(51) International Patent Classification:  
C07K 14/35 (2006.01)  
C12N 1/36 (2006.01)  
A61K 39/04 (2006.01)

(21) International Application Number:  
PCT/EP2015/062457

(22) International Filing Date:  
3 June 2015 (03.06.2015)

(25) Filing Language:  
English

(26) Publication Language:  
English

(30) Priority Data:  
62007161  
3 June 2014 (03.06.2014)  
US


(74) Agent: REGIMBEAU; 20, rue de Chazelles, F-75847 Paris Cedeax 17 (FR).

(81) Designated States (unless otherwise indicated, for every kind of national protection available):  

(84) Designated States (unless otherwise indicated, for every kind of regional protection available):  

Declarations under Rule 4.17:  
— if inventorship (Rule 4.17(iv))

(54) Title: RECOMBINANT MYCOBACTERIUM BOVIS BCG EXPRESSING ANTIGENS OF THE MYCOBACTERIUM MARINUM ESX-1 SECRETION SYSTEM

(57) Abstract: Recombinant strains of Mycobacterium bovis bacille Calmette-Guerin (M. bovis BCG) comprising a heterologous nucleic acid sequence of Mycobacterium marinum (M. marinum) are provided. In some embodiments the heterologous nucleic acid sequence of M. marinum comprises a plurality of open reading frames, wherein the plurality of open reading frames comprise open reading frames that encode proteins each at least 95% homologous to the Mycobacterium marinum (M. marinum) proteins MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455. In some embodiments the plurality of open reading frames further comprise an open reading frame that encodes a protein at least 95% homologous to the M. marinum proteins MMAR5443, MMAR5444, and MMAR5457. In some embodiments the plurality of open reading frames further comprise open reading frames that encode proteins at least 95% homologous to the M. marinum proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5456, MMAR5458, MMAR5459, MMAR5460, and MMAR5461. Pharmaceutical compositions, kits, methods of inducing an immune response against M. tuberculosis in a subject, and methods of treating an M. tuberculosis infection in a subject are also provided, among other things.
Published:
— with international search report (Art. 21(3))
— with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description (Rules 13bis.4(d)(i) and 48.2(a)(viii))
— with sequence listing part of description (Rule 5.2(a))
RECOMBINANT MYCOBACTERIUM BOVIS BCG EXPRESSING ANTIGENS OF THE MYCOBACTERIUM MARINUM ESX-1 SECRETION SYSTEM

INTRODUCTION

Tuberculosis

Tuberculosis (TB) represents a major public health challenge [1]. In 2012, the World Health Organization (WHO) estimates that 8.6 million people developed the disease and 1.3 million people died from TB [2]. Humans have been scourged by TB for millennia. First, PCR- confirmed cases date back to about 200 years B.C. [3] and still today approximately one third of the world population are latently infected by the bacilli. The reasons for this intriguing evolutionary success of Mycobacterium tuberculosis (Mtb or M. tuberculosis), the etiologic agent of human TB, are manifold. The slow growth of Mtb necessitates long antibiotic therapy rendering treatment susceptible to failure due to non-adherence of the patients [4]. The drugs used involve unpleasant side effects and additionally, absence from work as well as expensive travel to the treatment posts pose economical difficulties to patients, resulting in their drop out of the treatment regimen [5]. Notably, treatment failure is the major fuel for the development of drug resistances [6] with Totally Drug Resistant cases already being reported in several countries [7]. The perilous union with HIV represents a likewise challenging public health priority, as it weakens our most effective barrier against TB, our immune system [8]. A further reason for the difficult and lengthy treatment is the ability of Mtb to transform into a dormancy, or persister state when faced with stress conditions [9]. Cellular metabolism and replication are halted to persist, and thereby escaping the host immune response and chemotherapy [10].

Mycobacteria are part of the new order Corynebacteriaceae [11] that feature an unusually complex and thick cell wall. Hallmarks are long-chain fatty acids called mycolic acids that surround the bacterial cytoplasmic membrane. These mycolic acids are covalently linked to the cell-matrix creating a second hydrophobic barrier outside the cytomembrane [12].
While mycobacteria are considered gram-positive the second membrane executes biological functions comparable to the outer membrane of gram-negative bacteria, such as the uptake of small hydrophilic nutrient via special membrane channels [13]. This protective outer membrane plays an important role in securing the bacillus' integrity in the face of harsh environmental conditions [14]. Yet, at the same time this complicates the secretion process of bacterial effectors across this complex barrier into the host cytosol. To this end mycobacteria employ novel secretion systems that have only recently been uncovered.

**Type VII Secretion Systems**

The breakthrough of discovering the Type VII Secretion Systems (T7SS) in mycobacteria resulted from the detection of unknown secreted antigens in Mtb culture filtrates [15]. Strikingly, these proteins were deficient of the conventional N-terminal signal sequence of about 20 amino acids, such as the well characterized secreted T-cell Antigen 85A, that would direct them to one of the known secretion systems (SS) [16,17]. Additionally, the genome sequence of the paradigm reference strain Mtb H37Rv [18] predicted the flanking genes of these detected secreted antigens to be trans-membrane proteins, conceivably forming chaperones and channels for their secretion [19]. This substantiated the prospect of the presence of a novel secretion apparatus.

Bacteria inherently rely on secretion of effectors into the host phagosome or cytoplasm to shape and adapt to the milieu in their in vivo niche [20]. Due to their complex membrane it was traditionally the gram-negative bacteria that attracted more attention to unravel the mechanisms of secretion [21]. Among the first SS to be characterized was the Sec pathway [22] which recognizes a specific N-terminal sequence of amino acids to guide and propel the proteins into the cytosol. Another secretion pathway employed by Mtb is the Twin-arginine Translocation that exports fully folded proteins [23]. In the past the different SS have been named I - to - VI, with the Type VI Secretion System discovered as recently as 2006 [24]. Considering the presence of the unique group of proteins without signal sequence, its occurrence only in gram-positive bacteria and the mysterious co-dependency
of the different players this recently discovered novel SS was termed Type VII SS in line with the previous nomenclature [25].

The T7SS are dedicated to the secretion of low-molecular-weight proteins, notably the "6-kD Early Secreted Antigenic Target" ESAT-6, or EsxA, and its protein partner "10-kD Culture Filtrate Protein" CFP-10, or EsxB [15]. In the genome of Mtb 11 homologues with sequence similarities to the ESAT-6/CFP-10 family can be found [18,26]. In five cases these genes are encompassed by genes encoding integral inner-membrane proteins, ATP-binding proteins and cell-wall-associated mycosins [27,28], i.e. components of a putative secretion machinery, leading to the postulation that in Mtb, five T7SSs can be distinguished (Figure 1) [29]. Subsequently, these have been named ESX-1-5 (Early Secreted Antigen 6 kD System) whereby the ESX-1 represents the paradigm T7SS.

Phylogenetically, the ESX loci are of ancient origin and have derived from gene duplication [27]. Interestingly, in silico analyses predicted that 52% of the genome of Mtb has resulted from gene duplication events suggesting this to be the major source of genetic variation [30]. The most ancient putative T7SS is ESX-4. It contains the smallest set of genes and its role in infection has not yet been established, questioning its functionality [31]. The ESX-1 locus has been most extensively studied and will be presented in more detail below.

The role of ESX-2 which encodes all core genes has not yet been described. Mutants for ESX-2 genes were shown to be viable suggesting that ESX-2 is not necessary for survival of Mtb [32]. ESX-3, like ESX-1 contains all the basic core components and is conserved in all mycobacterial genomes available today [19]. Its substrates and ESAT-6/CFP-10 homologues can be detected in culture supernatants [33] underlining the functioning of this ESX system. ESX-5 is the second most studied T7SS next to ESX-1 and its functions have recently been characterized [34-38]. It is the ESX system that has most recently evolved in mycobacteria and is only found in slow-growing mycobacteria such as Mtb, M. bovis and M. leprae [39]. In the fish pathogen M. marinum ESX-5 manipulates the macrophage immune response by induction of anti-inflammatory cytokines and host cell death to
promote bacterial spread [36,38]. The importance of the ESX-5 system for viability and
virulence in Mtb has recently been shown using ESX-5 knockout/deletion mutants [40,41].

Comparative genomics revealed the presence of T7SS also in other bacteria that could
provide insights into its role and function in pathogenicity and virulence. In Staphylococcus
aureus, failure to secrete the EsxA and EsxB homologues reduces virulence, dissemination
and colonization [42]. Likewise, a distant ESX locus homologue could be identified in
Bacillus subtilis [43]. T-cell responses against the ESAT-6 homologue in M. leprae [44]
suggest the presence of a T7SS which is further supported by the genome sequence [45].
Yet, the un-culturable status of the leprae bacilli renders efforts into T7SS research
difficult. Across these species the unifying feature of the T7SS is the shared amino acid
(AA) motif Trp-Xaa-Gly of approximately 100 AA, therefore called WXG-100 [17,46].
This suggests that T7SS are present in many more gram-positive bacteria carrying
homologues of the WXG-100 protein family and raises the question about the role of T7SS
in virulence

ESX-1 Locus

Components

The ESX-1 locus is considered the paradigm T7SS as it is partly absent in the TB vaccine
strain M. bovis BCG contributing to its attenuation [47]. According to the recently
established nomenclature for T7SS [29] genes that are present in at least four of the five
ESX loci are called ESX-conserved component (Ecc), with an alphabetic suffix according
to gene order in the paradigm ESX-1 system, i.e. EccA, B, C. Genes that are involved in
secretion but encoded outside the ESX-1 locus are termed ESX-1 secretion-associated
proteins, i.e. Esp.
The ESX-1 locus, in combination with the EspACD operon, consists of 23 genes (12 shown in Table 1) and is centred around *esxA* and *B*, encoding the secreted proteins ESAT-6 and CFP-10. The core genes are *eccAi* which encodes a putative cytoplasmic chaperone with an AAA+ ATPase domain, *eccB, eccD, and eccE* which encode transmembrane proteins. EccCai and EccCbi form a FtsK/SpoIIIE-like ATPase. The mycosin MycP is a membrane associated protease. The genetic cluster encoding *espA, espC, and espD*, which are secreted via ESX-1 are conserved in pathogenic mycobacteria. They represent most likely gene duplicates as they are homologues of the ESX-1 genes *espE, espF*, and *espH* [19].

