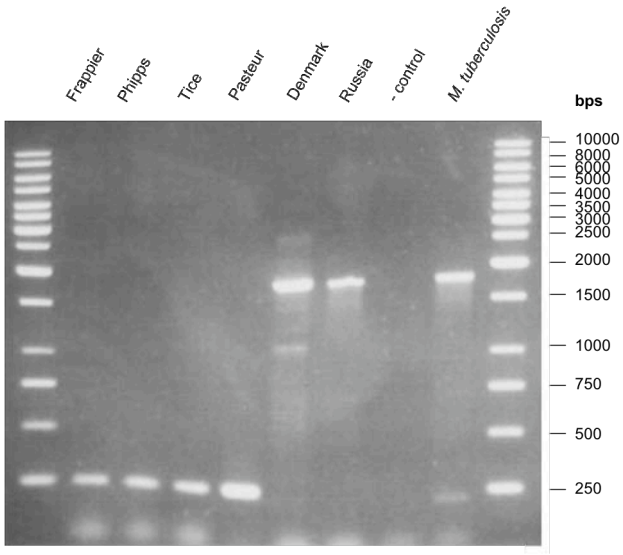



**Figure 37.** 1% agarose gel, ethidium bromide stained demonstrating amplificates of nRD18 region.

### *3.2.3.1.3 RD14*

RD14 is a hallmark of BCG Pasteur. RD14 may be analyzed in two separate PCR reactions, which should give complementary results with respect to a positive or negative amplification. A 518-bp PCR product in the reactions with primers RD14-flank.F and RD14-flank.R is indicative of a deletion, whereas a PCR amplifica-

tion product with primers RD14-Rv1769.int.F and RD14-Rv1769.int.R is indicative of a *M. tuberculosis* like sequence. Out of six analysed substrains, only BCG Pasteur showed a deletion at RD14. The PCR products were analysed on two separate gels. As expected the PCR reaction with primers RD14-Rv1769.int.F and RD14-Rv1769.int.R gave no product for the Pasteur strain, whereas a product was obtained with two primers flanking the deletion (Figure 38).

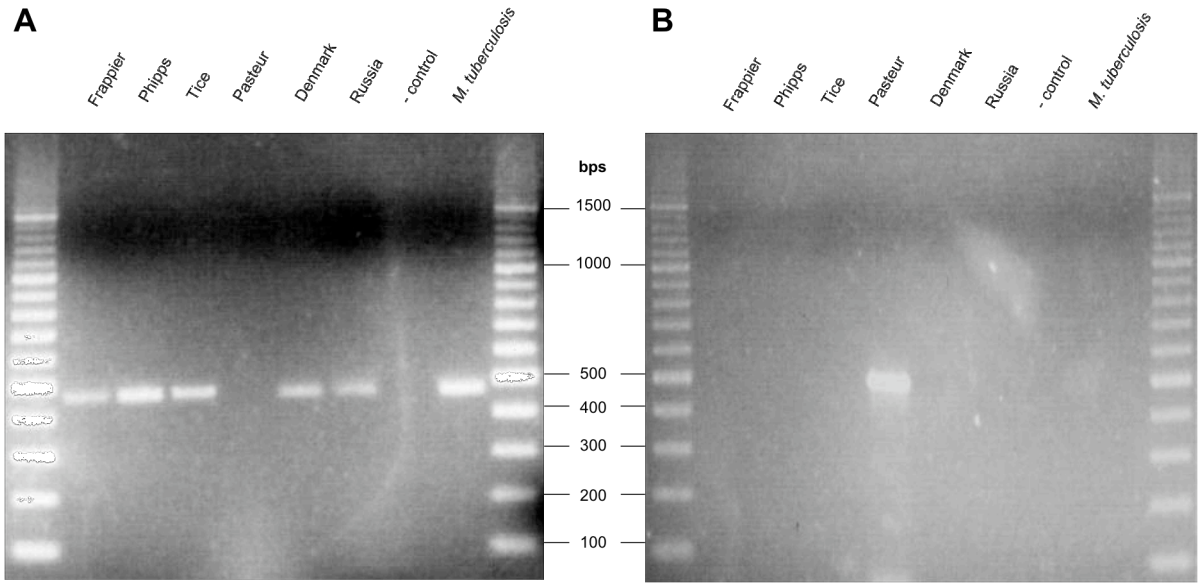



**Figure 38.** 2% agarose gels, ethidium bromide stained. (A) shows PCR amplificates of the RD14 deletion site obtained with a flanking and an internal primer, B) shows the amplificate, which was obtained with two flanking primers.

### *3.2.3.1.4 RDDenmark/Glaxo*

RDDenmark/Glaxo is a 726-bp deletion specific for BCG Denmark and a derivative substrain, BCG Glaxo. Presence or absence of this region was demonstrated in a PCR containing region-flanking primers RDDenmark-flank.F and RDDenmark-flank.R. These primers amplify a 234-bp fragment if the region is absent and a 960-bp fragment if the region is present. RDDenmark/Glaxo was deleted in BCG Denmark 1331, no other of the examined strains presented with this deletion (Figure 39).

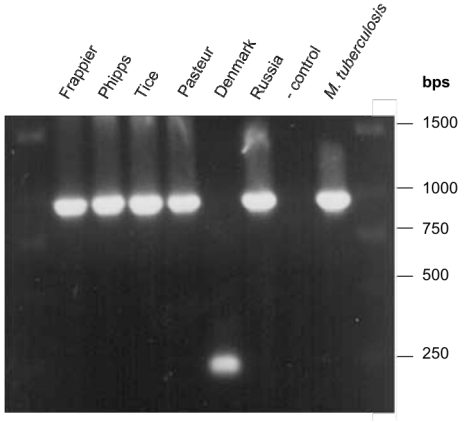



**Figure 39.** 1% agarose gel, ethidium bromide stained. Amplificates over RDDenmark/Glaxo.

### *3.2.3.1.5 RD2*

RD2 differentiates BCG Russia from all other BCG substrains and from *M. tuberculosis* H37Rv. RD2 was investigated by a multiplex PCR using primers RD2-flank.D, RD2-flank.R, and RD2-Rv1979-int.R in an assay similar to the RD1 test developed by Talbot et al. (see 2.7.1.8). A deletion in RD2 results in a PCR product of 1043 bps, which is generated by the annealing of the two flanking primers. If the sequence at the site of RD2 is like in *M. tuberculosis* H37Rv, an 1884-bp product results from amplification with primers RD2-flank.F and RD2-Rv1979-int.R. With the exception of BCG Russia and obviously *M. tuberculosis* H37Rv, RD2 was deleted in all BCG substrains. The respective PCR produced sharply delimitable fragments without unspecific amplificates.

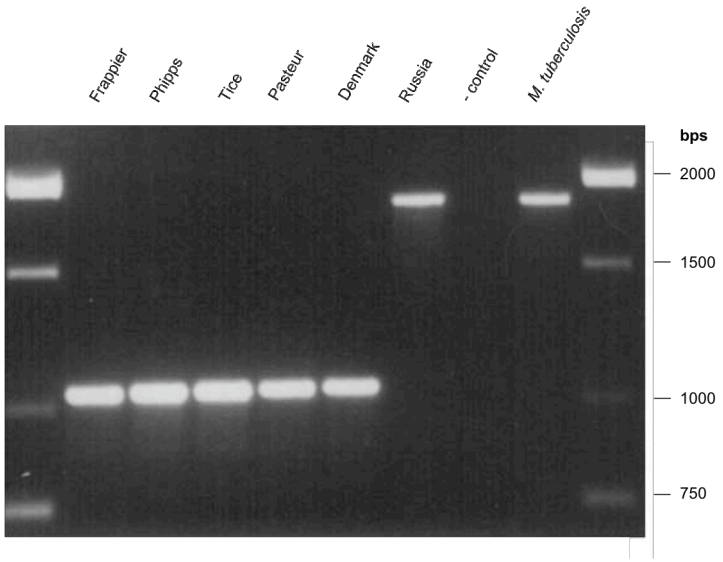



**Figure 40.** 1% agarose gel, ethidium bromide stained. RD2 PCR products.

### *3.2.3.1.6 RDFrappier*

RDFrappier is a hallmark of BCG Frappier. An analogous multiplex PCR setup as for RD1 (see 2.7.1.8) was chosen. If RDFrappier is deleted, two flanking primers (RDFrappier-fl.F and RDFrappier-fl.R) produce a 229-bp amplificate; otherwise the internally laying primer RDFrappier-Rv3495c-int.R can anneal thereby favouring a 177-bp amplificate together with primer RDFrappier-flank.F. The amplificates can be separated in 2% agarose gel. The fragment sizes are clearly distinguishable. RDFrappier was deleted in BCG Frappier only; no other strain displayed a PCR product indicative for this deletion.

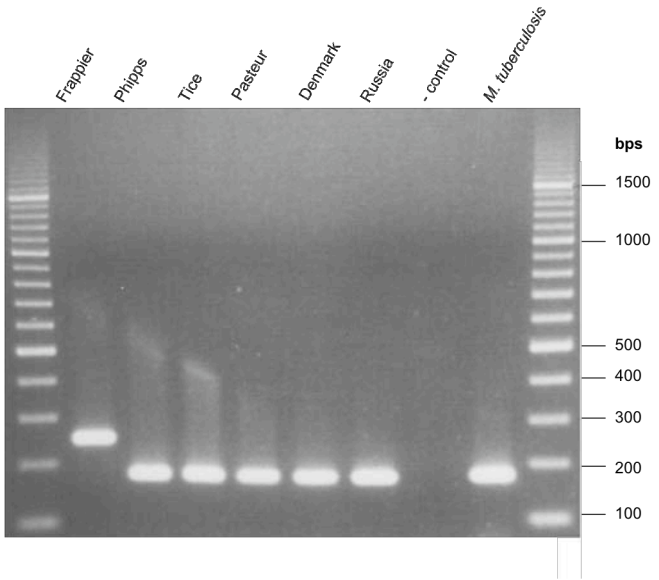



**Figure 41.** 2% agarose gel, ethidium bromide stained. Amplificates of PCR over RDFrappier.

### *3.2.3.1.7 RDRussia*

Besides RD1, RDRussia is the only described region of difference in the genome of BCG Russia. RDRussia separates BCG Russia from other BCG substrains and *M. tuberculosis* H37Rv. The underlying deletion has a lenght of 1603 bps. A PCR with two primers flanking RDRussia (RDRussia-flank.F and RD-Russia-flank.R) was used to investigate RDRussia. The PCR produces a 1956-bp amplificate in H37Rv and BCG substrains with intact RDRussia. If RDRussia is deleted, a 355-bp product results. Unfortunately the PCR over RDRussia produced unspecific amplificates and optimisation of the reaction was tried with variation of $[Mg^{2+}]$, annealing temperature, DNA strand dissolving agents (DMSO), and different polymerase blends. No measure led to the elimination of unspecific bands. Figure 42 shows that the expected fragment of approx 2 kbs for a present RDRussia was demonstrably there in all substrains but BCG Russia which had a distinct fragment of about 350 bps instead. Because a band of similar lenght was present in the PCR product featuring *M. tuberculosis* H37Rv template too, the 2-kb amplificates of *M. tuberculosis* and BCG Denmark were excised from agarose gel and sequenced using primers RDRussia-flank.F and RDRussia-flank.R. The approximate 350-bp fragments from PCR of *M. tuberculosis* and BCG Russia were prepared and sequenced too. Sequencing demonstrated, that both analysed 2-kb fragments (BCG Denmark and H37Rv) have a 100% identical sequence (same sequence as annotated *M. tuberculosis* H37Rv at the RDRussia site). The sequence of the 355-bp fragment obtained from BCG Russia corresponds to a sequence with a deletion in RDRussia. A BLAST search reveals that the 350-bp fragment obtained from *M. tuberculosis* H37Rv corresponds to a repetitive element of PE-PGRS genes present in *M. tuberculosis* complex namely *Rv0980c* (*PE_PGRS18*). The fact, that the chosen primers allow amplification and sequencing of an *Rv0980c* fragment, demonstrates the primers tendency towards mispriming. This is due to a high-degree homology of the respective priming sites (sequence flanking RDRussia and sequence within *Rv0980c*).

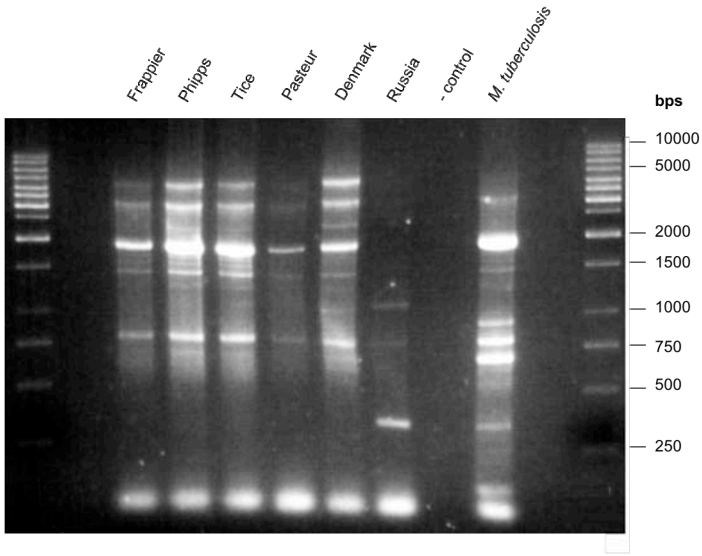



**Figure 42.** 1% agarose gel, ethidium bromide stained. PCR amplificates over RDRussia.

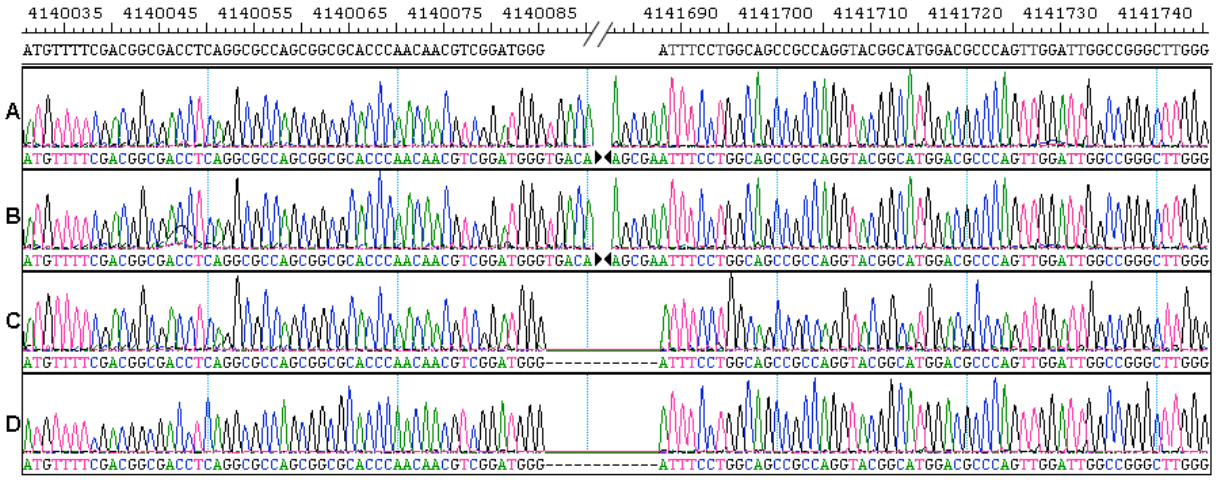



**Figure 43.** Multiple sequence alignment of sequences derived from RDRussia-PCR using primers RDRussia-flank.F and RDRussia-flank.R. **(A)** is the sequence that was obtained from a 2-kbs amplificate using *M. tuberculosis* H37Rv DNA as a template. **(B)** is the sequence of the *M. bovis* BCG Denmark 2-kb amplificate. **(C)** is the sequence of the BCG Russia 350-bp amplificate using RDRussia-flank.F as forward sequencing primer; **(D)** is the reverse complement sequence of the same fragment using RDRussia-flank.R as reverse sequencing primer. **(A)** and **(B)** have the same sequence as *M. tuberculosis* H37Rv at the RDRussia site. In BCG Russia RDRussia is deleted (marked by -).

### *3.2.3.1.8 RD1*

RD1 is deleted in all BCG substrains. This is the single most important region of difference that separates attenuated BCG from *M. bovis* wild-type and from *M. tuberculosis*. RD1 encodes for parts of the ESAT-6 system, which is an important contributor to the virulence of *M. tuberculosis* complex. The extended RD1 region encodes for a dedicated secretion system ESX-1 (assures export of the ESAT-6 protein and its partner, the 10-kDa culture filtrate protein CFP-10) (*140*). Since the tragical Lübeck incident (*35*), where 77 of 251 infants died after they were vaccinated with a BCG vaccine batch contaminated with *M. tuberculosis*, all studies cultivating BCG vaccines should assure, that only non-virulent BCG is present in the culture media. Therefore a multiplex PCR was developed by Talbot et al. (*119*). They developed a diagnostically evaluated multiplex PCR applying three primers: ET1 and ET3 flank RD1, whereas ET2 anneals within RD1 (see Figure 11). In case RD1 is delete ET1 and ET3 amplify a 200-bp fragment, otherwise ET1 and ET3 produce a 150-bp fragment indicating that

RD1 sequence is present. The multiplex PCR method for detection of the RD1 deletion was shown to be highly sensitive and specific for the detection of BCG. The multiprimer assay for RD1 showed that all preparations from BCG substrains had an absent RD1, which excluded contamination of all vaccine strain preparations with *M. bovis* wild-type or even *M. tuberculosis* DNA.

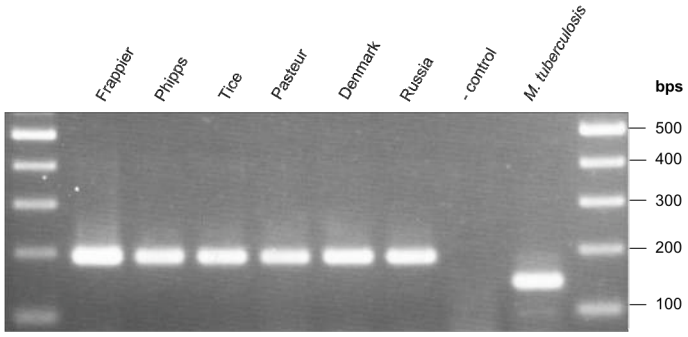



**Figure 44.** 2% agarose gel, ethidium bromide stained. Amplificates of RD1 PCR reaction.

### 3.2.3.2 Mb3700/Rv3676 (cAMP receptor protein, CRP)

The *M. tuberculosis Rv3676* encodes for a cyclic AMP (cAMP) receptor-like protein ($CRP_{Mt}$), a global transcription regulator belonging to the CPR/FNR family, and may play an important role during tuberculosis infection. In general this regulator family controls the expression of many genes in the metabolism of carbons, nitrogen, sulphur, aerobic respiration, denitrification, nitrogen fixation, aromatic acid degradation, bacterial luminescence, and hemolysin synthesis. The $CRP_{Mt}$ ortholog in *Mycobacterium bovis* BCG, $CRP_{BCG}$, is dysfunctional in an *E. coli* CRP competition assay and has been proposed as a potential contributor to *M. bovis* BCG's attenuation (*141*). Many transcriptional regulators form oligomers when bound to DNA. The CRP protein of *M. tuberculosis* recognizes a CRP-like palindromic sequence, suggesting that it also functions as a dimer consisting of two identical proteins. The N-terminal domain of the CRP protein contains a cAMP-binding site and the C-terminal domain contains a helix-turn-helix (HTH) DNA-binding motif. Malltow et al. published that the gene products of *Rv3676* (*M. tuberculosis* H37Rv) and *Mb3700* (*M. bovis* BCG) exhibited different two dimensional electrophoretic mobility (*142*), Spreadbury et al. found specific point mutations in the *Mb3700* gene allowing to differentiate BCG substrains (*116*). To further assure the identity of the substrains that were used in this study, the *Mb3700* locus was amplified and sequenced. Characteristic point mutations were confirmed. The findings were consistent with the published literature: All BCG strains included in this study have a point mutation in comparison to *M. tuberculosis* and *M. bovis* affecting the first helix of the DNA-binding HTN motif. $Guanosine_{532}$ is mutated to an Adenosine, resulting in a $Glu_{178}Lys$ replacement. There is a second point mutation, which is present in "younger" BCG substrains only. $Thymidine_{140}$ is mutated to cytosine which results in amino acid change $Leu_{47}Pro$. This mutation affects the cNMP-binding site of the CRP protein. $Leu_{47}$ is usually conserved in homologous proteins, even in *E. coli* CRP (Table 8). Of the BCG strains in this study all but BCG Russia have the $T_{140} \rightarrow C$ point mutation. There was an additional silent point mutation in $thymidine_{205}$ which is changed to cytosine in all BCG vaccine strains.

**Table 8.** Mutation of $Glu_{178}$ in HTH motif, which is present in all BCG strains examined, comparative chart. Spreadbury et al. (*116*) have not analyzed strains in bold letters so far. Percentage of amino acid identity is indicated with reference to *M. tuberculosis* H37Rv.

| Strain | Ident. (%) | AA# | AA Sequence |
|---|---|---|---|
| *M. tuberculosis* H37Rv | 100 | 176 | TQEEIAQLVGASRETVNKALADF |
| *M. bovis* AF2122/97 | 100 | 176 | ..E..... |
| BCG **Frappier**/**Phipps**/**Tice**/Pasteur/Denmark/Russia | 99 | 176 | ..**K**..... |
| *M. leprae* | 96 | 176 | ..E..... |
| *M. avium* | 96 | 176 | ..E..... |
| *M. marinum* | 95 | 176 | ..E..... |
| *M. smegmatis* MC$^2$ 155 | 97 | 176 | ..E..... |
| *C. glutamicum* | 79 | 179 | ..E.....T. |
| *C. diphteriae* | 78 | 179 | ..E.....T. |
| *S. coelicolor* | 54 | 176 | ..E..... |
| *E. coli* | 32 | 168 | .RQ..G.I..C....GRI.KML |
| | | | Helix Turn Helix |

**Table 9.** Comparative chart displaying the mutation $Leu_{47}$ which marks genealogically younger BCG substrains

| Strain | Ident. (%) | AA# | AA Sequence |
|---|---|---|---|
| *M. tuberculosis* H37Rv | 100 | 25 | QLQPVDFPRGHTVFAEGEPGDRLYIIIS |
| *M. bovis* AF2122/97 | 100 | 25 | .....L..... |
| BCG Russia | 99 | 25 | .....L..... |
| BCG **Frappier**/**Phipps**/**Tice**/Pasteur/Denmark | 99 | 25 | .....**P**..... |
| *M. leprae* | 96 | 25 | .....T.....L...V. |
| *M. avium* | 96 | 25 | .....L...V. |
| *M. marinum* | 95 | 25 | .....L...VA |
| *M. smegmatis* MC$^2$ 155 | 97 | 25 | .....L..... |
| *C. glutamicum* | 79 | 28 | DMET.R....A.I.D.....L...T. |
| *C. diphteriae* | 78 | 28 | D.ES.R....A.I.D.....L...T. |
| *S. coelicolor* | 54 | 25 | SMSE.TLA..D.L.H..D....L.VVTE |
| *E. coli* | 32 | 17 | HCHIHKY.SKSTLIHQGEKAETL.Y.VK |

In summary, molecular investigations of RD8, nRD18, RD14, RDDenmark/Glaxo, RDFrappier, RDRussia, RD1, and *crp* (*Rv3676*) confirmed identity of the BCG substrains.

## 3.3 Generation of *M. bovis* BCG *recA* mutants

Following *M. bovis* BCG substrain characterisation generation of *recA* knock-out mutants was attempted. A total of 11 lots of electrocompetent cells were prepared for transformation experiments. With the exception of BCG Pasteur two different lots of the same substrain were transformed with the suicide knock-out generation plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph* (Figure 25, panel B). Due to yield variations in the preparation procedure of electrocompetent BCG a control to determine the number of bacteria that are spontaneously resistant to hygromycin B or kanamycin could not be included in every transformation experiment. In case of low yield, suicide vector transformation had priority, as the primary intention of this study was to generate *recA* deficient mutants. A total of 49 batches of electrocompetent BCG cells were electroporated.

### 3.3.1 Transformation efficiencies in the different BCG substrains

Suicide vectors integrate with a very low frequency. Therefore the bacteria's competency to take up DNA is of critical importance. To measure transformation efficiency 1 μg of an integrative plasmid with a cassette conferring resistance to hygromycin and kanamycin was used (pMV361-*hyg*). Plasmid pMV361-*hyg* carries an integrase (*int*) and an attachment site (*attP*) of mycobacteriophage L5 (*103*). Integration of plasmid pMV361-*hyg* is based on site-specific recombination between plasmid encoded *attP* and bacterial *attB* site. The *attP* and *attB* sites share a 43-bps core region of identical sequence that overlaps the 3' end of a $tRNA_{Gly}$ gene of mycobacteria such that the integrity of the tRNA gene is maintained following prophage formation (*143*). Site-specific recombination is several orders of magnitude more efficient than homologous recombination of suicide vectors. The assessment of transformation efficiency was performed as outlined in chapter 2.4.2. The first transformation experiments with BCG Frappier, Phipps and Tice showed that the method provided only a raw estimate of the transformation efficiency if the number of CFU per plate exceeded 4000 due to CFU number estimation difficulties. For further experiments the cultures of the control transformations were diluted in three tenfold steps. All dilutions were plated and transformation efficiency was calculated per diluted assay and as a mean value of the three logarithmic dilutions. Negative transformation control was performed by plating of cells that were electroporated without the addition of plasmid DNA. An equivalent of approximately 330 μL (after centrifugation of a diluted overnight culture) of the initial 400 μL transformational sample was plated on 7H10 media containing kanamycin or hygromycin to determine the background of spontaneous kanamycin resistant bacteria among the transformants. The effective number of CFU in the plated sample was determined. The cultures on hygromycin media have never grown spontaneously resistant BCG colonies whereas spontaneous kanamycin resistance could be noted in all BCG substrains. The number of spontaneous kanamycin resistant bacteria in a transformation sample varied between 1 and 300. This confirms observations by Pavelka et al. (*144*) who have never observed spontaneous resistance to hygromycin in the generation of a lysin auxotrophic mutants in BCG Pasteur and BCG Connaught neither.