**Mechanisms of Secretion**

While the exact number of genes required for ESX-1 secretion is still under debate [48] the current working model is presented here (Figure 2). The ESX-1 components form a multi-subunit structure spanning the cell envelope to facilitate export of bacterial effectors across the two lipid membranes [25]. ESAT-6 and CFP-10 form a 2-helix hairpin structure stabilizing each other by multiple hydrophobic interactions [49] before secretion in form of a heterodimer [50-52] (Figure 2). This four-helix bundle is a common feature of all ESX-1
substrates and their homologues [14]. A short secretion motif on the C-terminal flexible tail of CFP-10 is uninvolved in dimerization but attaches as secretion signal to EccCb, a membrane bound ATPase [53]. The same motif has also been identified in other ESX-1 substrates [54]. To catapult the ESAT-6/CFP-10 heterodimer or other substrates into the cytosol EccCb is associated with the membrane bound EccCa [50] which subsequently assemble next to the 11 translocation channels formed by EccD [25]. Both EccCb and EccCa show similarities to SpoIIE/FtsK-like ATPase in Type IV secretion systems in gram-negative bacteria which perform an important part in directing the substrates to the transmembrane channel in an ATP dependent manner [17, 55]. EccCb/EccCa are thus predicted to provide the energy for protein export via T7SS. The pore forming channel in the outer mycomembrane has not yet been identified [14]. Data obtained from mass spectrometry analysis suggest that a single T7SS complex consists of several copies of the transmembrane proteins and ATPases [56].

Region of Difference 1

The fundamental role of a functioning ESX-1 system in virulence and pathogenicity was unravelled when the live TB vaccine M. bovis BCG (Bacille Calmette Guerin) was shown to have lost 38 open reading frames among which the Region of Difference 1 (RD1) of 9.5 kB, deleting essential ESX-1 components [57]. BCG is the only licenced TB vaccine and has been used in the past decades with an astonishing safety record [58,59]. Yet, its efficacy to protect against pulmonary TB remains controversial, ranging from 0-80% [60]. The absence of the RD1 is largely responsible for the reduced efficacy of BCG, and reintroduction of the ESX-1 locus into BCG restores secretion of ESX-1 immunogens, virulence and protective efficacy as vaccine [47]. In BCG, RD1 comprises nine open reading frames encoding nine genes that are deleted, among which the ESAT-6/CFP-10 partner and the EccD transmembrane channel. This results in failure to express the components of the ESX-1 secretion system and hence export its substrates, resulting in reduced virulence [61].
The deletion of RD1 and hence attenuation of BCG resulted most likely from 13 years of passaging M. bovis on glycerol immersed potato slices [62,63]. That RD1 is a major factor in BCG virulence could be shown as its virulence can partly be restored when complementing with RD1 [64]. Likewise, when Mtb is deleted for the RD1 region it looses its virulence [61,65]. However, BCG complemented with RD1 does not regain its complete level of virulence suggesting that there are other factors involved [47].

Comparative genomics are a powerful tool in understanding the impact of different genes on pathogenicity and to explain the variable virulence between Mtb, BCG and M. bovis [62,66]. The mycobacteria are an interesting target for comparative genomics as they represent a high genome homology of up to 99.9% in the M. tuberculosis complex, of which M. bovis BCG is part [67]. This allows extrapolation and interpretation of virulence factors across different mycobacteria. The mycobacterial species comprises around 80 members of mostly environmental, non pathogenic mycobacteria. While the Mtb complex is part of the more recently evolved slow-growing mycobacteria, a member of the more ancient fast-growing mycobacteria, M. marinum, is closely related to Mtb [39]. The agent of fish tuberculosis growths at 30-33°C explaining its non-pathogenicity for humans, except for mild superficial lesions. [68]. M. marinum encodes the paradigm T7SS ESX-1 system which has been shown to be involved in virulence [69]. The homologies between the genome sequence of the ESX-1 system of Mtb and M. marinum are compelling, with the ESAT-6/CFP-10 proteins having 91% and 97% sequence homology, respectively.

There is a need in the art for new strains of M. bovis BCG that combine the ability to induce a greater protective immune response than that induced by M. bovis BCG when introduced into a subject while at the same time having a virulence equal to or lower than the virulence of M. bovis BCG. Such strains will find several uses, such as vaccination against and treatment of M. tuberculosis infections. This invention meets these and other needs in the art.
SUMMARY

This invention encompasses recombinant strains of *Mycobacterium bovis* bacille Calmette-Guerin (*M. bovis* BCG) comprising a heterologous nucleic acid sequence of *Mycobacterium marinum* (*M. marinum*). In some embodiments the heterologous nucleic acid sequence of *M. marinum* comprises a plurality of open reading frames, wherein the plurality of open reading frames comprise open reading frames that encode proteins each at least 95% homologous to the *Mycobacterium marinum* (*M. marinum*) proteins MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455. In some embodiments the plurality of open reading frames further comprise an open reading frame that encodes a protein at least 95% homologous to the *M. marinum* proteins MMAR5443, MMAR5444, and MMAR5457. In some embodiments the plurality of open reading frames further comprise open reading frames that encode proteins at least 95% homologous to the *M. marinum* proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455. In some embodiments the plurality of open reading frames further comprise open reading frames that encode proteins at least 97% homologous to the listed *M. marinum* proteins. In some embodiments the plurality of open reading frames further comprise open reading frames that encode proteins at least 99% homologous to the listed *M. marinum* proteins.

In some embodiments the plurality of open reading frames comprise open reading frames that encode *M. marinum* proteins MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455. In some embodiments the plurality of open reading frames further comprise an open reading frame that encodes the *M. marinum* proteins MMAR5443, MMAR5444, and MMAR5457. In some embodiments the plurality of open reading frames further comprise open reading frames that encode the *M. marinum* proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437,
In some embodiments the open reading frames that encode the listed proteins have the sequence of the corresponding open reading frames present in the nucleic acid sequence of Figure 3. In some embodiments the heterologous nucleic acid sequence comprises the nucleic acid sequence of Figure 3.

In some embodiments the heterologous nucleic acid sequence is present on a plasmid.

In some embodiments the heterologous nucleic acid sequence is integrated into the *M. bovis* BCG chromosome.

In some embodiments the recombinant strain secretes the CFP-10 and ESAT-6 proteins of *M. marinum*.

In some embodiments, the recombinant strain induces a protective immune response greater than the parent *M. bovis* BCG when introduced into a subject, and/or the virulence of the recombinant strain is equal to or lower than the virulence of the parent *M. bovis* BCG.

In some embodiments the heterologous nucleic acid sequence comprises the *M. marinum* nucleic acid sequence inserted in the recombinant pYUB412 vector carried by the bacteria deposited at the CNCM under the reference number 1-4858 on June 3, 2014.

This invention also encompasses pharmaceutical compositions comprising at least one of the recombinant strains of *M. bovis* BCG of the invention. In some embodiments the pharmaceutical compositions further comprises at least one isolated recombinant protein or peptide antigen of a *mycobacterium*. In some embodiments the isolated recombinant protein or peptide antigen of a *mycobacterium* is from a strain selected from *M. bovis* BCG, *M. marinum*, and *Mycobacterium tuberculosis* (*M. tuberculosis*). In some embodiments the isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof.
This invention also encompasses kits comprising at least one of the recombinant strains of *M. bovis* BCG of the invention in a container. In some embodiments the kits further comprise at least one isolated recombinant protein or peptide antigen of a mycobacterium. In some embodiments the isolated recombinant protein or peptide antigen of a mycobacterium is from a strain selected from *M. bovis* BCG, *M. marinum*, and *Mycobacterium tuberculosis* (*M. tuberculosis*). In some embodiments the isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof.

The invention also encompasses methods of inducing an immune response against *M. tuberculosis* in a subject. In some embodiments the immune response is a protective immune response against *M. tuberculosis*. In some embodiments the methods comprise administering an effective dose of a pharmaceutical composition of the invention to a subject and inducing an immune response in the subject that is protective against *M. tuberculosis*. In some embodiments the methods further comprise administering at least one isolated recombinant protein or peptide antigen of a mycobacterium to the subject. In some embodiments the at least one isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof. In some embodiments the methods further comprise administering at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, HI-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and *M. vaccae* to the subject.

The invention also encompasses methods of treating an *M. tuberculosis* infection in a subject. In some embodiments the methods comprise administering an effective dose of a pharmaceutical composition of the invention to a subject and inducing an immune response in the subject that is protective against *M. tuberculosis*. In some embodiments the methods further comprise administering at least one isolated recombinant protein or peptide antigen of a mycobacterium to the subject. In some embodiments the at least one isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein,
and peptides thereof. In some embodiments the methods further comprise administering at least one subunit vaccine to the subject. In some embodiments the methods further comprise administering at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, H1-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and *M. vaccae* to the subject.