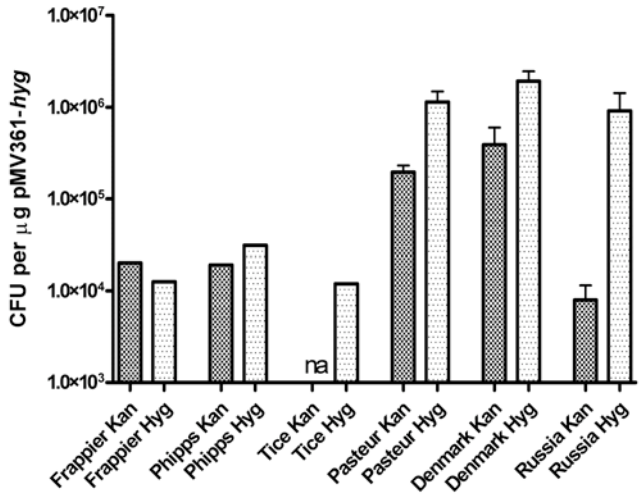



**Figure 45.** Transformation efficiencies (CFU per μg of integrative plasmid pMV361-*hyg*) in different substrains of BCG. Error bars indicate one standard deviation (where applicable).

**Table 10.** Comparison of transformation efficiencies (calculated as mean value of all interpretable transformation control cultures of the six substrains, in brackets number of interpretable samples on which the value bases on). n indicates the number of transformation experiments per strain, $Kan^R$/$Hyg^R$ gives the number of CFU that arose from plating of a negative transformation control sample on the respective antibiotics.

| BCG substrain | n | transformation efficiency per µg of pMV361-*hyg* | | spontaneously resistant | |
|---|---|---|---|---|---|
| | | culture on 7H10-Kanamycin | culture on 7H10-Hygromycin | $Kan^R$ | $Hyg^R$ |
| BCG Frappier | 2 | $2.02 \cdot 10^4$ (1) | $1.26 \cdot 10^4$ (1) | 28 | 0 |
| BCG Phipps | 2 | $1.92 \cdot 10^4$ (1) | $3.11 \cdot 10^4$ (1) | 300 | 0 |
| BCG Tice | 2 | not available[¶] | $1.20 \cdot 10^4$ (1) | 40 | 0 |
| BCG Pasteur | 1 | $1.95 \cdot 10^5 \pm 6.32 \cdot 10^4$ (3) | $1.14 \cdot 10^6 \pm 4.81 \cdot 10^5$ (2) | 1 | 0 |
| BCG Denmark | 2 | $3.89 \cdot 10^5 \pm 4.64 \cdot 10^5$ (5) | $1.92 \cdot 10^6 \pm 1.09 \cdot 10^6$ (4) | 10 | 0 |
| BCG Russia | 2 | $7.95 \cdot 10^3 \pm 6.20 \cdot 10^3$ (3) | $9.09 \cdot 10^5 \pm 7.37 \cdot 10^5$ (2) | 2 | 0 |

[¶] No reliable data available: No logarithmic dilutions of control culture were performed, platting of undiluted control assay resulted in bacterial overgrowth of the Kanamycin plate (> 10000 CFU equaling a transformation efficiency of > $1 \cdot 10^5$ $Kan^R$ CFU per µg of pMV361-*hyg*).

### 3.3.2 Transformation efficiency with suicide *recA* targeting plasmid

As described by Hinds et al. (*102*) an increase in the efficiency of recombinational integrations of alleles on suicide plasmids in the mycobacterial chromosome may be achieved by UV irradiation of the plasmids. For this study an assay to examine this previously established thesis was established. In four BCG substrains (Frappier, Phipps, Tice, and Pasteur) electrocompetent cells were transformed with untreated suicide knock-out vector and a sample of UV irradiated vector (1.5 µg plasmid for each experiment). Recombination efficiency was compared by counting of the $Hyg^R$ (and $Kan^R$ in case of BCG Pasteur) CFU after transformation.

**Table 11.** Results of vector UV irradiation experiments: Two different preparations (cls. 1 and 7) of suicide knock-out vectors were irradiated with 100 $mJcm^{-2}$ and electroporated in 400 µL competent BCG cell preparations. The electroporated cells were resuspended in 5 mL 7H9 medium. Two different amounts of cells were plated. The contents of 1.0 mL and of 4.0 mL 7H9 medium were plated on solid media containing the mentioned antibiotics. No plating on Kanamycin media was carried out in BCG Frappier, Phipps, and Tice.

| BCG substrain | Vector # | Amount plated | $Hyg^R$ colonies | $Hyg^R$ colonies | $Kan^R$ colonies | $Kan^R$ colonies |
|---|---|---|---|---|---|---|
| | | mL | UV irradiated vector | control (not irradiated) | UV irradiated vector | control (not irrad.) |
| BCG Frappier | 1 | 1.0 | 0 | 0 | na | na |
| | 1 | 4.0 | 0 | 30 | | |
| | 7 | 1.0 | 1 | 0 | | |
| | 7 | 4.0 | 50 | 0 | | |
| BCG Phipps | 1 | 1.0 | 0 | 0 | na | na |
| | 1 | 4.0 | 0 | 30 | | |
| | 7 | 1.0 | 1 | 0 | | |
| | 7 | 4.0 | 50 | 0 | | |
| BCG Tice | 1 | 1.0 | 0 | 3 | na | na |
| | 1 | 4.0 | 0 | 10 | | |
| | 7 | 1.0 | 0 | 0 | | |
| | 7 | 4.0 | 1 | 0 | | |
| BCG Pasteur | 1 | 1.0 | 1 | 2 | 0 | 2 |
| | 1 | 4.0 | 2 | 8 | 16 | 25 |
| | 7 | 1.0 | 1 | 1 | 6 | 1 |
| | 7 | 4.0 | 16 | 23 | 25 | 73 |
| BCG Denmark | na | na | na | na | na | na |
| BCG Russia | na | na | na | na | na | na |

### 3.3.3 Selection of transformants and identification of homologous recombinants

Prior to counterselection of transformants, these had to be differentiated form spontaneously resistant $Kan^R/Hyg^R$ bacteria. Therefore colonies growing on hygromycin containing selective media were amplified and analysed by PCR (see 3.4.1) and Southern blot analysis (3.4.2) after transformation with the suicide knock-out plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*. Advantageously the fast PCR screening may be performed from half a colony of one millimetre diameter (the rest may be streaked on a agar plate). It readily separates transformants from spontaneously $Hyg^R/Kan^R$ bacteria but it does not differentiate whether homologous or illegitimate recombination events led to integration of the knock-out plasmid in the bacterial chromosome, which is most reliably determined by Southern blot analysis. In order to identify single-crossover transformants for the *recA* locus, only Southern blot analysis proved to be reliable due to occurrence of illegitimate recombinants (Table 12).

Single-crossover transformants were identified in BCG Frappier, Phipps, Tice, Pasteur, and Denmark, but not in BCG Russia.

**Table 12.** Analysis of recombinants in BCG substrains Pasteur, Frappier, Denmark, and Russia following transformation with *recA* targeting plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*.

| **BCG substrain** | **# $Hyg^R$** | **# analyzed** | **transformants** | | **$Hyg^R$** |
|---|---|---|---|---|---|
| | | **(Southern)** | **HR** | **IR** | with uncertain *recA* genotype |
| Frappier | 90 | 9 | 2 | 1 | 6 |
| Phipps | 83 | 7 | 1 | 0 | 6 |
| Tice | 20 | 7 | 1 | 2 | 4 |
| Pasteur | 54 | 11 | 2 | 3 | 6 |
| Denmark | 377 | 21 | 10 | 2 | 9 |
| Russia | 72 | 23 | 0 | 16 | 7 |

For counterselection single-crossover transformants were subcultured in liquid Middlebrook 7H9 and streaked on sucrose containing Middlebrook 7H10 agar in 10-fold dilution series. A precise calculation of the counterselective efficiency could not be established for the selection of allelic exchange mutants in *M. bovis* BCG using the *sacB* gene as counterselective marker due to throughput limitations in the subsequent screening of candidate allelic exchange mutant (PCR and Southern analysis). Media containing less than 10% sucrose proved inadequate for the selection of allelic exchange mutants; almost no difference in CFU counts could be seen as compared to controls. Therefore only 10% sucrose containing counterselective media were used.

In such a manner *recA* knock-out mutants were isolated in BCG Frappier, Pasteur, and Denmark (Table 13). In BCG Russia, illegitimate integration of the knock-out plasmid was observed in 16 out of 16 transformants analyzed by Southern blot (Figure 48).

**Table 13.** Isolates after counterselection (CS): Southern blot analysis results. $c_{CS}$ gives the number of clones that grew on counterselective plates., n gives the number of clones that were analysed by Southern blot after a *recA* deletion specific PCR; $n_i$ gives the count of interpretable results (equals 100%).

| BCG substrain | Southern | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | $n_i$ | knock-out | 3' sco | | 5' sco | | wild-type | | IR | | no signal |
| Frappier | 7 | 6 | 50% (3/6) | 33.3% | (2/6) | 0% | (0/6) | 16.7% | (1/6) | 0% | (0/6) | 1/7 |
| Phipps | 8 | 8 | 0% (0/8) | 0% | (0/8) | 12.5% | (1/8) | 87.5% | (7/8) | 0% | (0/8) | 0/8 |
| Tice | 36 | 23 | 0% (0/23) | 0% | (0/23) | 0% | (0/23) | 100% | (23/23) | 0% | (0/23) | 13/36 |
| Pasteur | 16 | 15 | 20% (3/15) | 0% | (0/15) | 0% | (0/15) | 75% | (12/15) | 0% | (0/15) | 1/16 |
| Denmark | 12 | 10 | 50% (5/10) | 0% | (0/10) | 0% | (0/10) | 40% | (4/10) | 10% | (1/10) | 2/12 |

## 3.4 Characterisation of *recA* mutants

Putative BCG mutants were analysed in different manners: i.e. PCR amplification at the site of deletion, Southern blot analysis of genomic DNA with respect to the induced mutation, and Western blot analysis after SOS-response induction.

### 3.4.1 Specific PCR for *recA* locus

A specific PCR was established to screen BCG transformants for a mutated *recA* allele. Primers lining within the *recA* gene, adjacent to the deletion site, were chosen. Primer recAPstI_forw and recAPstI_rev were used in the screening PCR. The amplification product from wild-type genomic DNA has a size of 1489 bps whereas the muted allele is significantly shorter with 274 bps. The two different fragments can be discerned easily by agarose gel electrophoresis. The template DNA was gathered by simple boiling preparation either from single BCG colonies or from liquid media grown cells (approximately 0.5 mL densely grown culture).

The flaw of this PCR reaction is the inability to properly distinguish single crossover transformants from knock-out mutants. Because of preferential amplification of the short fragment over the long the presence of a 250-bp and absence of a 1.5-kb fragment does not proof allelic replacement; the screening test does not always produce two amplificates if the template is a single-crossover transformant (see Table 14). It is also susceptible to a false positive reaction: In the screening of BCG Phipps clones after counterselection 7/8 reactions produced 274 bps amplificates even though in Southern blot only 1 clone had a fragment indicative of the deleted *recA* allele. Therefore PCR positive knock-out candidates were subsequently analysed by Southern blotting.