The invention also encompasses methods of making a recombinant strain of *M. bovis* BCG of the invention. In some embodiments the methods comprise providing a vector comprising a heterologous nucleic acid sequence comprising a plurality of open reading frames, introducing the vector into *M. bovis* BCG cells, and selecting *M. bovis* BCG cells that stably maintain the heterologous nucleic acid sequence comprising the plurality of open reading frames. In some embodiments the vector is an integrating vector and the method further comprises selecting *M. bovis* BCG cells in which the heterologous nucleic acid sequence comprising the plurality of open reading frames has integrated into the host cell chromosome. In some embodiments the vector is the recombinant pYUB412 vector carried by the bacteria deposited at the CNCM under the reference number 1-4858 on June 3, 2014.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**Figure 1:** The five ESX loci of *Mtb* as shown in reference [29]. The ESX-4 region represents the most ancient ESX system encoding supposedly the core genes required for secretion. The esxA/B family are found in all SS as well as the ATPases of the FtsK/SpoIIE family. Detailed information on mechanisms of secretion have only been described for the systems 1 and 5 [14].

**Figure 2:** The current working model of the ESX-1 SS. CFP-10 and ESAT-6 form a heterodimer and attach via the C-terminal end of CFP-10 to the ATPase EccCb. Latter interacts with EccCa which facilitates transport through the transmembrane protein EccD. How the substrates pass the outer membrane is not yet known. ESX-1 associated proteins EspA and C are secreted in the same fashion. With permission from Laleh Majlessi.
Figure 3: A schematic drawing of the pRDl-MAR-construct.

Figure 4: Core region of insert in integrating cosmids containing the *Mycobacterium marinum* strain ESX-1 region (in comparison to the ESX-1 or RD1 region of *Mycobacterium tuberculosis* H37Rv).

Figure 5: ORFs comprised in the cloned fragment.

Figure 6: The Region of Difference 1 in BCG. The genes enclosed by the black frame are absent from BCG resulting in the loss of ESX-1 function. As a result BCG is less virulent but is unable to induce adaptive immunity against ESX-1 protective antigens [62].

Figure 7: The Xhol restriction profile of the cosmid vector used and PCR analysis of electroporated cells are shown. A) The restriction profile is compared to earlier profiles obtained in the laboratory. B) Two clones, A and B, were PCR positive for all three primers. BCG Pasteur was used for negative and the cosmid vector with the ESX--1 region of *M. marinum* was used as positive control.

Figure 8: Comparison of the gene homology between Mtb and *M. marinum*. The genes encoding the ESAT–6 and CFP–10 proteins show 91% and 97% sequence homology, respectively. Other genes, such as Rv3877 encoding the transmembrane channel EccD are less homologous and differs in 22% of base pairs from its *M. marinum* homologue.

Figure 9: Immunoblot analysis of mid-logIO phase cultures of BCG::ESX-1MM A and B for the secreted proteins ESAT-6 and CFP-10. BCG::pYUB served as negative control not encoding and secreting these proteins. For positive control BCG::RD1 2F9 and *M. marinum* were used. CL = Cell lysate, SN = Supernatant.

Figure 10: T-cell epitope mapping of EspA. Splenocytes from *Mtb*-immunized C3H (H-2K) mice were stimulated with each of the individual 15-mer peptides of the EspA pepscan and IFN–γ was quantified in the culture supernatants using ELISA. Two immunodominant epitopes could be identified, namely the 1–15 and 326–340. Several subdominant epitopes such as 361–375 or 246-260 were not chosen for further analysis.
Figure 11: Screening of T-cell hybridoms specific to CFP—10:1 1-25, C3H-derived BM-DC were infected with H37Rv WT, or ESX--1 deficient mutants or were loaded with the homologous peptide. After incubation overnight the hybridoms were added and IL--2 was quantified. Only the hybridoma that responded to both the peptide as well as the protein secreted by H37Rv WT were chosen, namely XC8, XE12 and YH6.

Figure 12: The XE12 hybridoma produces IL-2 in response to CFP-10:1 l-25 starting at concentrations of less than 15 nM. This is more sensitive and specific than analysis using Western Blot techniques and presents a powerful tool to detect secretion of proteins at even minute levels.

Figure 13: Antigen presenting assay showing mycobacterial secreted pro- teins presented by BM-DCs infected at an M.O.I, of 4 with subsequent twofold dilutions. The bell-shaped curve demonstrates the lowest M.O.I, necessary to ensure presentation of antigens and the highest M.O.I, when mortality interferes with presentation.

Figure 14: Evaluation of the immunogenicity of the BCG::ESX- IMM clone A. T-cell responses of splenocytes from C3H immunized mice against PPD, CFP-10:1 l-25 peptide or CFP-10 recombinant protein. MalE was used as a negative recombinant control antigen. Both BCG::ESX-1Mtb and BCG::ESX-1MM clone A are able to induce strong T-cell responses against CFP-10 in vivo. It has been previously shown that the immunogenicity of ESX-1 substrates is directly proportional to their proper secretion by the ESX-1 machinery.

Figures 15: Virulence studies in SCID mice suggest that the ESX-1 region of M. marinum is less virulent than its Mtb homologue (A in lung, B in spleen). While both BCG::ESX-1MM A and BCG::pYUB have similar levels of CFUs on day 1 the virulence of the former strain is lower as indicated by the CFUs counted on day 21. The initial CFUs for the BCG::ESX-1 Mtb strain seem to be too low but its virulence is demonstrated by high mycobacterial load on day 21 confirming earlier results [64].
Figure 16: FRET based study of phagosomal rupture of the different mycobacterial strains. In A, the phagosomal rupture induced by the strains H37Rv WT is evident by the increased blue (447nm) / green (520nm) ratio. The H37Rv Δ ESX-1 which does not encode the ESAT-6 family anymore shows a much weaker shift towards blue. Interestingly, the recombinant BCG encoding the ESX-1 M. marinum region is able to induce phagosomal rupture (B) compared to the negative controls, exhibiting a similar blue / green ratio as H37Rv WT. Cells were infected with an M.O.I. of 1 and analyzed 4 days after infection.

Figure 17: Survival experiment of SCID mice that were intravenously infected with ~ 10^4 CFU per mouse. The different lines indicate the weight gain or loss of infected SCID mice over time, until they reach the humane endpoint (= 20% loss of their bodyweight, when mice were killed). These virulence studies show that the ESX-1 region of M. marinum confers almost no additional virulence compared to the regular BCG Pasteur, particularly, as the initial dose tested (indicated on the figure) on two mice is slightly higher for BCG :RD1-marinum (=BCG : ESX - 1 M. marinum). The time to humane endpoint experiment was conducted on 10 mice per group, for which the mean weight is shown here. The graphs for Mycobacterium microti infected mice are shown because these strains were tested in the same experiment, using the same BCG control, however, the data are not subject of this application.

Figure 18: Survival experiment of SCID mice that were intravenously infected with ~ 10^4 CFU per mouse. The different lines indicate the weight gain or loss of infected SCID mice over time, until they reach the humane endpoint (= 20% loss of their bodyweight, when mice were killed). These virulence studies show that the ESX-1 region of M. marinum confers much less virulence than the one from M. tuberculosis. The survival of SCID mice infected with BCG :RD1-marinum (=BCG : ESX - 1 M. marinum) is very similar to the survival time of SCID mice infected with BCG Pasteur. The time to humane endpoint experiment was conducted on 10 mice per group, for which the individual weight curves are shown here.
Figure 19: Survival experiment of SCID mice that were intravenously infected with \( \sim 10^6 \) CFU per mouse. The different lines indicate the weight gain or loss of infected SCID mice over time, until they reach the humane endpoint (= 20% loss of their bodyweight, when mice were killed). These virulence studies show that the ESX-1 region of \( M. marinum \) confers much less virulence than the one from \( M. tuberculosis \). The survival of SCID mice infected with BCG::RD1-marinum (=BCG::ESX-1-marinum) is very similar to the survival time of SCID mice infected with BCG Pasteur. The time to humane endpoint experiment was conducted on 10 mice per group, for which the mean weight is shown here.

Figure 20: Induction of CD4+ T-cell responses (as determined by IFN-gamma release by splenocytes) specific to ESX-1 or ESX-5 substrates in C57BL/6 (H-2b) mice, immunized with BCG::pyub (vector control), BGC::RD1-marinum (=BCG::ESX-1M.marinum), BCG::RD1-Mtb strains and \( M. tuberculosis \) H37Rv as control strain. The light gray arrows point to the responses against ESX-1 antigens, the dark-gray arrows point to responses against epitopes from ESX-5 antigens.

Figure 21: Induction of CD4+ T-cell responses (as determined by IFN-gamma release by splenocytes) specific to ESX-1 substrates in C3H (H-2k) mice, immunized with BCG::pyub (vector control), BCG::RD1-marinum (=BCG::ESX-1M.marinum), BCG::RD1-Mtb strains and \( M. tuberculosis \) H37Rv as control strain. The arrows point to the responses against EspA responses which were only observed for the \( M. tuberculosis \) control strain.

Figure 22: Readout of the vaccination experiment using BCG::pYUb (vector control), BCG::RD1-Mtb, BCG::RD1-marinum and the Appe25-pel9 attenuated Mtb vaccine (the latter described in Sayes et al., Cell Host Microbe 2012) in C57BL/6 mice. Mice \( (n = 5/\text{group}) \) were immunized subcutaneously by \( 1 \times 10^6 \) CFU of mycobacterial strains of interest. Four weeks post immunization, mice were challenged by aerosol with \( M. tuberculosis \) H37Rv virulent strain in order to deliver approximately 150 CFU/mouse. Mycobacterial loads in the spleen and lungs were determined at four weeks post challenge.
Figure 23. Virulence test of the BCG: ESX 1_M. marinum in guinea pigs. The recombinant BCG strains BCG ::RD1_M. marinum (=BCG ::RD1-marinum), BCG ::RD1_M. tuberculosis (=BCG ::RD1-2F9) and the virulent M. tuberculosis H37Rv strain are administered by aerosol to groups of 6 guinea pigs each. A. One day after aerosol infection of guinea pigs, the dose, measured in CFU, of each of tested mycobacterium strains is evaluated in lungs of two guinea pigs. B. Six weeks after aerosol infection of guinea pigs by the different mycobacterium strains, the dose of each mycobacterium strain, measured in CFU, is evaluated in lungs and spleen of the four guinea pigs of each group. Dark grey bars show CFU in lungs, light grey bars show CFU in spleens.