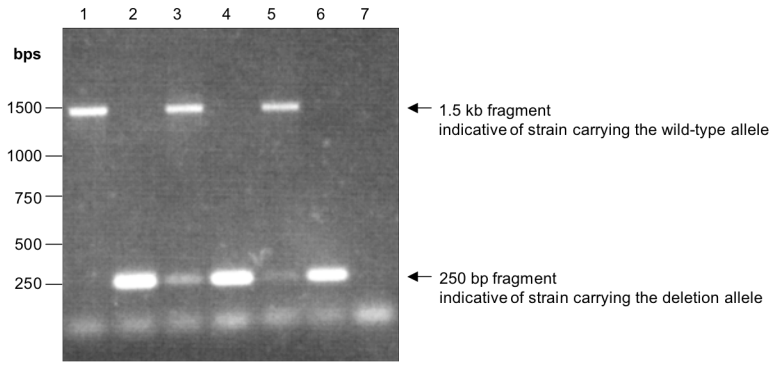



**Figure 46.** Ethidium bromide stained 1% agarosis gel electrophoresis (5 V $cm^{-1}$, 1 hour 20 min) of PCR products (screening of a *M. bovis* BCG Denmark *recA* 5' single crossover transformant [cl. 103] on 7H10 with 10 % sucrose). Lane 1 is the amplificate corresponding to wild-type allele; lanes 2, 4, and 6 contain amplificates of *recA* knock-out mutants; lanes 3 and 5 contain two different amplificates originating from the parental single crossover transformant; lane 7 is the negative control (PCR without template). These findings were subsequently confirmed by Southern blot analysis.

**Table 14.** Isolates after counterselection (CS): PCR screening results. $c_{CS}$ gives the number of colonies that grew on counterselective plates., n gives the number of clones that were analysed by a *recA* deletion specific PCR (see 3.4.1), $n_i$ is the number of interpretable PCR results (equals 100%). A PCR ampificate of 0.25 kb is indicative of a deleted *recA*; a 1.5 kb fragment is indicative of a wild-type *recA*.

| **BCG substrain** | **$c_{CS}$** | **PCR** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **n** | **$n_i$** | **0.25 kb (del.)** | | **1.5 kb (wt)** | | **not interpretable** |
| Frappier | 300 | 33 | 22 | 95.5% | (21/22) | 4.5% | (1/22) | 11/33 |
| Phipps | 9 | 8¶ | 7 | 100% | (7/7) | 0% | (0/7) | 1/8 |
| Tice | 48 | 45¶ | 2 | 50% | (1/2) | 50% | (1/2) | 43/45 |
| Pasteur | 16 | 16 | 3 | 100% | (3/3) | 0% | (0/3) | 13/16 |
| Denmark | 12 | 12 | 12 | 75% | (9/12) | 25% | (3/12) | 0/12 |

¶ One or several colonies could not be propagated from the counter-selective plate and was lost.

## 3.4.2 Southern blot analysis

Colonies raising on counterselective, 10% sucrose containing media from the different BCG substrains were plated on non-selective 7H10 agar dishes after transformation with the *recA* suicide knock-out plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*, and identification of single-crossover transformants on hygromycin containing 7H10 agar. Genomic DNA was extracted and treated with *Sma*I restriction endonuclease. After hybridization with a *Not*I-*Nhe*I *recA* fragment (412 bps) as probe, the transformants could be allocated to five categories: Mutants with wild-type allele of *recA* locus (1962 bps fragment), mutants with suggestive 5' upstream integration of the suicide knock-out plasmid containing the inactivated *recA* gene (1962 bps and 9579 bps fragments), mutants with 3' downstream integration of the plasmid with the mutated *recA* (4088 bps and 5634 bps fragments), mutants with orthologous integration of the inactivated *recA* gene, i.e. *recA* knock-out mutants (5906 bps fragment), and mutants with illegitimate integration of the inactivated *recA* gene (1961 bps wild-type fragment and additional variable fragment). To differentiate putative 5' single-crossover transformants from illegitimate recombinants genomic DNA was digested with *Not*I and Southern blot analysis was performed with the same probe. In this analysis a wild-type strain is characterized by a 6635-bb fragment, a 5' single-crossover transformant by an 8499-bp and a 4800-bp fragment, a 3' single-crossover transformant by a 6635-bp and 7266-kb fragment, and a knock-out strain by a 5402-bp fragment.

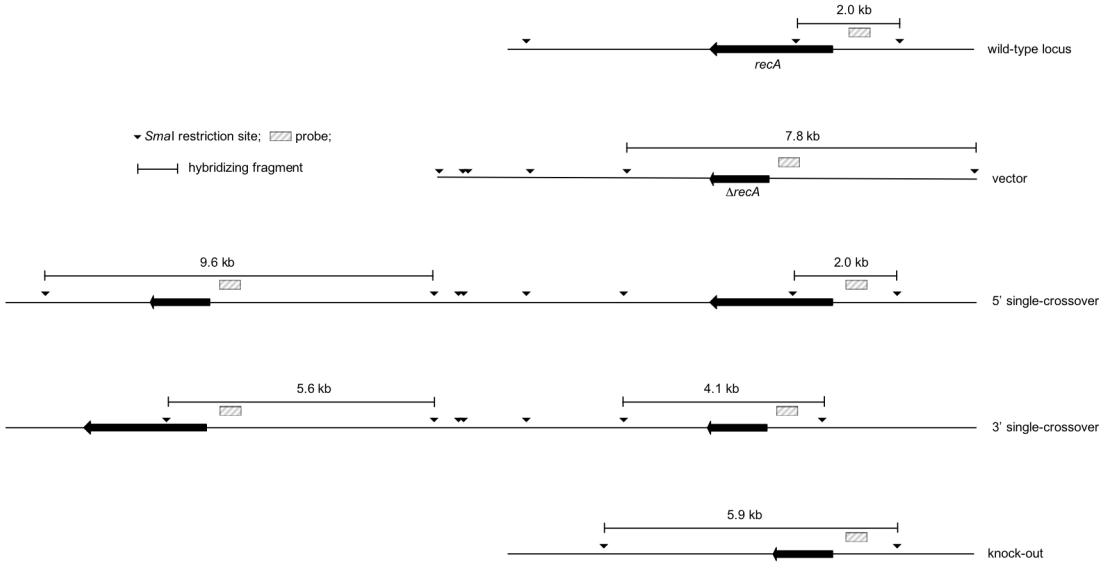



**Figure 47.** Schematic drawing of the BCG *recA* locus: the wild-type locus is shown along with the vector used for inactivation, 5' and 3' single-crossover transformant, knock-out mutant. *Sma*I sites (▾) and the size of the hybridizing fragments are indicated.

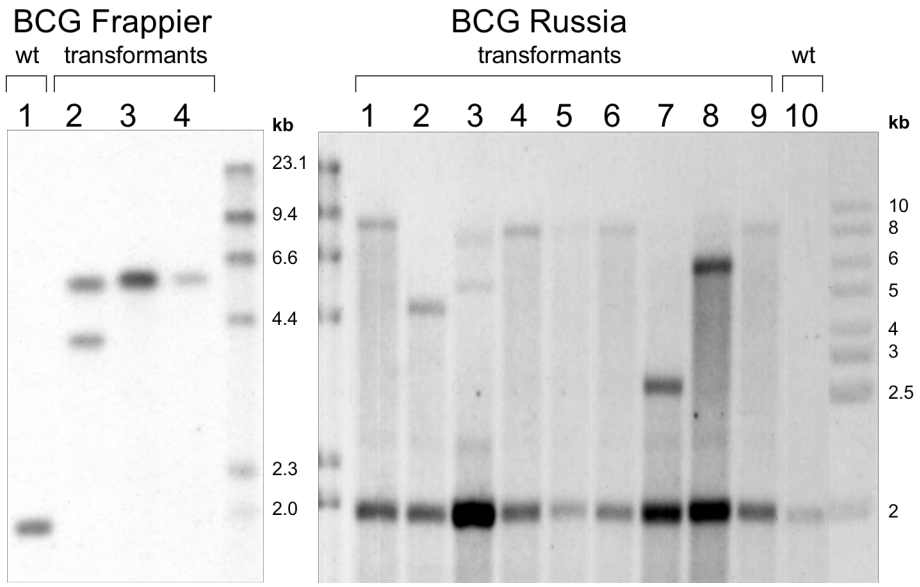



**Figure 48.** Southern blot analysis of the *recA* locus. **(A)** BCG Frappier (ATCC # 35735, CIP 105920). Lane 1, parental *M. bovis* BCG Frappier; lane 2, single-crossover transformant obtained with plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph* (3' single-crossover); lanes 3-4, *recA* (silent) knockout mutants obtained after counterselection of a single-crossover transformant on medium containing 10% sucrose. **(B)** BCG Russia (ATCC # 35740), lanes 1-9 transformants obtained with plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*; lane 10, parental strain. Approximately 500 ng of genomic DNA was digested with *Sma*I and hybridized to a *recA* probe. This blot shows a virtually at random integration of the suicide knockout plasmid containing the inactivated *recA* gene in substrain Russia. The presence of a 2.0-kb fragment corresponding to the original *recA* and the presence of a second fragment of varying size indicates non-homologous illegitimate recombination.

## 3.5 BCG Russia is a natural *recA* mutant

### 3.5.1 Genomic *recA* sequence

BCG Russia integrated the *recA* suicide knock-out plasmid by illegitimate recombination in 16 of 16 cases analyzed by Southern blot. This observation led to the suspicion, that this substrain might be defective in homologous recombination. Therefore, *recA* was amplified using primers recA_start and recA_stop and *Pfu* polymerase and sequenced as outlined in 2.10. A single insertional mutation 413 bps from *recA* start was found. The respective sequence was deposited with the GenBank database accession EU442641. The mutation leads to translational frameshift and truncation by a premature translational stop in the major central RecA domain (residues 31-269). The truncated RecA has a length of 139 instead of 790 amino acids and lacks the complete C-terminal part of the protein including loops L1 (residues 156-165) and L2 (195-210) implicated in DNA binding (*70*). BCG Frappier, Denmark, and Pasteur *recA* were identical with the annotated reference sequence encoding a functional RecA.

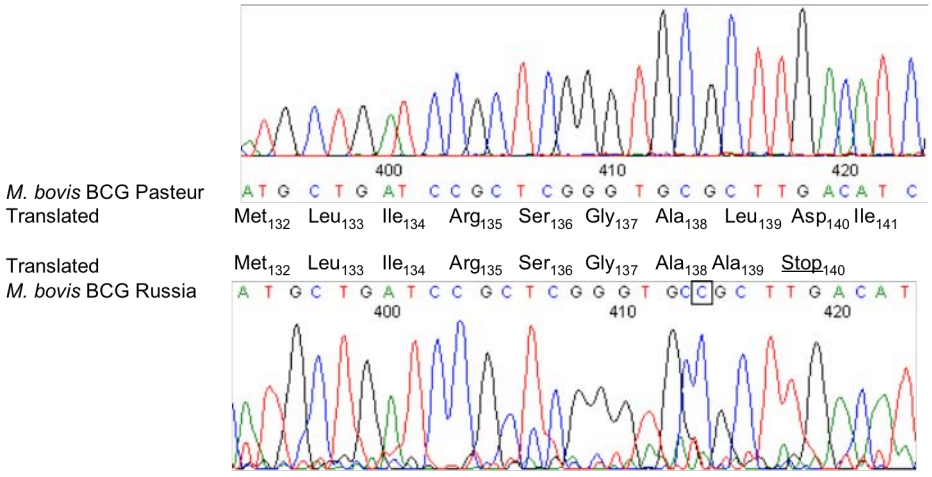



**Figure 49.** Comparative alignment of DNA sequencing electropherograms for *M. bovis* BCG Pasteur 1173P2 and BCG Russia displaying the inserational SNP ($C_{414}$) in the *recA* gene of BCG Russia. Thereby $Leu_{139}$ is changed to Ala, which is immediately followed by a premature stop codon (D140*, * instead of Asp). Nucleotide and amino acid counting starts at the *recA* initatiation codon.

### 3.5.2 RecA expression at protein level after SOS response induction

Bacteria respond to DNA-damage by coordinated expression of a multitude of genes involved in repair and control of cell division – the SOS response (*145*): SOS genes are negatively regulated by LexA respressor protein dimers, which are bound to a 20-bp LexA consensus site (SOS box) upstream of the regulated genes, such as *recA*/*recX* in *M. bovis* BCG and *M. tuberculosis* (see Figure 2). DNA damage with strand crosslinks results in blockage of DNA polymerase at replication forks; single stranded DNA (ssDNA) accumulates at these

sites. In an ATP-dependent fashion, RecA is bound to the ssDNA, forms hexameric filaments, and becomes activated. Activated RecA acts as a coprotease in the autocatalytic digestion of the LexA repressor.

The SOS response was induced by addition of DNA-damaging agent mitomycin C (0.2 $\mu$g mL$^{-1}$) to a BCG culture with an $OD_{600}$ of 0.6 and incubated for additional 24 hours (*80*). Western blot analysis using a RecA antibody demonstrated induction of RecA in strain BCG Pasteur, and absence and failure of induction of RecA in BCG Russia (Figure 50). A BCG Pasteur *recA* mutant was used as a negative control.