DETAILED DESCRIPTION

A. Introduction

A more effective vaccine than BCG is indispensable to achieve eradication of TB by 2050, the declared goal of the Stop TB partnership [70]. In the past decade, the research pipeline has been filled with a number of encouraging TB vaccine candidates [71]. However, recent clinical evidence suggests that subunit vaccines encoding only few TB antigens do not confer adequate protection [72]. Attenuated live vaccines such as recombinant BCG may therefore present as more attractive alternative exposing a variety of antigenic epitopes to host immune cells [73,74]. The experiments reported in the examples made use of the well characterized immunogenicity of the ESX-1 encoded proteins, ESAT-6 and CFP-10, that were shown to elicit protective immunity once reintroduced into BCG [64]. While the Mtb proteins render this vaccine candidate BCG::ESX-IMtb too virulent, the experiments reported herein heterologously expressed the ESX-1 proteins of M. marinum to achieve the same immunogenicity while reducing virulence. These experimental results demonstrate that the recombinant strains of Mycobacterium bovis bacille Calmette-Guerin (M. bovis BCG) comprising a heterologous nucleic acid sequence of Mycobacterium marinum (M. marinum) of this invention are promising vaccine candidates, linking the safety of BCG with the immunogenicity of ESX-1 proteins.
B. Definitions

Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include the plural and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of, biochemistry, enzymology, molecular and cellular biology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art. Certain references and other documents cited herein are expressly incorporated herein by reference. Additionally, all UniProt/SwissProt records cited herein are hereby incorporated herein by reference. In case of conflict, the present specification, including definitions, will control. The materials, methods, and examples are illustrative only and not intended to be limiting.


This disclosure refers to sequence database entries (e.g., UniProt/SwissProt records) for certain protein and gene sequences that are published on the internet, as well as other information on the internet. The skilled artisan understands that information on the internet, including sequence database entries, is updated from time to time and that, for example, the reference number used to refer to a particular sequence can change. Where reference is made to a public database of sequence information or other information on the internet, it is understood that such changes can occur and particular embodiments of
information on the internet can come and go. Because the skilled artisan can find equivalent information by searching on the internet, a reference to an internet web page address or a sequence database entry evidences the availability and public dissemination of the information in question.

Before the present proteins, compositions, methods, and other embodiments are disclosed and described, it is to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. It must be noted that, as used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

The term "comprising" as used herein is synonymous with "including" or "containing", and is inclusive or open-ended and does not exclude additional, unrecited members, elements or method steps.

The full name of amino acids is used interchangeably with the standard three letter and one letter abbreviations for each in this disclosure. For the avoidance of doubt, those are:

Alanine (Ala, A), Arginine (Arg, R), Asparagine (Asn, N), Aspartic acid (Asp, D), Cysteine (Cys, C), Glutamic Acid (Glu, E), Glutamine (Gin, Q), Glycine (Gly, G), Histidine (His, H), Isoleucine (Ile, I), Leucine (Leu, L), Lysine (Lys, K), Methionine (Met, M), Phenylalanine (Phe, F), Proline (Pro, P), Serine (Ser, S), Threonine (Thr, T), Tryptophan (Tip, W), Tyrosine (Tyr, Y), Valine (Val, V).

As used herein, the term "in vitro" refers to events that occur in an artificial environment, e.g., in a test tube or reaction vessel, in cell culture, in a Petri dish, etc., and not within an animal.

As used herein, the term "in vivo" refers to events that occur within a living animal.

As used herein, the term "isolated" refers to a substance or entity that has been (1) separated from at least some of the components with which it was associated when initially produced (whether in nature or in an experimental setting), and/or (2) produced, prepared,
and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or more of the other components with which they were initially associated. In some embodiments, isolated agents are more than about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components.

The "isolated" products of this invention, including isolated nucleic acids, proteins, polypeptides, and antibodies are not products of nature (i.e., "non-naturally occurring"). Rather, the "isolated" nucleic acids, proteins, polypeptides, and antibodies of this invention are "man-made" products. The "isolated" products of this invention can be "markedly different" or "significantly different" from products of nature. By way of non-limiting example, the isolated nucleic acids may be purified, recombinant, synthetic, labeled, and/or attached to a solid substrate. Such nucleic acids can be markedly different or significantly different than nucleic acids that occur in nature. By way of further non-limiting example, the "isolated" proteins, polypeptides, and antibodies of this invention may be purified, recombinant, synthetic, labeled, and/or attached to a solid substrate. Such proteins, polypeptides, and antibodies can be markedly different or significantly different from proteins, polypeptides, and antibodies that occur in nature.

The term "peptide" as used herein refers to a short polypeptide that contains at least 2 amino acids and typically contains less than about 50 amino acids and more typically less than about 30 amino acids. In some embodiments a peptide consists of from 2 to 50, from 2 to 20, from 2 to 10, from 5 to 10, from 5 to 15, from 5 to 20, from 10 to 20, from 10 to 30, from 10 to 40, from 10 to 50, from 20 to 40, or from 20 to 50 amino acids. The term as used herein encompasses analogs and mimetics that mimic structural and thus biological function.
The term "polypeptide" encompasses both naturally-occurring and non-naturally occurring proteins, and fragments, mutants, derivatives and analogs thereof. A polypeptide may be monomeric or polymeric. Further, a polypeptide may comprise a number of different domains each of which has one or more distinct activities. For the avoidance of doubt, a "polypeptide" may be any length greater two amino acids. Accordingly, a "polypeptide" may be a protein or a peptide.

The term "protein" refers to a polypeptide that comprises at least 50 amino acids. A "protein" may have the amino acid sequence of a naturally occurring protein or may be a modified derivative or mutein thereof.

The term "isolated protein" or "isolated polypeptide" is a protein or polypeptide that by virtue of its origin or source of derivation (1) is not associated with naturally associated components that accompany it in its native state, (2) exists in a purity not found in nature, where purity can be adjudged with respect to the presence of other cellular material (e.g., is free of other proteins from the same species) (3) is expressed by a cell from a different species, or (4) does not occur in nature (e.g., it is a fragment of a polypeptide found in nature or it includes amino acid analogs or derivatives not found in nature or linkages other than standard peptide bonds). Thus, a polypeptide that is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be "isolated" from its naturally associated components. A polypeptide or protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art. As thus defined, "isolated" does not necessarily require that the protein, polypeptide, peptide or oligopeptide so described has been physically removed from a cell in which it was synthesized.

The protein or polypeptide can be purified. Preferably, the purified protein or polypeptide is more than 50%, 75%, 85%, 90%, 95%, 97%, 98%, or 99% pure. Within the context of this invention, a purified protein that is more than 50% (etc.) pure means a purified protein sample containing less than 50% (etc.) other proteins. For example, a sample of a protein comprising can be 99% pure if it contains less than 1% contaminating host cell proteins.
The term "polypeptide fragment" as used herein refers to a polypeptide that has a deletion, e.g., an amino-terminal and/or carboxy-terminal deletion compared to a full-length polypeptide, such as a naturally occurring protein. In an embodiment, the polypeptide fragment is a contiguous sequence in which the amino acid sequence of the fragment is identical to the corresponding positions in the naturally-occurring sequence. Fragments typically are at least 5, 6, 7, 8, 9 or 10 amino acids long, or at least 12, 14, 16 or 18 amino acids long, or at least 20 amino acids long, or at least 25, 30, 35, 40 or 45, amino acids, or at least 50 or 60 amino acids long, or at least 70 amino acids long, or at least 100 amino acids long.

The term "fusion protein" refers to a polypeptide comprising a polypeptide or fragment coupled to heterologous amino acid sequences. Fusion proteins are useful because they can be constructed to contain two or more desired functional elements that can be from two or more different proteins. A fusion protein comprises at least 10 contiguous amino acids from a polypeptide of interest, or at least 20 or 30 amino acids, or at least 40, 50 or 60 amino acids, or at least 75, 100 or 125 amino acids. The heterologous polypeptide included within the fusion protein is usually at least 6 amino acids in length, or at least 8 amino acids in length, or at least 15, 20, or 25 amino acids in length. Fusions that include larger polypeptides, such as an IgG Fc region, and even entire proteins, such as the green fluorescent protein ("GFP") chromophore-containing proteins, have particular utility.

Fusion proteins can be produced recombinantly by constructing a nucleic acid sequence which encodes the polypeptide or a fragment thereof in frame with a nucleic acid sequence encoding a different protein or peptide and then expressing the fusion protein. Alternatively, a fusion protein can be produced chemically by crosslinking the polypeptide or a fragment thereof to another protein.

As used herein, a protein has "homology" or is "homologous" to a second protein if the nucleic acid sequence that encodes the protein has a similar sequence to the nucleic acid sequence that encodes the second protein. Alternatively, a protein has homology to a second protein if the two proteins have similar amino acid sequences (Thus, the term "homologous proteins" is defined to mean that the two proteins have similar amino acid
sequences). As used herein, homology between two regions of amino acid sequence (especially with respect to predicted structural similarities) is interpreted as implying similarity in function.

When "homologous" or «similar» is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein.

In cases where two or more amino acid sequences differ from each other by conservative substitutions, the percent sequence identity or degree of homology may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. See, e.g., Pearson, 1994, Methods Mol. Biol. 24:307-31 and 25:365-89.

The following six groups each contain amino acids that are conservative substitutions for one another: 1) Serine, Threonine; 2) Aspartic Acid, Glutamic Acid; 3) Asparagine, Glutamine; 4) Arginine, Lysine; 5) Isoleucine, Leucine, Methionine, Alanine, Valine, and 6) Phenylalanine, Tyrosine, Tryptophan.

Sequence homology or similarity for polypeptides, which is also referred to as percent sequence identity, is typically measured using sequence analysis software. See, e.g., the Sequence Analysis Software Package of the Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 910 University Avenue, Madison, Wis. 53705. Protein analysis software matches similar sequences using a measure of homology assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different
species of organisms or between a wild-type protein and a mutein thereof. See, e.g., GCG Version 6.1.


Exemplary parameters for BLASTp are: Expectation value: 10 (default); Filter: seg (default); Cost to open a gap: 11 (default); Cost to extend a gap: 1 (default); Max. alignments: 100 (default); Word size: 11 (default); No. of descriptions: 100 (default); Penalty Matrix: BLOWSUM62. The length of polypeptide sequences compared for homology will generally be at least about 16 amino acid residues, or at least about 20 residues, or at least about 24 residues, or at least about 28 residues, or more than about 35 residues. When searching a database containing sequences from a large number of different organisms, it may be useful to compare amino acid sequences. Database searching using amino acid sequences can be measured by algorithms other than blastp known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. Pearson, Methods Enzymol. 183:63-98 (1990). For example, percent sequence identity between amino acid sequences can be determined using FASTA with its default parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1, herein incorporated by reference.

In some embodiments, polymeric molecules (e.g., a polypeptide sequence or nucleic acid sequence) are considered to be "homologous" to one another if their sequences are at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% identical. The term
"homologous" necessarily refers to a comparison between at least two sequences (nucleotides sequences or amino acid sequences).

As used herein, a "modified derivative" refers to polypeptides or fragments thereof that are homologous in primary structural sequence to a reference polypeptide sequence but which include, e.g., in vivo or in vitro chemical and biochemical modifications or which incorporate amino acids that are not found in the reference polypeptide. Such modifications include, for example, acetylation, carboxylation, phosphorylation, glycosylation, ubiquitination, labeling, e.g., with radionuclides, and various enzymatic modifications, as will be readily appreciated by those skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes are well known in the art, and include radioactive isotopes such as $^{125}$I, $^{32}$P, $^{35}$S, and $^3$H, ligands that bind to labeled antiligands (e.g., antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands that can serve as specific binding pair members for a labeled ligand. The choice of label depends on the sensitivity required, ease of conjugation with the primer, stability requirements, and available instrumentation. Methods for labeling polypeptides are well known in the art. See, e.g., Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2002).

As used herein, "polypeptide mutant" or "mutein" refers to a polypeptide whose sequence contains an insertion, duplication, deletion, rearrangement or substitution of one or more amino acids compared to the amino acid sequence of a reference protein or polypeptide, such as a native or wild-type protein. A mutein may have one or more amino acid point substitutions, in which a single amino acid at a position has been changed to another amino acid, one or more insertions and/or deletions, in which one or more amino acids are inserted or deleted, respectively, in the sequence of the reference protein, and/or truncations of the amino acid sequence at either or both the amino or carboxy termini. A mutein may have the same or a different biological activity compared to the reference protein.
In some embodiments, a mutein has, for example, at least 85% overall sequence homology or similarity to its counterpart reference protein. In some embodiments, a mutein has at least 90% overall sequence homology or similarity to the wild-type protein. In other embodiments, a mutein exhibits at least 95% sequence identity, or 98%, or 99%, or 99.5% or 99.9% overall sequence identity.

As used herein, a "polypeptide tag for affinity purification" is any polypeptide that has a binding partner that can be used to isolate or purify a second protein or polypeptide sequence of interest fused to the first "tag" polypeptide. Several examples are well known in the art and include a His-6 tag, a FLAG epitope, a c-myc epitope, a Strep-TAGII, a biotin tag, a glutathione 5-transferase (GST), a chitin binding protein (CBP), a maltose binding protein (MBP), or a metal affinity tag.

As used herein, "recombinant" may refer to a biomolecule, e.g., a gene or protein, or to an organism. The term "recombinant" may be used in reference to cloned DNA isolates, chemically synthesized polynucleotides, or polynucleotides that are biologically synthesized by heterologous systems, as well as proteins or polypeptides and/or RNAs encoded by such nucleic acids. A "recombinant" nucleic acid may be a nucleic acid linked to a nucleotide or polynucleotide to which it is not linked in nature. A "recombinant" protein or polypeptide may be (1) a protein or polypeptide linked to an amino acid or polypeptide to which it is not linked in nature; and/or (2) a protein or polypeptide made by transcription and/or translation of a recombinant nucleic acid. Thus, a protein synthesized by a microorganism is recombinant, for example, if it is synthesized from an mRNA synthesized from a recombinant nucleic acid present in the cell. A "recombinant" organism is an organism comprising a "recombinant" biomolecule. For example, a "recombinant" strain of M. bovis BCG is a strain of M. bovis BCG that comprises a "recombinant" nucleic acid.

The term "polynucleotide", "nucleic acid molecule", "nucleic acid", or "nucleic acid sequence" refers to a polymeric form of nucleotides of at least 10 bases in length. The term includes DNA molecules (e.g., cDNA or genomic or synthetic DNA) and RNA molecules.
(e.g., mRNA or synthetic RNA), as well as analogs of DNA or RNA containing non-natural nucleotide analogs, non-native internucleoside bonds, or both. The nucleic acid can be in any topological conformation. For instance, the nucleic acid can be single-stranded, double-stranded, triple-stranded, quadruplexed, partially double-stranded, branched, hairpinned, circular, or in a padlocked conformation. The nucleic acid (also referred to as polynucleotides) may include both sense and antisense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers of the above. They may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphorami dates, carbamates, etc.), charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), pendant moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen, etc.), chelators, alkylators, and modified linkages (e.g., alpha anomic nucleic acids, etc.) Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule. Other modifications can include, for example, analogs in which the ribose ring contains a bridging moiety or other structure such as the modifications found in "locked" nucleic acids.

A "synthetic" RNA, DNA or a mixed polymer is one created outside of a cell, for example one synthesized chemically.

The term "nucleic acid fragment" as used herein refers to a nucleic acid sequence that has a deletion, e.g., a 5'-terminal or 3'-terminal deletion compared to a full-length reference nucleotide sequence. In an embodiment, the nucleic acid fragment is a contiguous sequence in which the nucleotide sequence of the fragment is identical to the corresponding positions in the naturally-occurring sequence. In some embodiments fragments are at least 10, 15, 20, or 25 nucleotides long, or at least 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120,
130, 140, or 150 nucleotides long. In some embodiments a fragment of a nucleic acid sequence is a fragment of an open reading frame sequence. In some embodiments such a fragment encodes a polypeptide fragment (as defined herein) of the protein encoded by the open reading frame nucleotide sequence.

The nucleic acid can be purified. Preferably, the purified nucleic acid is more than 50%, 75%, 85%, 90%, 95%, 97%, 98%, or 99% pure. Within the context of this invention, a purified nucleic acid that is at least 50% pure means a purified nucleic acid sample containing less than 50% other nucleic acids. For example, a sample of a plasmid can be at least 99% pure if it contains less than 1% contaminating bacterial DNA.

As used herein, an endogenous nucleic acid sequence in the genome of an organism (or the encoded protein product of that sequence) is deemed "recombinant" herein if a heterologous sequence is placed adjacent to the endogenous nucleic acid sequence. In this context, a heterologous sequence is a sequence that is not naturally adjacent to the endogenous nucleic acid sequence, whether or not the heterologous sequence is itself endogenous (originating from the same host cell or progeny thereof) or exogenous (originating from a different host cell or progeny thereof). By way of example, a promoter sequence can be substituted (e.g., by homologous recombination) for the native promoter of a gene in the genome of a host cell, such that this gene has an altered expression pattern. This gene would now become "recombinant" because it is separated from at least some of the sequences that naturally flank it.

A nucleic acid is also considered "recombinant" if it contains any modifications that do not naturally occur to the corresponding nucleic acid in a genome. For instance, an endogenous coding sequence is considered "recombinant" if it contains an insertion, deletion or a point mutation introduced artificially, e.g., by human intervention. A "recombinant nucleic acid" also includes a nucleic acid integrated into a host cell chromosome at a heterologous site and a nucleic acid construct present as an episome.
As used herein, the phrase "degenerate variant" of a reference nucleic acid sequence encompasses nucleic acid sequences that can be translated, according to the standard genetic code, to provide an amino acid sequence identical to that translated from the reference nucleic acid sequence. The term "degenerate oligonucleotide" or "degenerate primer" is used to signify an oligonucleotide capable of hybridizing with target nucleic acid sequences that are not necessarily identical in sequence but that are homologous to one another within one or more particular segments.