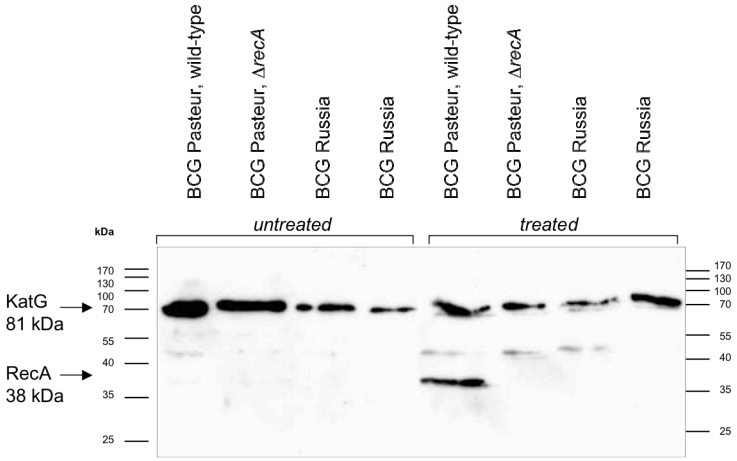



**Figure 50.** Western blot analysis of cell-free protein extract from different BCG preparations (with and without SOS response induction by mitomycin C; 50 µg protein per lane). Detection was carried out using a polyclonal mouse α-RecA antibody (1:1000), mature RecA protein has an approximate molecular weight of 38 kDa. A polyclonal mouse α-KatG antibody was used as loading control (1:2500); KatG protein has a molecular weight of 81 kDa. Bands migrating around 50 kDa are unspecific; secondary antibody was HRP-conjugated goat-anti-mouse antibody (1:2500).

# 4 Discussion

## 4.1 General aspects of *recA* inactivation

### 4.1.1 RecA function and supposed impact on genome evolution

In many bacterial species RecA is a multifunctional and evolutionary conserved protein, that is involved in general recombination and DNA repair. It regulates expression of DNA repair proteins (SOS response) and is the central enzyme of homologous recombination. Besides point mutations, intramolecular homologous recombination, supposedly contributes to genomic evolution in pathogenic mycobacteria due to lack of other evolutionary possibilities because of the high specialization on their ecological niche within phagocytic cells that shield them from contact with other bacteria. Comparative genomics of mycobacteria indicates that recombination events are a major driving force of mycobacterial evolution. Many differences in strains of *M. tuberculosis* can be attributed to RecA-dependent genetic rearrangements (*146, 147*). There is extensive evidence for large-scale rearrangements, duplications and deletions arising from homologous recombination events in *M. leprae* (*148*), *M. tuberculosis* (*149*) and *M. bovis* BCG (*59*). Half of the proteins present in the tubercle bacillus have arisen from gene duplication and adaptation events. Tandem duplications are generally caused by unequal cross-over between homologous sequences or by recombination of short DNA homologies. Homologous recombination between similar sequences may invert or delete genes. Several deletions in *M. tuberculosis* H37Rv genome resulted from recombination between adjacent repeats of IS6110 elements (*149, 150*). Sometimes, the molecular mechanisms underlying alterations at particular loci remain obscure and subsequent alterations may mask initial events. E.g. tandem duplication DU2 in *M. bovis* BCG Pasteur arose from duplication of a 100-kb genomic segment that subsequently incurred an internal deletion of 64 kbs (*59*). In analogy to *M. tuberculosis, recA* inactivation in *M. bovis* BCG should shield the genome from further intramolecular rearrangements. The finding, that an inactivated *recA* coexists in a background of a highly conserved genome in *M. bovis* BCG Russia substantiates the herein proposed impact of RecA mediated events in BCG *in vitro* genome evolution (Fig. 26).

### 4.1.2 Importance of RecA in interspecies gene exchange

RecA inactivation could prevent uptake of virulence genes from the pathogenic *M. tuberculosis* complex into the genome of the attenuated *M. bovis* BCG vaccine in case such mechanisms will be uncovered. In the generation of enterohaemorrhagic *E. coli* (EHEC) live vaccines, the authors denote *recA* deletion as an important safety measure to prevent recombination of virulence genes with the pathogenic strains (*151*). Whether RecA mediates interspecies horizontal gene transfer in mycobacteria has not been clearly demonstrated, while in *Salmonella* RecA-dependent recombination mediates genetic rearrangements resulting in increased genetic instability (*2*).

### 4.1.3 RecA dependent virulence and *in vivo* persistence

For the construction of a live vaccine based on *M. bovis* BCG, the relevancy of RecA function with regards to intraphagocytic persistence and thereon depending immunogenicity had to be assessed. RecA expres-

sion in BCG Pasteur is only upregulated by 1.46 fold 2 hours after professional macrophage cell invasion and not elevated anymore after 24 hours of intracellular survival (*152*). Known virulence factors of *M. tuberculosis* are upregulated up to 4.5 fold after 2 hours in the same microarray analysis. It has been demonstrated in a mouse vaccination and infection model (*31*), that *recA* is not a persistence associated factor in BCG. In fact *recA* deficient BCG survives in the host as long as parental BCG and has equal immunogenetic properites which has been demonstrated by an *M. tuberculosis* infection challenge (*5*). In contrast to BCG, for which *recA* dependence of virulence and thereon depending persistence in the host can be excluded, *recA* has been shown to act as a virulence factor in different bacterial species in numerous reports: It promotes DNA repair mechanisms, which are essential for the intracellular survival of *Salmonella thyphimurium* (*153, 154*). In this species, RecA plays a crucial role in protecting the DNA form the phagocyte oxidative burst during intracellular survival: *recA* mutants were susceptible to killing by macrophages *in vitro* and were shown to be attenuated in a mouse infection model. In *Campylobacter fetus* inversion of surface layer protein (SLP) gene cassettes – a process, which is essential for host immune defense evasion – is RecA dependent (*155*). Immunoescape is facilitated by a mechanism of nested DNA rearrangements involving the inversion of either a single *sapA* promoter-containing element or more tandemly arrayed SLP-cassettes. Intragenic recombination within a single pilin gene or multiple, distinct, partial pilin gene copies (silent, storage loci) is accounted for the Pilin antigenetic diversity in *Neisseria gonorroeae* and changes in the pilation phase (on-off switch after passage through host epithelium) (*156*). In the plant pathogen *Erwinia carotovora* RecA has proteolytical activity and cleaves a gene repressor protein (RgdA, repressor of a pectin-lyase gene) involved in the degradation of plant cell wall constituents (*157*). Noteworthy for this study, *recA* inactivation does not alter measurable parameters of virulence in *Campylobacter jejuni (158)*, *Corynebacterium pseudotuberculosis (4)*, *Brucella abortus (159, 160)*, and certain *Vibrio cholerae* strains (*161*). These findings underline, that *recA* is not necessarily needed for intracellular persistence in all pathogenic bacterial species.

### 4.1.4 Probable compensatory mechanisms for *recA* inactivation

In *E. coli recA* mutants it has been shown, that *radA* compensates for *recA* inactivation to some extent with respect to DNA repair and homologous recombination (*162*). Like RecA, RadA is highly conserved in bacteria; in archea it plays an important role in homologous recombination (*163*). The previously mentioned study by Sander et al. (*5*) exclude that *radA* homolgous gene *Rv3585* (*Mb3616*) compensates for *recA* inactivation related DNA-repair deficits in BCG. This is assumed because *recA* mutants are sensitive to DNA-damaging agents to a comparable extent as *recA* mutants in other species. It is suggested that BCG vaccine strains have a non-redundant *recA* gene; its inactivation decreases homologous recombination events (*31*).

## 4.2 BCG vaccine substrains

The selection of BCG substrains for this study based on genomic analyses more than on description of the relative protective efficacy: It was my goal to include substrain for the whole width of BCG's phylogenetic tree. A review of the published literature about BCG substrain's protective efficacy based on meta-analysis of major field-trials does not conclusively recommend a single most protective BCG substrain (*46*). Therefore only a broad selection of existing BCG strains, that are still in use and commercially available can assure future usability of the generated recombinant strains.

The identity of the BCG substrains included in this work was confirmed by a set of specific region of difference (RD) targeted PCR assays, and phenotypical observations (single colony aspect). The results confirm the identity of the substrains used. Even though this study's strains have different provenience (in four of six cases) than the matching strains that had been analyzed in detail by Behr et al. (*43*), the distinct deletions were also present in the included strains. Targeted PCR assays do not allow the detection of novel regions of difference, therefore total genomic identity of this study's strain selection in comparison the ATCC held reference strains remains to be affirmed.

The RD1 multiprimer PCR assay, which allows discrimination of apathogenic BCG from pathogenic *M. tubeculosis complex*, is the only RD specific PCR assay with published primer sequences (*119, 164*). Mostowy et al. used Affimetrix GeneChip™ to identify absent RDs in different BCG strains. PCR with flanking primers was subsequently applied to identify the exact extent of detected RDs (*115*). For this study, new flanking primers had to be designed, as the primer sequences of the previously mentioned study were not accessible.

In some BCG substrains, single nucleotide polymorphisms supposedly further attenuated different substrains. The orthologs of *M. tuberculosis* H37Rv ORFs *Rv3676* (*crp*), *Rv0445c* (*sigK*), and *Rv0758* (*phoR*), have incurred inactivating mutations in younger BCG substrains (*41, 116, 165*). This study has added evidence to the findings of Spreadbury et al. that point mutation in the DNA- and cAMP-binding domain of *crp* (*Rv2676*) occurred in a certain timeline affecting only younger BCG strains. BCG substrains (Frappier, Tice, and Phipps), that have not been analyzed for *crp* mutations previously, have the same mutations that affected other younger substrains like BCG Pasteur. The results of these sequencing analyses further assure the identity of the substrains used in this study.

## 4.3 WHO vaccine production guidelines

WHO vaccine production guidelines (*104*) were followed wherever possible. The procedures applied in this study allowed generating *M. bovis* BCG *recA* deficient mutants in a vaccine preproduction setting. The feasibility of the herein demonstrated *recA* inactivating strategy under WHO guidelines for BCG vaccine production thus should be possible.

Lot numbers and passage counts were noted and the vials obtained were part of seed lot batches with the lowest passage numbers on the market in order to allow further preclinical studies with the generated mutants.

The mutant generated in BCG Pasteur, and Frappier are derived directly from primary seed lots. The mutant generated in BCG Denmark is based on a production lot with higher passage number. If a *recA* mutant of BCG Denmark shall be used for vaccination trials, generation of the mutant directly form a primary seed lot batch should be considered. The parental strains of the mutant strains represent a selection of BCG substrains that could be used in a study fully complying with WHO vaccine production guidelines.

Central cornerstones in good laboratory practice have been followed; especially contamination with pathogenic *M. tuberculosis complex* is highly unlikely, as those have never been handled in the same laboratory. RD1 multiprimer PCR assay demonstrated absence of RD1 in all vaccine substrains before the experiments.

## 4.4 Generation of unmarked mutants in mycobacteria

Considering further use as a live vaccine or antigen delivery system of the herein described *recA* mutant BCG strains a strategy to generate unmarked knock-out mutants was chosen. The already existing *recA* mutant in BCG Pasteur (*31*) carries a kanamycin resistance marker. Plasmids containing kanamycin resistance have been allowed for US Food and Drug Administration (FDA) approved products in the past but future allowance is uncertain (*6*). Currently, penicillin resistance genes are not allowed to be present in FDA approved live vaccines. Absence of antibiotic resistance leaves the full antibiotic spectrum to the therapist in case a vaccinee is immunocompromised (e.g. has an HIV-infection).

The markerless BCG mutants were generated by application of an allelic replacement strategy as defined by Pelicic and others (*7*). This strategy was successfully applied for the generation of several unmarked mutations in the slow growing *M. tuberculosis* or *M. bovis* BCG: *lprG*-Rv1410 operon (*166*), *hbhA* (Rv0475) (*167*), *lysA* (Rv1293) (*144*), *ureC* (Rv1850) (*168*), and *erp* (*169*). As suggested in parts of the literature (*7, 83, 168*) a sucrose concentration of 10% had to be used in the counterselection process for BCG; in *M. smegmatis* a 2% concentration is efficient. There are concerns that this unphysiological, high concentration could select for unwanted adaptional mutations in BCG, which would alter properties of the recombinant vaccine strains. In a circumventory assay, selection of mutants on 2% and 5% sucrose containing media was tried but with the suicide knock-out plasmid pGEM7-*recA*::3xstop-*hsp60-sacB-hyg-aph*, which was used for all *recA* knock-out experiments, only persistence of integrated plasmid could be noted after counterselection on 2% or 5% sucrose. The conclusion could be that the employed mycobacterial *hsp60* (*groEL1* regulatory region) promoter that regulates *sacB* expression in my construct, does not yield to enough SacB production in order to introduce a lethal effect at 2%/5% sucrose concentration. In the construction of a lysine auxotrophic mutant in BCG Pasteur and Connaught mentioned above Pavelka and Jacobs used 2% sucrose media for the counterslective process (*144*). Their suicide knock-out plasmid pYUB657 contained the same artificial *hsp60-sacB* fusion that was employed in this study. Interestingly, they found a certain frequency of $Suc^r$ $Hyg^r$ isolates that had not undergone a second recombination event after counterselction on 2% sucrose. They suggested, that inactivation of the *sacB* gene had occurred by the action of mobile insertion elements (*170*). In this study a non-representative number of clones that were obtained after plating on 2 or 5% sucrose media, phenotypically $Suc^r$ $Hyg^r$ isolates were analyzed for presence and nucleotide alterations of *sacB*. Thereby no insertional elements could be found in the *sacB* gene. Since only a few $Suc^r$ $Hyg^r$ mutants that haven't undergone a second recombination event could be isolated from 10% sucrose media, *sacB* inactivation is not a major cause of failure of counterselection on low concentration sucrose plates.

So far, a counterselection strategy employing the *sacB* gene is the best-documented way to generate markerless mutants in *M. tuberculosis*, *M. bovis* and *M. smegmatis*. The results of this study confirm pervious findings concerning feasibility and efficiency of such a strategy.

## 4.5 Concluding remarks and outlook

Availability of *recA* mutants in three substrains of BCG is of interest for further development of novel *M. tuberculosis* vaccines. These mutants are deprived of the major recombination pathway and therefore have a more stable genome. This property is important when BCG is considered as antigen delivery system. The in-

creased genetic stability is an advantage in the production of vaccine lots because it reduces the need to keep low passage numbers from primary seed lots.

Comparative genome and transcriptome analyses indicate that BCG Russia is an ancient BCG strain most closely related to the original strain attenuated by Calmette and Guérin (*41*). The finding, that this ancient BCG substrain is a natural *recA* mutant, gave an important clue on probable mechanisms of in vitro genome evolution of BCG.

Considering the tremendously long terms that have to be observed in clinical field trials and the good safety record of certain BCG substrains, development of novel tuberculosis vaccines should relay on improvements and adoptions of BCG rather then on attenuation of virulent *M. tuberculosis (18)*. Due to regulatory requirements safety and potency assessments of live recombinant *M. tuberculosis* vaccines supposedly is more arduous: As long as safety and incapacity to revert to wild-type virulence of a recombinant *M. tuberculosis* vaccine candidate is not proven, production of test vaccine lots needs to be performed under biosafety level 3 conditions. Challenges will be to find appropriate quality dedicated BL-3 facilities and to maintain batch to batch variations as low as possible *(171)*. In the production of BCG or recombinant BCG existing facilities and protocols can be used.

The combination of improved recombinant BCG vaccine with an antigen boost system might prove more efficient than BCG vaccination alone in middle term (*10*). The increase in incidence of tuberculosis in economically poor regions can only be counterfeited by an affordable efficient vaccination.

The herein presented recombinant BCG substrains and the natural *recA* mutant in BCG Russia might be a good starting point for the construction and evaluation of novel tuberculosis vaccines.

# 5 References


1. **Tamas, I. et al.** 50 million years of genomic stasis in endosymbiotic bacteria. *Science. 296* (5577): 2376-9 (2002).
2. **Johnson, B.N. et al.** Construction of *Salmonella* strains with both antigen O4 (of group B) and antigen O9 (of group D). *J Bacteriol. 174* (6): 1911-5 (1992).
3. **Mekalanos, J.J.** Live bacterial vaccines: environmental aspects. *Curr Opin Biotechnol. 5* (3): 312-9 (1994).
4. **Pogson, C.A. et al.** Cloning and manipulation of the *Corynebacterium pseudotuberculosis recA* gene for live vaccine vector development. *FEMS Microbiol Lett. 142* (2-3): 139-45 (1996).
5. **Sander, P. et al.** *Mycobacterium bovis* BCG *recA* deletion mutant shows increased susceptibility to DNA-damaging agents but wild-type survival in a mouse infection model. *Infect Immun. 69* (6): 3562-8 (2001).
6. **Rowland, S.S. et al.** Advancing TB vaccines to Phase I clinical trials in the US: regulatory/ manufacturing/licensing issues. *Tuberculosis (Edinb). 85* (1-2): 39-46 (2005).
7. **Pelicic, V. et al.** Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol Microbiol. 20* (5): 919-25 (1996).
8. **Corbett, E.L. et al.** The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch Intern Med. 163* (9): 1009-21 (2003).
9. **WHO.** Fact sheet N°104: Tuberculosis. In: WHO Media Centre., ed. Geneva, WHO, 2007, pp. 1.
10. **Maartens, G. & Wilkinson, R.J.** Tuberculosis. *Lancet* (2007).
11. **Espinal, M.A. et al.** Global trends in resistance to antituberculosis drugs. World Health Organization-International Union against Tuberculosis and Lung Disease Working Group on Anti-Tuberculosis Drug Resistance Surveillance. *N Engl J Med. 344* (17): 1294-303 (2001).
12. **Nettleman, M.D.** Multidrug-resistant tuberculosis: news from the front. *Jama. 293* (22): 2788-90 (2005).
13. **Drobniewski, F. et al.** Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia. *Jama. 293* (22): 2726-31 (2005).
14. **Zignol, M. et al.** Global incidence of multidrug-resistant tuberculosis. *J Infect Dis. 194* (4): 479-85 (2006).
15. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs – worldwide, 2000-2004. *MMWR Morb Mortal Wkly Rep. 55* (11): 301-5 (2006).
16. **WHO.** Global tuberculosis control: surveillance, planning, financing. *WHO/HTM/TB/2006.362* (2006).
17. **Orme, I.M. et al.** Tuberculosis vaccine development: recent progress. *Trends Microbiol. 9* (3): 115-8 (2001).
18. **Orme, I.M.** Current progress in tuberculosis vaccine development. *Vaccine. 23* (17-18): 2105-8 (2005).
19. **Pathan, A.A. et al.** Boosting BCG with Recombinant Modified Vaccinia Ankara Expressing Antigen 85A: Different Boosting Intervals and Implications for Efficacy Trials. *PLoS ONE. 2* (10): e1052 (2007).
20. **Beveridge, N.E. et al.** Immunisation with BCG and recombinant MVA85A induces long-lasting, polyfunctional Mycobacterium tuberculosis-specific $CD4^+$ memory T lymphocyte populations. *Eur J Immunol. 37* (11): 3089-100 (2007).
21. **Ibanga, H.B. et al.** Early clinical trials with a new tuberculosis vaccine, MVA85A, in tuberculosis-endemic countries: issues in study design. *Lancet Infect Dis. 6* (8): 522-8 (2006).
22. **McShane, H. et al.** Recombinant modified vaccinia virus Ankara expressing antigen 85A boosts BCG-primed and naturally acquired antimycobacterial immunity in humans. *Nat Med. 10* (11): 1240-4 (2004).

23. **Kita, Y. et al.** Novel recombinant BCG and DNA-vaccination against tuberculosis in a *cynomolgus* monkey model. *Vaccine. 23* (17-18): 2132-5 (2005).

24. **Horwitz, M.A. & Harth, G.** A new vaccine against tuberculosis affords greater survival after challenge than the current vaccine in the guinea pig model of pulmonary tuberculosis. *Infect Immun. 71* (4): 1672-9 (2003).

25. **Horwitz, M.A. et al.** Recombinant bacillus Calmette-Guérin (BCG) vaccines expressing the *Mycobacterium tuberculosis* 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. *Proc Natl Acad Sci U S A. 97* (25): 13853-8 (2000).

26. **Horwitz, M.A. et al.** A novel live recombinant mycobacterial vaccine against bovine tuberculosis more potent than BCG. *Vaccine. 24* (10): 1593-600 (2006).

27. **Derrick, S.C. et al.** A polyvalent DNA vaccine expressing an ESAT6-Ag85B fusion protein protects mice against a primary infection with *Mycobacterium tuberculosis* and boosts BCG-induced protective immunity. *Vaccine. 23* (6): 780-8 (2004).

28. **Dietrich, J. et al.** Exchanging ESAT6 with TB10.4 in an Ag85B fusion molecule-based tuberculosis subunit vaccine: efficient protection and ESAT6-based sensitive monitoring of vaccine efficacy. *J Immunol. 174* (10): 6332-9 (2005).

29. **Grode, L. et al.** Increased vaccine efficacy against tuberculosis of recombinant *Mycobacterium bovis* Bacille Calmette-Guérin mutants that secrete listeriolysin. *J Clin Invest. 115* (9): 2472-9 (2005).

30. **Hess, J. et al.** Mycobacterium bovis bacille Calmette-Guérin strains secreting listeriolysin of *Listeria monocytogenes*. *Proc Natl Acad Sci U S A. 95* (9): 5299-304 (1998).

31. **Sander, P. et al.** A *recA* deletion mutant of *Mycobacterium bovis* BCG confers protection equivalent to that of wild-type BCG but shows increased genetic stability. *Vaccine. 21* (27-30): 4124-7 (2003).

32. **Bloom, B.R. & Fine, P.E.** The BCG experience: implications for future vaccines against tuberculosis. In: Bloom, B.R., ed. *Tuberculosis: Pathogenesis, Protection and Controll*. Washington, DC, ASM Press, 1994, pp. 531-557.

33. **Calmette, A. et al.** La vaccination préventive contre la tuberculose par le "BCG". Paris, Masson et cie, 1927.

34. **Weill-Hallé, B. & Turpin, R.** Premiers essais de vaccination antituberculeuse de l'enfant par le Bacille Calmette-Guérin (BCG). *Bull Mem Soc Med l'Hosp de Paris 49*: 1589-1601 (1925).

35. **Rieder, H.L.** [Clarification of the Luebeck infant tuberculosis] article in German. *Pneumologie. 57* (7): 402-5 (2003).

36. **Aronson, J.D.** Protective vaccination against tuberculosis with special reference to BCG vaccination. *Am Rev Tuberc. 58*: 255-281 (1948).

37. **Fine, P.E. & Rodrigues, L.C.** Modern vaccines. Mycobacterial diseases. *Lancet. 335* (8696): 1016-20 (1990).

38. **Guérin, C.** Vaccination Against Tuberculosis. In: Rosenthal, S.R., Little et al., eds. *BCG Vaccination Against Tuberculosis*. Boston, MA, 1957, pp. 48.

39. **Osborn, T.W.** BCG vaccine: an investigation of colony morphology from four different strains after their introduction as seed for vaccine preparation in four production laboratories. *J Biol Stand. 11* (1): 19-27 (1983).

40. **Osborn, T.W.** Changes in BCG strains. *Tubercle. 64* (1): 1-13 (1983).

41. **Brosch, R. et al.** Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A. 104* (13): 5596-601 (2007).

42. **Corbel, M.J.** Reasons for instability of bacterial vaccines. *Dev Biol Stand. 87*: 113-24 (1996).

43. **Behr, M.A. & Small, P.M.** A historical and molecular phylogeny of BCG strains. *Vaccine. 17* (7-8): 915-22 (1999).

44. **Colditz, G.A. et al.** The efficacy of bacillus Calmette-Guérin vaccination of newborns and infants in the prevention of tuberculosis: meta-analyses of the published literature. *Pediatrics. 96* (1 Pt 1): 29-35 (1995).

45. **Brewer, T.F.** Preventing tuberculosis with bacillus Calmette-Guérin vaccine: a meta-analysis of the literature. *Clin Infect Dis. 31 Suppl 3*: S64-7 (2000).

46. **Colditz, G.A. et al.** Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *Jama. 271* (9): 698-702 (1994).

47. **Fine, P.E.** BCG: the challenge continues. *Scand J Infect Dis. 33* (4): 243-5 (2001).

48. **Behr, M.A. & Small, P.M.** Has BCG attenuated to impotence? *Nature. 389* (6647): 133-4 (1997).

49. **WHO Expert Committee on Biological Standardization.** Requirements for dried BCG Vaccine. *WHO Tech Rep Ser. 329*: 23 (1966).

50. **Corbel, M.J. et al.** Report on a WHO consultation on the characterisation of BCG strains, Imperial College, London 15-16 December 2003. *Vaccine. 22* (21-22): 2675-80 (2004).

51. **Ho, M. et al.** Report on a WHO consultation on the characterisation of BCG vaccines, WHO, Geneva, Switzerland 8–9 Dec 2004. *Vaccine. 23* (50): 5700–5704 (2005).

52. **Paulus, H.** Inteins as targets for potential antimycobacterial drugs. *Front Biosci. 8*: s1157-65 (2003).

53. **Perler, F.B. et al.** Compilation and analysis of intein sequences. *Nucleic Acids Res. 25* (6): 1087-93 (1997).