The term "percent sequence identity" or "identical" in the context of nucleic acid sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over a stretch of at least about nine nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 32, and even more typically at least about 36 or more nucleotides. There are a number of different algorithms known in the art which can be used to measure nucleotide sequence identity. For instance, polynucleotide sequences can be compared using FASTA, Gap or Bestfit, which are programs in Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wis. FASTA provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. Pearson, Methods Enzymol. 183:63-98 (1990). For instance, percent sequence identity between nucleic acid sequences can be determined using FASTA with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) or using Gap with its default parameters as provided in GCG Version 6.1, herein incorporated by reference. Alternatively, sequences can be compared using the computer program, BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990); Gish and States, Nature Genet. 3:266-272 (1993); Madden et al, Meth. Enzymol. 266:131-141 (1996); Altschul et al, Nucleic Acids Res. 25:3389-3402 (1997); Zhang and Madden, Genome Res. 7:649-656 (1997)), especially blastp or tblastn (Altschul et al, Nucleic Acids Res. 25:3389-3402 (1997)).
"Stringent hybridization conditions" and "stringent wash conditions" in the context of nucleic acid hybridization experiments depend upon a number of different physical parameters. Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, solvents, the base composition of the hybridizing species, length of the complementary regions, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. One having ordinary skill in the art knows how to vary these parameters to achieve a particular stringency of hybridization.

In general, "stringent hybridization" is performed at about 25°C below the thermal melting point (Tm) for the specific DNA hybrid under a particular set of conditions. "Stringent washing" is performed at temperatures about 5°C lower than the Tm for the specific DNA hybrid under a particular set of conditions. The Tm is the temperature at which 50% of the target sequence hybridizes to a perfectly matched probe. See Sambrook et al., Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), page 9.51. For purposes herein, "stringent conditions" can be defined for solution phase hybridization as aqueous hybridization (i.e., free of formamide) in 6xSSC (where 20xSSC contains 3.0 M NaCl and 0.3 M sodium citrate), 1% SDS at 65°C for 8-12 hours, followed by two washes in 0.2xSSC, 0.1% SDS at 65°C for 20 minutes. It will be appreciated by the skilled worker that hybridization at 65°C will occur at different rates depending on a number of factors including the length and percent identity of the sequences which are hybridizing.

As used herein, an "expression control sequence" refers to polynucleotide sequences which affect the expression of coding sequences to which they are operatively linked. Expression control sequences are sequences which control the transcription, post-transcriptional events and translation of nucleic acid sequences. Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (e.g., ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance
protein secretion. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to encompass, at a minimum, any component whose presence is essential for expression, and can also encompass an additional component whose presence is advantageous, for example, leader sequences and fusion partner sequences.

As used herein, "operatively linked" or "operably linked" expression control sequences refers to a linkage in which the expression control sequence is contiguous with the gene of interest to control the gene of interest, as well as expression control sequences that act in trans or at a distance to control the gene of interest.

As used herein, a "vector" is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid," which generally refers to a circular double stranded DNA loop into which additional DNA segments may be ligated, but also includes linear double-stranded molecules such as those resulting from amplification by the polymerase chain reaction (PCR) or from treatment of a circular plasmid with a restriction enzyme. Other vectors include cosmids, bacterial artificial chromosomes (BAC) and yeast artificial chromosomes (YAC). Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome (discussed in more detail below). Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., vectors having an origin of replication which functions in the host cell). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and are thereby replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "recombinant expression vectors" (or simply "expression vectors"). The integrating cosmid vector pYUB412 is an example of a "vector".
The term "recombinant host cell" (or simply "recombinant cell" or "host cell"), as used herein, is intended to refer to a cell into which a recombinant nucleic acid such as a recombinant vector has been introduced. In some instances the word "cell" is replaced by a name specifying a type of cell. For example, a "recombinant microorganism" is a recombinant host cell that is a microorganism host cell. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "recombinant host cell," "recombinant cell," and "host cell", as used herein. A recombinant host cell may be an isolated cell or cell line grown in culture or may be a cell which resides in a living tissue or organism.

As used herein, the term "mammal" refers to any member of the taxonomic class mammalia, including placental mammals and marsupial mammals. Thus, "mammal" includes humans, primates, livestock, and laboratory mammals. Exemplary mammals include a rodent, a mouse, a rat, a rabbit, a dog, a cat, a sheep, a horse, a goat, a llama, cattle, a primate, a pig, and any other mammal. In some embodiments, the mammal is at least one of a transgenic mammal, a genetically-engineered mammal, and a cloned mammal.

C. **ESX-1 Region of M. marinum**

A nucleotide sequence from the ESX-1 region of *M. marinum* is provided in Figure 3. Figure 7 provides a map of the specific *M. marinum* nucleic acid sequence used in the Examples. Figure 21 provides the nucleotide sequences of the open reading frames present in the nucleotide sequence shown in SEQ ID NO :1. Only part of the open reading frame for MMAR5429 is present in the sequence SEQ ID NO :1. SEQ ID NO :3-34 provides the amino acid sequences of the proteins encoded by the open reading frames present in the nucleotide sequence shown in SEQ ID NO :1.
The nucleotide sequence presented in SEQ ID NO:1 comprises open reading frames for the following *M. marinum* proteins: MMAR5429 (partial open reading frame, SEQ ID NO:68), MMAR5430 (SEQ ID NO:3), MMAR5431 (SEQ ID NO:4), MMAR5432 (SEQ ID NO:5), MMAR5433 (SEQ ID NO:6), MMAR5434 (SEQ ID NO:7), MMAR5435 (SEQ ID NO:8), MMAR5436 (SEQ ID NO:9), MMAR5437 (SEQ ID NO:10), MMAR5438 (SEQ ID NO:11), MMAR5439 (SEQ ID NO:12), MMAR5440 (SEQ ID NO:13), MMAR5441 (SEQ ID NO:14), MMAR5442 (SEQ ID NO:15), MMAR5443 (SEQ ID NO:16), MMAR5444 (SEQ ID NO:17), MMAR5445 (SEQ ID NO:18), MMAR5446 (SEQ ID NO:19), MMAR5447 (SEQ ID NO:20), MMAR5448 (SEQ ID NO:21), MMAR5449 (SEQ ID NO:22), MMAR5450 (SEQ ID NO:23), MMAR5451 (SEQ ID NO:24), MMAR5452 (SEQ ID NO:25), MMAR5453 (SEQ ID NO:26), MMAR5454 (SEQ ID NO:27), MMAR5455 (SEQ ID NO:28), MMAR5456 (SEQ ID NO:29), MMAR5457 (SEQ ID NO:30), MMAR5458 (SEQ ID NO:31), MMAR5459 (SEQ ID NO:32), MMAR5460 (SEQ ID NO:33), and MMAR5461 (SEQ ID NO:34) proteins of *M. marinum*.

The *M. marinum* genes/proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, MMAR5454, MMAR5455, MMAR5456, MMAR5457, MMAR5458, MMAR5459, MMAR5460, and MMAR5461 are known by various names in the art. In most cases homologues in other mycobacterium species are named using a different nomenclature. With reference to the *M. marinum* genes/proteins, the MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, MMAR5454, MMAR5455, MMAR5456, MMAR5457, MMAR5458, MMAR5459, MMAR5460, and MMAR5461 genes/proteins the nucleotide and amino acid
sequences of the genes and proteins are provided in SEQ ID NO:68, SEQ ID NO:3-34 and SEQ ID NO: 35-67, respectively.

For the avoidance of doubt, as used herein the "CFP-10" protein is coded by esxB (or Rv3874) gene in *M. tuberculosis* and the MMAR 5449 gene in *M. marinum* (SEQ ID NO :55). As used herein the "ESAT-6" protein is coded by esxA (or Rv3875) gene in *M. tuberculosis* and the MMAR 5450 gene in *M. marinum* (SEQ ID NO :56). The names are used interchangeably herein.

**D. Recombinant Strains of *M. bovis* BCG**

This invention encompasses recombinant strains of *M. bovis* BCG. The recombinant strains comprise a heterologous amino acid sequence that comprises a plurality of open reading frames. The plurality of open reading frames comprise open reading frames that encode proteins that are the same as or are homologous to *M. marinum* proteins. In some embodiments each of the encoded proteins are at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the corresponding *M. marinum* proteins. In some embodiments all or some of the encoded proteins are identical to the *M. marinum* proteins. In some embodiments all or some of the open reading frames are functionally linked to endogenous *M. marinum* expression regulatory sequences. In some embodiments all or some of the open reading frames are functionally linked to endogenous *M. marinum* expression regulatory sequences present in the nucleotide sequence SEQ ID NO :1. In some embodiments all or some of the open reading frames are functionally linked to heterologous *M. marinum* expression regulatory sequences.

In some embodiments the plurality of open reading frames encode proteins that comprise the *M. marinum* proteins MMAR5445 (SEQ ID NO :18), MMAR5446 (SEQ ID NO :19), MMAR5447 (SEQ ID NO :20), MMAR5448 (SEQ ID NO :21), MMAR5449 (SEQ ID NO :22), MMAR5450 (SEQ ID NO :23), MMAR5451 (SEQ ID NO :24), MMAR5452
(SEQ ID NO:25), MMAR5453 (SEQ ID NO:26), and MMAR5455 (SEQ ID NO:28) or homologues thereof. In some embodiments the plurality of open reading frames further encode proteins that comprise the *M. marinum* proteins MMAR5443 (SEQ ID NO:16), MMAR5444 (SEQ ID NO:17), and MMAR5457 (SEQ ID NO:30) or a homologues thereof. In some embodiments the plurality of open reading frames further encode at least one of the *M. marinum* proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, MMAR5454, MMAR5455, MMAR5456, MMAR5457, MMAR5458, MMAR5459, MMAR5460, and MMAR5461, or homologues thereof. In some embodiments the plurality of open reading frames encode proteins that comprise the *M. marinum* proteins MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455. In some embodiments, the plurality of open reading frames further comprise an open reading frame that encodes the *M. marinum* proteins MMAR5443, MMAR5444, and MMAR5457. In some embodiments the plurality of open reading frames further comprise at least one open reading frame that encodes an *M. marinum* protein selected from MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455. In some embodiments, the plurality of open reading frames further comprise an open reading frame that encodes the *M. marinum* proteins MMAR5443, MMAR5444, and MMAR5457.
In some embodiments the plurality of open reading frames comprise open reading frames that encode *M. marinum* proteins MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5458, MMAR5459, MMAR5460, and MMAR5461. In some embodiments the recombinant *M. bovis* BCG strain comprises heterologous proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5458, MMAR5459, MMAR5460, and MMAR5461.

In some embodiments the open reading frames that encode the listed proteins have the sequence of the corresponding open reading frame presented in SEQ ID NO:35-67 (nucleotide sequence). In some embodiments the recombinant strain of *M. bovis* BCG comprises a heterologous nucleic acid sequence that hybridizes to at least one nucleic acid sequence of SEQ ID NO:35-67 (nucleotide sequence). In some embodiments the recombinant strain of *M. bovis* BCG comprises a heterologous nucleic acid sequence that is at least at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to at least one of the nucleic acid sequences of SEQ ID NO:35-67 (nucleotide sequence). In some embodiments the recombinant strain of *M. bovis* BCG comprises a heterologous nucleic acid sequence that is at least at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to the nucleic acid sequence of SEQ ID NO:1.

In some embodiments the heterologous nucleic acid sequence is present on a plasmid. In some embodiments the heterologous nucleic acid sequence is integrated into the *M. bovis* BCG chromosome. In some embodiments the recombinant *M. bovis* BCG strain comprises...
a single copy of the heterologous nucleic acid sequence integrated on its chromosome. In some embodiments the recombinant *M. bovis* BCG strain comprises multiple copies of the heterologous nucleic acid sequence integrated on its chromosome.

The recombinant *M. bovis* BCG strains of the invention may be made by any suitable method known in the art. In some embodiments an integrating shuttle vector is electroporated into a strain of *M. bovis* BCG and recombinant *M. bovis* BCG cells comprising the heterologous nucleic acid sequence integrated into the host cell chromosome are identified. In some embodiments the integrating shuttle vector is a cosmid. In some embodiments the integrating shuttle vector is pYUB412.

In some embodiments the recombinant strain of *M. bovis* BCG secretes the CFP-10 and ESAT-6 proteins of *M. marinum*.

In some embodiments the recombinant strain of *M. bovis* BCG induces a protective immune response greater than the parent *M. bovis* BCG when introduced into a mammal. In some embodiments the virulence of the recombinant strain is equal to or lower than the virulence of the parent *M. bovis* BCG in a mammal. In some embodiments the recombinant strain of *M. bovis* BCG induces a protective immune response greater than the parent *M. bovis* BCG when introduced into a mammal. In another embodiment, the virulence of the recombinant strain is equal to or lower than the virulence of the parent *M. bovis* BCG in the mammal. In some embodiments the mammal is a human.

In some embodiments the recombinant strain of *M. bovis* BCG further comprises an antibiotic marker linked to the plurality of open reading frames encoding proteins that are the same as or are homologous to *M. marinum* proteins. In some embodiments the antibiotic marker is removed from the recombinant strain of *M. bovis* BCG after integration of the plurality of open reading frames, so as the recombinant strain could be used for GMP production.
E. Methods of Making Recombinant Strains of M. bovis BCG

The invention also encompasses methods of making a recombinant strain of M. bovis BCG of the invention. In some embodiments the methods comprise providing a vector comprising a heterologous nucleic acid sequence comprising a plurality of open reading frames, introducing the vector into M. bovis BCG cells, and selecting M. bovis BCG cells that stably maintain the heterologous nucleic acid sequence comprising the plurality of open reading frames. The heterologous nucleic acid sequence may be any heterologous nucleic acid sequence described in Section D above, for example. In some embodiments the vector is an integrating vector and the method further comprises selecting M. bovis BCG cells in which the heterologous nucleic acid sequence comprising the plurality of open reading frames has integrated into the host cell chromosome.

M. bovis BCG cells used to make a recombinant strain of M. bovis BCG of the invention are preferably from a commercially available BCG strain which has been approved for use in humans such as Pasteur, Frappier, Connaught (Toronto), Tice (Chicago), RIVM, Danish 1331, Glaxo-1077, Tokyo-172 (Japan), Evans, Prague, Russia, China, Sweden, Birkhaugh, Moreau and Phipps.

F. Pharmaceutical Compositions

The invention further encompasses compositions, particularly pharmaceutical compositions, comprising a recombinant strain of M. bovis BCG according to this disclosure and a pharmaceutically-acceptable carrier. The compositions may be, for example, for use for inducing a protective immune response against M. tuberculosis in a subject and/or for use for treating an M. tuberculosis infection in a subject.

In addition to the strains, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the living vaccine into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest
In some embodiments the pharmaceutical composition is suitable for oral and/or subcutaneous and/or intradermal and/or inhalation and/or intravesicular administration.

The determination of the effective dose is well within the capability of those skilled in the art. An effective dose refers to that amount of active ingredient, i.e. the number of cells administered, which induces an immune response against *M. tuberculosis* and/or ameliorates the symptoms of *M. tuberculosis* infection. Efficacy and toxicity may be determined by standard pharmaceutical procedures in experimental animals, e.g., ED50 (the dose effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. Of course, ED50 is to be modulated according to the mammal to be treated or vaccinated. In this regard, the invention contemplates a composition suitable for human administration as well as a veterinary composition.

The invention also encompasses a vaccine comprising a recombinant strain of *M. bovis* BCG according to this disclosure and a suitable carrier. This vaccine is especially useful for preventing tuberculosis.

The invention therefore target said vaccine for use for preventing tuberculosis.

The recombinant strain of *M. bovis* BCG according to this disclosure are also useful as a carrier for the expression and presentation of foreign antigens or molecules of interest that are of therapeutic or prophylactic interest. Examples of such foreign antigens can be found in patents and patent applications U.S. Pat. No. 6,191,270 for antigen LSA3, U.S. Pat. Nos. 6,096,879 and 5,314,808 for HBV antigens, EP 201,540 for HIV-1 antigens, U.S. Pat. No. 5,986,051 for H. pylori antigens and FR 2,744,724 for P. falciparum MSP-1 antigen.
G. Kits

The invention further encompasses a kit for use for inducing a protective immune response against *M. tuberculosis* in a subject and/or for treating an *M. tuberculosis* infection in a subject. Said kit comprises a recombinant strain of *M. bovis* BCG according to this disclosure and a container. In some embodiments, said kit further comprises at least one isolated recombinant protein or peptide antigen of a mycobacterium. In some embodiments, said mycobacterium is a strain selected from *M. bovis* BCG, *M. marinum*, and *M. tuberculosis*. In some embodiments, said isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof. In some embodiments, said kit further comprises at least one subunit vaccine. In some embodiments, said kit further comprises at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, H1-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and *M. vaccae*.

H. Methods of Inducing a Protective Immune Response Against *M. tuberculosis*

The invention also encompasses methods for inducing an immune response against *M. tuberculosis* in a subject. In some embodiments, the immune response is a protective immune response. The subject may be any mammal. In some embodiments the subject is a human. Generally the methods comprise administering an effective dose of a pharmaceutical composition of this disclosure to a subject and inducing an immune response in the subject that is protective against *M. tuberculosis*. In some embodiments, the methods further comprise administering at least one isolated recombinant protein or peptide antigen of a mycobacterium to the subject. In some embodiments, said isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof. In some embodiments, the methods further comprise administering at least one mycobacterium subunit vaccine to the subject. In some embodiments the methods further comprise administering at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, H1-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and *M. vaccae* to the subject.
In methods comprising administering the pharmaceutical composition of this disclosure and a second agent to the subject, the pharmaceutical composition of this disclosure and the second agent may be administered simultaneously or at separate times. In embodiments in which the pharmaceutical composition of this disclosure and the second agent are administered at separate times the order of administration may be (1) that the pharmaceutical composition of this disclosure is administered first and the second agent is administered at a later point in time; or (2) that the second agent is administered first and the pharmaceutical composition of this disclosure is administered at a later point in time. In some embodiments the second agent is administered second and is used to boost an immune response of the subject that was raised against the recombinant strains of *M. bovis* BCG present in the pharmaceutical composition of this disclosure. In some embodiments one or both of the pharmaceutical composition of this disclosure and the second agent is administered at multiple timepoints.

In some embodiments the pharmaceutical composition is administered by an oral and/or subcutaneous and/or intradermal and/or inhalation and/or intravesicular mode.

1. **Methods ofTreating *M. tuberculosis* Infection**

The invention also encompasses methods for treating an *M. tuberculosis* infection in a subject. The subject may be any mammal. In some embodiments the subject is a human. Generally the methods comprise administering an effective dose of a pharmaceutical composition of this disclosure to a subject infected with *M. tuberculosis* and inducing an immune response in the subject that ameliorates the *M. tuberculosis* infection. In some embodiments, the *M. tuberculosis* infection is cured and the subject becomes free of *M. tuberculosis*. In some embodiments, the methods further comprise administering at least one isolated recombinant protein or peptide antigen of a mycobacterium to the subject. In some embodiments, said isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof. In some embodiments, the methods further comprise administering at least one mycobacterium subunit vaccine to the subject.
In some embodiments, the methods further comprise administering at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, H1-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and *M. vaccae* to the subject.