54. **Matic, I. et al.** Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell. 80* (3): 507-15 (1995).

55. **Springer, B. et al.** Lack of mismatch correction facilitates genome evolution in mycobacteria. *Mol Microbiol. 53* (6): 1601-9 (2004).

56. **Sreevatsan, S. et al.** Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A. 94* (18): 9869-74 (1997).

57. **Fleischmann, R.D. et al.** Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J Bacteriol. 184* (19): 5479-90 (2002).

58. **Brosch, R. & Behr, M.A.** Comparative genomics and evolution of *Mycobacterium bovis* BCG. In: Cole, S.T., Eisenach, K.D. et al., eds. *Tuberculosis and the Tubercle Bacillus*, 1 ed. Washington DC, ASM Press, 2004, pp. 584.

59. **Brosch, R. et al.** Comparative genomics uncovers large tandem chromosomal duplications in *Mycobacterium bovis* BCG Pasteur. *Yeast. 17* (2): 111-23 (2000).

60. **Edlund, T. & Normark, S.** Recombination between short DNA homologies causes tandem duplication. *Nature. 292* (5820): 269-71 (1981).

61. **Davis, E.O. et al.** Novel structure of the *recA* locus of *Mycobacterium tuberculosis* implies processing of the gene product. *J Bacteriol. 173* (18): 5653-62 (1991).

62. **Davis, E.O. et al.** Protein splicing in the maturation of *M. tuberculosis recA* protein: a mechanism for tolerating a novel class of intervening sequence. *Cell. 71* (2): 201-10 (1992).

63. **Papavinasasundaram, K.G. et al.** Construction and complementation of a *recA* deletion mutant of *Mycobacterium smegmatis* reveals that the intein in *Mycobacterium tuberculosis recA* does not affect RecA function. *Mol Microbiol. 30* (3): 525-34 (1998).

64. **Guhan, N. & Muniyappa, K.** *Mycobacterium tuberculosis* RecA intein possesses a novel ATP-dependent site-specific double-stranded DNA endonuclease activity. *J Biol Chem. 277* (18): 16257-64 (2002).

65. **Guhan, N. & Muniyappa, K.** *Mycobacterium tuberculosis* RecA intein, a LAGLIDADG homing endonuclease, displays Mn(2+) and DNA-dependent ATPase activity. *Nucleic Acids Res. 31* (14): 4184-91 (2003).

66. **Movahedzadeh, F. et al.** Determination of DNA sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA-damaging agents suggests that two modes of regulation exist. *J Bacteriol. 179* (11): 3509-18 (1997).

67. **Davis, E.O. et al.** DNA damage induction of *recA* in *Mycobacterium tuberculosis* independently of RecA and LexA. *Mol Microbiol. 46* (3): 791-800 (2002).

68. **Gopaul, K.K. et al.** Characterization of the two *Mycobacterium tuberculosis recA* promoters. *J Bacteriol. 185* (20): 6005-15 (2003).

69. **Davis, E.O. et al.** Evidence of selection for protein introns in the *recA*s of pathogenic mycobacteria. *Embo J. 13* (3): 699-703 (1994).

70. **Datta, S. et al.** Crystal structures of *Mycobacterium tuberculosis* RecA and its complex with ADP-AlF(4): implications for decreased ATPase activity and molecular aggregation. *Nucleic Acids Res. 28* (24): 4964-73 (2000).

71. **Shingledecker, K. et al.** Reactivity of the cysteine residues in the protein splicing active center of the *Mycobacterium tuberculosis* RecA intein. *Arch Biochem Biophys. 375* (1): 138-44 (2000).

72. **Holliday, R.** The Induction of Mitotic Recombination by Mitomycin C in *Ustilago* and *Saccharomyces*. *Genetics. 50*: 323-35 (1964).

73. **Potter, H. & Dressler, D.** On the mechanism of genetic recombination: electron microscopic observation of recombination intermediates. *Proc Natl Acad Sci U S A. 73* (9): 3000-4 (1976).

74. **Kumar, R.A. et al.** Functional characterization of the precursor and spliced forms of RecA protein of *Mycobacterium tuberculosis*. *Biochemistry. 35* (6): 1793-802 (1996).

75. **Ganesh, N. & Muniyappa, K.** Characterization of DNA strand transfer promoted by *Mycobacterium smegmatis* RecA reveals functional diversity with *Mycobacterium tuberculosis* RecA. *Biochemistry. 42* (23): 7216-25 (2003).

76. **Vaze, M.B. & Muniyappa, K.** RecA protein of *Mycobacterium tuberculosis* possesses pH-dependent homologous DNA pairing and strand exchange activities: implications for allele exchange in mycobacteria. *Biochemistry. 38* (10): 3175-86 (1999).

77. **Kalpana, G.V. et al.** Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proc Natl Acad Sci U S A. 88* (12): 5433-7 (1991).

78. **Balasubramanian, V. et al.** Allelic exchange in *Mycobacterium tuberculosis* with long linear recombination substrates. *J Bacteriol. 178* (1): 273-9 (1996).

79. **Reyrat, J.M. et al.** The urease locus of *Mycobacterium tuberculosis* and its utilization for the demonstration of allelic exchange in *Mycobacterium bovis* bacillus Calmette-Guérin. *Proc Natl Acad Sci U S A. 92* (19): 8768-72 (1995).

80. **Papavinasasundaram, K.G. et al.** Slow induction of RecA by DNA damage in *Mycobacterium tuberculosis*. *Microbiology. 147* (Pt 12): 3271-9 (2001).

81. **Husson, R.N. et al.** Gene replacement and expression of foreign DNA in mycobacteria. *J Bacteriol. 172* (2): 519-24 (1990).

82. **Frischkorn, K. et al.** Investigation of mycobacterial *recA* function: protein introns in the RecA of pathogenic mycobacteria do not affect competency for homologous recombination. *Mol Microbiol. 29* (5): 1203-14 (1998).

83. **Pelicic, V. et al.** Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A. 94* (20): 10955-60 (1997).

84. **Dedonder, R.** Levansucrase from *Bacillus subtilis*. *Methods Enzymol. 8*: 500-505 (1966).

85. **Blomfield, I.C. et al.** Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. *Mol Microbiol. 5* (6): 1447-57 (1991).

86. **Gay, P. et al.** Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J Bacteriol. 164* (2): 918-21 (1985).

87. **Ried, J.L. & Collmer, A.** An *nptI-sacB-sacR* cartridge for constructing directed, unmarked mutations in gram-negative bacteria by marker exchange-eviction mutagenesis. *Gene. 57* (2-3): 239-46 (1987).

88. **Jager, W. et al.** Expression of the *Bacillus subtilis sacB* gene leads to sucrose sensitivity in the gram-positive bacterium *Corynebacterium glutamicum* but not in *Streptomyces* lividans. *J Bacteriol. 174* (16): 5462-5 (1992).

89. **Pelicic, V. et al.** Expression of the *Bacillus subtilis sacB* gene confers sucrose sensitivity on mycobacteria. *J Bacteriol. 178* (4): 1197-9 (1996).

90. **Oettinger, T. et al.** Development of the *Mycobacterium bovis* BCG vaccine: review of the historical and biochemical evidence for a genealogical tree. *Tuber Lung Dis. 79* (4): 243-50 (1999).

91. **Behr, M.A. et al.** Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science. 284* (5419): 1520-3 (1999).

92. News Bulletin. *Oncologist. 2* (1): 62-63 (1997).

93. **Ferguson, R.G. & Simes, A.B.** BCG vaccination of Indian infants in Saskatchewan. *Tubercle. 30*: 5-11 (1949).

94. **Aronson, N.E. et al.** Long-term efficacy of BCG vaccine in American Indians and Alaska Natives: A 60-year follow-up study. *Jama. 291* (17): 2086-91 (2004).

95. **Reid, J.K. et al.** The effect of neonatal bacille Calmette-Guérin vaccination on purified protein derivative skin test results in Canadian aboriginal children. *Chest. 131* (6): 1806-10 (2007).

96. **Dubos, R.J. et al.** Differential characteristics in vitro and in vivo of several substrains of BCG. III. Multiplication and survival in vivo. *Am Rev Tuberc. 74* (5): 683-98 (1956).

97. **Clements, J.C.** Vaccination – The Current Status of BCG. In: Kaufmann S.H. & Hahn, H., eds. *Mycobacteria and TB*, 1 ed. Basel, Karger, 2003, pp. 46-66.

98. **Honda, I. et al.** Identification of two subpopulations of bacillus Calmette-Guérin (BCG) Tokyo172 substrain with different RD16 regions. *Vaccine. 24* (23): 4969-74 (2006).

99. **Bottger, E.C.** High-efficiency generation of plasmid cDNA libraries using electro-transformation. *Biotechniques. 6* (9): 878-80 (1988).

100. **Kremer, K. et al.** Storage of mycobacterial strains. In: Parish, T. & Stoker, N.G., eds. *Mycobacterium tuberculosis protocols*, 1 ed. New Jersey, Humana Press, 2001, pp. 359-365.

101. **Swiss Agency for the Environment, F.a.L.** Richtlinien: Einstufung von Organismen, Bakterien. In: Federal Department of Environment, Transport, Energy and Communications, ed, Swiss Agency for the Environment, Forest and Landscape, 2003, pp. 98.

102. **Hinds, J. et al.** Enhanced gene replacement in mycobacteria. *Microbiology. 145 ( Pt 3)*: 519-27 (1999).

103. **Stover, C.K. et al.** New use of BCG for recombinant vaccines. *Nature. 351* (6326): 456-60 (1991).

104. **Donikian, R. et al.** Requirements for dried BCG vaccine. *WHO Tech Rep Ser. 745*: 60-62 (1987).

105. **European Pharmacopoeia Commission.** Freeze-dried BCG vaccine. In: Council of Europe, ed. *European Pharmacopoeia*, 3 ed. Strassbourg, Council of Europe, 1997, pp. 442-443.

106. **Harboe, M. & Nagai, S.** MPB70, a unique antigen of *Mycobacterium bovis* BCG. *Am Rev Respir Dis. 129* (3): 444-52 (1984).

107. **Behr, M.A. et al.** A point mutation in the *mma3* gene is responsible for impaired methoxymycolic acid production in *Mycobacterium bovis* BCG strains obtained after 1927. *J Bacteriol. 182* (12): 3394-9 (2000).

108. **Minnikin, D.E. et al.** Mycolic acid patterns of representatives of *Mycobacterium bovis* BCG. *J Gen Microbiol. 130* (10): 2733-6 (1984).

109. **Fomukong, N.G. et al.** Use of gene probes based on the insertion sequence IS986 to differentiate between BCG vaccine strains. *J Appl Bacteriol. 72* (2): 126-33 (1992).

110. **Frothingham, R. & Meeker-O'Connell, W.A.** Genetic diversity in the *Mycobacterium tuberculosis complex* based on variable numbers of tandem DNA repeats. *Microbiology. 144 ( Pt 5)*: 1189-96 (1998).

111. **Behr, M.A.** Comparative genomics of BCG vaccines. *Tuberculosis (Edinb). 81* (1-2): 165-8 (2001).

112. **Behr, M.A.** Correlation between BCG genomics and protective efficacy. *Scand J Infect Dis. 33* (4): 249-52 (2001).

113. **Mostowy, S. & Behr, M.A.** Comparative genomics in the fight against tuberculosis: diagnostics, epidemiology, and BCG vaccination. *Am J Pharmacogenomics. 2* (3): 189-96 (2002).

114. **Behr, M.A.** BCG - different strains, different vaccines? *Lancet Infect Dis. 2* (2): 86-92 (2002).

115. **Mostowy, S. et al.** The in vitro evolution of BCG vaccines. *Vaccine. 21* (27-30): 4270-4 (2003).

116. **Spreadbury, C.L. et al.** Point mutations in the DNA- and cNMP-binding domains of the homologue of the cAMP receptor protein (CRP) in *Mycobacterium bovis* BCG: implications for the inactivation of a global regulator and strain attenuation. *Microbiology. 151* (Pt 2): 547-56 (2005).

117. **Boettger, E.C. & Kirschner, P.** Species Identification of *Mycobacteria* Using rDNA Sequencing. In: Parish, T. & Stoker, N.G., eds. *Mycobacteria Protocols*, 1 ed. Totowa, Humana Press, 1998, pp. 349-361.

118. **Institut Pasteur, Unitée de Génetique Microbienne.** GenoList genome browser. Insitut Pasteur, 2008. http://genolist.pasteur.fr/ (accessed on February 5th, 2008).

119. **Talbot, E.A. et al.** PCR identification of *Mycobacterium bovis* BCG. *J Clin Microbiol. 35* (3): 566-9 (1997).

120. **Garnier, T. et al.** The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A. 100* (13): 7877-82 (2003).

121. **Prober, J.M. et al.** A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science. 238* (4825): 336-41 (1987).

122. **Short, J.M. et al.** Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Res. 16* (15): 7583-600 (1988).

123. **Hoheisel, J.D.** Creation of a rare cutter multiple cloning site. *Biotechniques. 17* (3): 456, 458, 460 (1994).

124. **Sutcliffe, J.G.** Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harb. Symp. Quant. Biol. 43* (1): 77-90 (1979).

125. **Lydiate, D.J. et al.** The *Streptomyces* plasmid SCP2*: its functional analysis and development into useful cloning vectors. *Gene. 35* (3): 223-35 (1985).

126. **Hobbie, S.N. et al.** Engineering the rRNA decoding site of eukaryotic cytosolic ribosomes in bacteria. *Nucleic Acids Res. 35* (18): 6086-93 (2007).

127. **Lenz, O. et al.** The *Alcaligenes eutrophus* H16 *hoxX* gene participates in hydrogenase regulation. *J Bacteriol. 176* (14): 4385-93 (1994).

128. **Pulido, D. et al.** The *hyg* gene promoter from *Streptomyces hygroscopicus*: a novel form of Streptomyces promoters. *Biochem Biophys Res Commun. 151* (1): 270-4 (1988).

129. **National Center for Biotechnology Information.** Entrez Nucleotide database. NCBI, 2005. http://www.ncbi.nlm.nih.gov/sites/entrez (accessed on February 5th, 2008).

130. **Bradford, M.M.** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem. 72*: 248-54 (1976).

131. **National Center for Biotechnology Information.** GenBank. NCBI, 2005. http://www.ncbi.nlm.nih.gov/Genbank/ (accessed on February 5th, 2008).

132. **European Bioinformatics Institute.** EMBL Nucleotide Sequence Database. European Bioinformatics Institute, 2005. http://www.ebi.ac.uk/embl/ (accessed on February 5th, 2008).

133. **Cole, S.T. et al.** Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature. 393* (6685): 537-44 (1998).

134. **TIGR.** TIGR Comprehensive Microbial Resource. J. Craig Venter Institute, 2005. http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi (accessed on Februaray 5th, 2008).

135. **Altschul, S.F. et al.** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25* (17): 3389-402 (1997).

136. **Kirschner, P. & Boettger, E.C.** Species Identification of *Mycobacteria* Using rDNA Sequencing. In: Parish, T. & Stoker, N.G., eds. *Mycobacteria Protocols*, 1 ed. Totowa, N.J., U.S.A., Humana Press, 1998, pp. 349-361.

137. **Bedwell, J. et al.** Identification of substrains of BCG vaccine using multiplex PCR. *Vaccine. 19* (15-16): 2146-51 (2001).

138. **Chen, J.M. et al.** *Mycobacterium bovis* BCG vaccines exhibit defects in alanine and serine catabolism. *Infect Immun. 71* (2): 708-16 (2003).

139. **Philipp, W.J. et al.** Physical mapping of *Mycobacterium bovis* BCG Pasteur reveals differences from the genome map of *Mycobacterium tuberculosis* H37Rv and from *M. bovis*. *Microbiology. 142 ( Pt 11)*: 3135-45 (1996).

140. **Brodin, P. et al.** Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. *Infect Immun. 74* (1): 88-98 (2006).

141. **Bai, G. et al.** The *Mycobacterium bovis* BCG cyclic AMP receptor-like protein is a functional DNA binding protein in vitro and in vivo, but its activity differs from that of its *M. tuberculosis* ortholog, Rv3676. *Infect Immun. 75* (11): 5509-17 (2007).

142. **Mattow, J. et al.** Identification of proteins from *Mycobacterium tuberculosis* missing in attenuated *Mycobacterium bovis* BCG strains. *Electrophoresis. 22* (14): 2936-46 (2001).

143. **Lee, M.H. & Hatfull, G.F.** Mycobacteriophage L5 integrase-mediated site-specific integration in vitro. *J Bacteriol. 175* (21): 6836-41 (1993).

144. **Pavelka, M.S., Jr. & Jacobs, W.R., Jr.** Comparison of the construction of unmarked deletion mutations in *Mycobacterium smegmatis*, *Mycobacterium bovis* bacillus Calmette-Guérin, and *Mycobacterium tuberculosis* H37Rv by allelic exchange. *J Bacteriol. 181* (16): 4780-9 (1999).

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

146. **Brosch, R. et al.** Genomic analysis reveals variation between *Mycobacterium tuberculosis* H37Rv and the attenuated *M. tuberculosis* H37Ra strain. *Infect Immun. 67* (11): 5768-74 (1999).

147. **Fang, Z. et al.** IS6110-mediated deletions of wild-type chromosomes of *Mycobacterium tuberculosis*. *J Bacteriol. 181* (3): 1014-20 (1999).

148. **Cole, S.T. et al.** Massive gene decay in the leprosy bacillus. *Nature. 409* (6823): 1007-11 (2001).

149. **Ho, T.B. et al.** Comparison of *Mycobacterium tuberculosis* genomes reveals frequent deletions in a 20 kb variable region in clinical isolates. *Yeast. 17* (4): 272-82 (2000).

150. **Gordon, S.V. et al.** Genomics of *Mycobacterium bovis*. *Tuberculosis (Edinb). 81* (1-2): 157-63 (2001).

151. **Fuchs, S. et al.** Influence of RecA on in vivo virulence and Shiga toxin 2 production in *Escherichia coli* pathogens. *Microb Pathog. 27* (1): 13-23 (1999).

152. **Rohde, K.H. et al.** *Mycobacterium tuberculosis* invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe. 2* (5): 352-64 (2007).

153. **Buchmeier, N.A. et al.** DNA repair is more important than catalase for *Salmonella* virulence in mice. *J Clin Invest. 95* (3): 1047-53 (1995).

154. **Buchmeier, N.A. et al.** Recombination-deficient mutants of *Salmonella typhimurium* are avirulent and sensitive to the oxidative burst of macrophages. *Mol Microbiol. 7* (6): 933-6 (1993).

155. **Dworkin, J. & Blaser, M.J.** Molecular mechanisms of *Campylobacter fetus* surface layer protein expression. *Mol Microbiol. 26* (3): 433-40 (1997).

156. **Koomey, M. et al.** Effects of *recA* mutations on pilus antigenic variation and phase transitions in Neisseria gonorrhoeae. *Genetics. 117* (3): 391-8 (1987).

157. **Liu, Y. et al.** RecA relieves negative autoregulation of *rdgA*, which specifies a component of the RecA-Rdg regulatory circuit controlling pectin lyase production in *Erwinia carotovora ssp. carotovora*. *Mol Microbiol. 22* (5): 909-18 (1996).

158. **Guerry, P. et al.** Development and characterization of *recA* mutants of *Campylobacter jejuni* for inclusion in attenuated vaccines. *Infect Immun. 62* (2): 426-32 (1994).

159. **Roux, C.M. et al.** RecA and RadA proteins of *Brucella abortus* do not perform overlapping protective DNA repair functions following oxidative burst. *J Bacteriol. 188* (14): 5187-95 (2006).

160. **Tatum, F.M. et al.** Construction of a *Brucella abortus* RecA mutant and its survival in mice. *Microb Pathog. 14* (3): 177-85 (1993).

161. **Stroeher, U.H. et al.** Gene sequence of *recA*+ and construction of *recA* mutants of *Vibrio cholerae*. *Mol Gen Genet. 244* (3): 295-302 (1994).

162. **Beam, C.E. et al.** Role for *radA*/*sms* in recombination intermediate processing in *Escherichia coli*. *J Bacteriol. 184* (24): 6836-44 (2002).

163. **Seitz, E.M. et al.** RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev. 12* (9): 1248-53 (1998).

164. **Mahairas, G.G. et al.** Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent M. bovis. *J Bacteriol. 178* (5): 1274-82 (1996).

165. **Charlet, D. et al.** Reduced expression of antigenic proteins MPB70 and MPB83 in *Mycobacterium bovis* BCG strains due to a start codon mutation in sigK. *Mol Microbiol. 56* (5): 1302-13 (2005).

166. **Bigi, F. et al.** The knockout of the lprG-Rv1410 operon produces strong attenuation of *Mycobacterium tuberculosis*. *Microbes Infect. 6* (2): 182-7 (2004).

167. **Mueller-Ortiz, S.L. et al.** Decreased infectivity despite unaltered C3 binding by a Δ*hbhA* mutant of *Mycobacterium tuberculosis*. *Infect Immun. 70* (12): 6751-60 (2002).

168. **Pelicic, V. et al.** Positive selection of allelic exchange mutants in *Mycobacterium bovis* BCG. *FEMS Microbiol Lett. 144* (2-3): 161-6 (1996).

169. **Berthet, F.X. et al.** Attenuation of virulence by disruption of the *Mycobacterium tuberculosis erp* gene. *Science*. *282* (5389): 759-62 (1998).

170. **Cirillo, J.D. et al.** A novel transposon trap for mycobacteria: isolation and characterization of IS1096. *J Bacteriol. 173* (24): 7772-80 (1991).

171. **Brennan, M.J.** The tuberculosis vaccine challenge. *Tuberculosis (Edinb). 85* (1-2): 7-12 (2005).

# 6 Acknowledgements

First of all, I would like to thank Prof. Dr. Erik C. Böttger for providing me with the great opportunity to contribute to this promising vaccine project at Institut für Medizinische Mikrobiologie.

I want to thank PD Dr. Peter Sander for his introduction into the *recA* project and the continuing supervision of this study with interesting and motivating impacts.

I express my gratitude to the Niedersächsischer Verein zur Bekämpfung der Tuberkulose, Lungen- und Bronchialerkrankungen e. V. and the Universität Zürich for their financial support which allowed working with a representative number of different BCG substrains.

My thanks go to TD Allergen for supplying this study with a free delivery of the Russian substrain of *Mycobacterium bovis* BCG. In this context I would like to thank Dr. Ivan Matuschka for several phone calls to Russia in the process of ordering this specific substrain.

Dr. Elaine O. Davis and Dr. Steward T. Cole are kindly acknowledged for providing RecA and KatG antibodies respectively.

I'd like to thank Dr. Maria Magdalena Senn for providing the methylation deficient DH5α *E. coli* strain and Dr. Sven Hobbie for the construction and provision of pGEM7-*hsp60-sacB*.

I want to thank all lab mates for the good introduction into the laboratory working field and the inspiring atmosphere. In particular, I want to mention Sonja Kuhn and Martin Bosshard for their kind introduction into laboratory techniques. All members of Peter Sander's workgroup have contributed with suggestions in several lab meetings. I'd like to thank Prof. Dr. Brigitte Berger-Bächi for helpful questions in a seminary talk.

In the last I express my gratitude to my parents Barbara and Christian Keller for their generous support and their interest in my research work.

# 7 Curriculum vitae

Peter Michael Keller

**Persönliche Angaben**

| | |
|---|---|
| Geburtsdatum und Geburtsort | 20. Februar 1980, Brugg AG, Schweiz |
| Nationalität | Schweizer |
| Heimatorte | Mandach AG und Hottwil AG, Schweiz |
| Zivilstand | Ledig |
| Fremdsprachen | Englisch, Französisch, Italienisch, Russisch |

**Ausbildung**

| | |
|---|---|
| 1987 - 1992 | Primarschule, Brugg, Schweiz |
| 1992 - 1996 | Bezirksschule, Brugg, Schweiz |
| 1996 - 2000 | Alte Kantonsschule, Aarau, Schweiz |
| Juni 2000 | Matura Typus B (Latein) mit Auszeichnung |
| Oktober 2000 - Oktober 2006 | Studium der Humanmedizin, Universität Zürich, Schweiz |
| Juli 2001 | Erste Vorprüfung für Ärzte |
| August 2002 | Zweite Vorprüfung für Ärzte |
| August 2003 | Schlussprüfung für Ärzte, erster Teil |
| August 2004 | United States Medical Licensing Examination, erster Teil (USMLE I) |
| Oktober 2006 | Staatsexamen an der Universität Zürich |

**Weiterbildung und jetzige Beschäftigung**

| | |
|---|---|
| November 2006 - März 2007 | Militärarzt in Aarau und Birmenstorf |
| März 2007 - jetzt | Assistent Forschung & Diagnostik, Institut für Medizinische Mikrobiologie, Universität Zürich |

**Extrakurrikuläre Aktivitäten**

| | |
|---|---|
| Oktober 2003 - Februar 2005 | Studentenvertreter in der Medizinischen Fakultät der Universität Zürich |
| April 2003 - April 2005 | Vertreter des Verbands Schweizer Medizinstudierender (VSM) |
| Oktober 2003 - Juli 2005 | Vertreter des Verbands Schweizer Medizinstudierender in der Schweizerischen Interfakultätskommission (SMIFK) |