In methods comprising administering the pharmaceutical composition of this disclosure and a second agent to the subject, the pharmaceutical composition of this disclosure and the second agent may be administered simultaneously or at separate times. In embodiments in which the pharmaceutical composition of this disclosure and the second agent are administered at separate times the order of administration may be (1) that the pharmaceutical composition of this disclosure is administered first and the second agent is administered at a later point in time; or (2) that the second agent is administered first and the pharmaceutical composition of this disclosure is administered at a later point in time. In some embodiments the second agent is administered second and is used to boost an immune response of the subject that was raised against the recombinant strains of *M. bovis* BCG present in the pharmaceutical composition of this disclosure. In some embodiments one or both of the pharmaceutical composition of this disclosure and the second agent is administered at multiple timepoints.

In some embodiments the pharmaceutical composition is administered by an oral and/or subcutaneous and/or intradermal and/or inhalation and/or intravesicular mode.

In some embodiments the subject has an active tuberculosis infection. In some embodiments the subject has a latent tuberculosis infection.

**J. Methods for treating other pathologies**

We hereby demonstrate that the BCG::ESX-1 marinum strain has a stronger immunogenicity phenotype than normal BCG. This enhanced immunogenicity could be beneficial for the treatment of other pathologies known to be cured by BCG, such as bladder cancer.
The compositions of the invention may be, for example, for use for treating bladder cancer in a subject in need thereof.

Moreover, other diseases may be cured, such as leprosy. As a matter of fact, wild-type *Mycobacterium leprae* also carries an ESX-1 cluster, so that the presence of a heterologous ESX-1 system may reinforce the immunity generated by the vaccine.

The compositions of the invention may be, for example, for use for inducing a protective immune response against *Mycobacterium leprae* in a subject and/or for use for treating an *Mycobacterium leprae* infection or leprosy in a subject in need thereof.

**EXAMPLES**

A. Materials and Methods

1. Generation of a recombinant BCG

   a. Preparation of Cosmid DNA

A genetic construct containing the ESX-1 region of *M. marinum* integrated in the cosmid shuttle vector pYUB412 was prepared. After cloning and amplification in *E. coli* DH10B (Invitrogen Corporation, Cergy Pontoise, France) cosmid DNA was obtained. In short, the bacterial suspension was centrifuged and 250µ1 of resuspension buffer (Qiagen, Venlo, The Netherlands), Lysis buffer and K+ Acetate (pH 4.8) were added to the pellet. Upon centrifugation the pellet was homogenized with 700µ1 of isopropanol and incubated on ice for 1 hour. The suspension was centrifuged, washed with 70% EtOH and once more centrifuged. To verify the size and integrity of the pYUB cosmid carrying the ESX-1 region of *M. marinum* the Xhol (New England Biolabs, Ipswich, MA) restriction profile was compared to digestion patterns obtained earlier (R. Brosch, unpublished) on 0.8% agarose gel. Finally the DNA was dialysed to remove salts that may interfere with the electroporation.
b. **Electroporation of pYUB::ESX-1\textsuperscript{MM} into BCG.**

*M. bovis* BCG Pasteur 1173P2, held at Pasteur Institute, was grown at 37°C on Middlebrook 7H11 medium (Difco) supplemented with oleic acid-albumin-dextrose-catalase (OADC; Difco). To obtain electrocompetent cells bacteria from solid culture were transferred into 200ml of 7H9 medium complemented with ADC (Difco) and grown for 10 days. Cells were harvested collected and washed twice with H\textsubscript{2}O and once with 10% Glycerol at room temperature. The pellet was resuspended in 2 ml of 10% Glycerol. The cell suspension was mixed with the integrative ESX-1 cosmids and electroporated using a Bio-Rad gene pulser XCell at 2500 Volt. Electroporated cells were cultured overnight at 37°C and plated on 7H11 medium containing hygromycin 50µg ml\textsuperscript{-1}. Antibiotic resistant colonies were collected after about 3 weeks and analysed for the presence of the integrated cosmid using PCR.

c. **PCR on resistant colonies**

Bacilli from colonies appearing on the plates were heat-killed and denaturated at 95°C for 3 minutes to allow access to bacterial DNA. PCR reactions comprised 200ng in 2.5µl and were conducted according to the standard protocols in this laboratory [47]. The cyclic reactions were performed using a PTC-100 Programmable Thermal Controller (MJ Research, St. Bruno, Canada) starting with an initial denaturation step of 1 min at 95°C, followed by 30 cycles of 30 s at 94%, 1 min at 57°C, and 4 min at 72°C. Samples were transferred onto 0.8% Agarose gel and electrophoresed for 30 minutes at 135 V. Primers used were purchased at Invitrogen. Only PCR positive clones were subjected to immunoblotting analyses.

d. **Immunoblotting of recombinant BCG::ESX-1\textsuperscript{MM}**

Pre-cultures of cells were grown in 7H9 + ADC and transferred into Sauton Media (Institut Pasteur) at O\textsubscript{D}0\textsubscript{600nm} of 0.015. Cell lysates and supernatants were obtained from early log-phase cultures indicated by reaching a O\textsubscript{D}0\textsubscript{600nm} of 0.3 as follows: Cells were centrifuged 4000 r.p.m. for 10 minutes at 4°C and the supernatant was filtered through 0.44 µm pores.
(Sartorius Stedim Biotech, Gottingen, Germany). The whole cell extract was washed in 50 mM sodium phosphate buffer (pH 7.0) and resuspended in 1 ml of PBS and Anti-protease mix (Complete EDTA free, Roche, Basel, Switzerland) containing 500μl glass beads unwashed (Sigma). Cells were lysed by 10 min at level 30 in the Tissue Lyser II (Qiagen). Glass beads were washed down and whole cell lysate was aliquoted at -20°C. Samples were prepared with NuPage reducing agent and sample buffer (Invitrogen) and denatured at 70°C for 10 ml before loading on 12% NuPage Bis-Tris gel (Invitrogen). Gels were consequently electro-botted using Gel Transfer Stacks Nitrocellulose (Novex Life Technologies, Israel) and the iBlot apparatus (Invitrogen). Upon saturation of membranes with 5% nonfat dry milk in TBS 0.1% Tween20 (Sigma-Aldrich, St. Quentin-Fallavier, France) for 30 min the primary antibody (ESAT-6 from Antibodyshop, Lot F0703-03, CFP-10 from in house production, and Anti-GroEl (MBiosource, San Diego, CA) was added in TBS 0.1% Tween20 ON at 4°C. Upon 1h of incubation with secondary antibody (ECL Anti-Mouse IgG, Amersham Biosciences, Glattbrug, Switzerland) the blot was revealed using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific, IL, USA).

e. Immunogenicity and virulence studies

Bacterial suspensions for injection were prepared from 50 ml 7H9 + ADC (Difco) liquid cultures. Bacilli were spun down, washed twice and resuspended in PBS before brief sonification. To allow residual aggregates to settle the suspensions were kept still for 1h before aliquotation and storing at -80°C. One aliquot was used to determine the CFUs of the bacterial suspension. For immunological studies 6 week old female C3H and C57BL/6 mice (Janvier Labs, Le Genest-Saint-Isle, France) were immunized subcutaneously with 10^6 CFUs in 200μl PBS of BCG::ESX-1 MM A (n = 4), BCG::ESX-Mtb in = 4) or BCG::pYUB412 (n = 2). After 3 weeks mice were sacrificed and the splenocytes were obtained and cultured on 96 well plates (1 x 10^5 cells / well) using HL-1 medium complemented with β-mercaptoethanol and 1% Streptomycin and Penicillin (Invitrogen) in the presence of various mycobacterial antigens. After 72h of incubation IFN-γ was quantified in culture supernatants using ELISA according to standard protocols.
(capture and detection antibodies and streptavidin-HRP were purchased from BD, le Pont de Claix, France). For virulence studies six-week-old SCID mice (n = 4 per group) (Janvier) were infected via the lateral tail vein with 10^6 CFUs. Upon sacrifice after 3 weeks of inoculation lungs and spleens were homogenized using Tissue Lyser II (Qiagen) and plated on previously prepared 7H11 agar plates in serial dilutions and incubated at 37°C. After 2-3 weeks CFU were counted.

2. Generation of an anti-CFP-10 and anti-EspA T-cell Hybridoma

a. T-cell epitope mapping of EspA

A Pepscan® composed of 77 peptides of 15-mers offset by 5 amino acids scanning the complete sequence of the EspA protein (Mimotopes, Claiton, Australia) was used to identify the immunogenic regions. Female C57BL/6 (H-2b), BALB/c (H-2d) or C3H (H-2K) mice (Janvier, France) were immunized subcutaneously with 3 x 106 CFUs of M. tuberculosis H37Rv strain in 200µl (n = 2). One mouse of each strain was used as non-infected control. The MHC-II H-2K-restricted immunodominant T-cell epitope in the CFP-10:1-25 peptide has been described earlier [75]. Pooled splenocytes were harvested 2 weeks post immunization and cultured at 106 cells/well in 96 well plates in RPMI 1640 Glutamax complemented with 5% FCS and 1% Penicillin-Streptomycin as duplicates in the presence of x-to-y µg/ml of individual peptides from the EspA pepscan. After 72h of incubation at 37°C and 5% CO2 IFN-γ was quantified in the culture supernatants using ELISA. The identified EspA immunodominant sequences were then synthetized by Polypeptide. The identified EspA immunodominant sequences were then synthetized by Polypeptide. (Strasbourg, France).

b. Generation of the T-cell hybridoma

C3H mice (n = 2) were immunized subcutaneously with 50µg of the synthetized peptides in 200µl in Incomplete Freuds Adjuvant (Sigma Aldrich). Two weeks post-infection splenocytes were removed, washed, pooled and cultured using complemented RPMI with 10% FCS. To verify successful immunization splenocytes were distributed in a 96 well plate (1x106 cells/well in 100µl) and stimulated by various concentrations of the homologous peptides ranging from 5µg/ml to 0.03 µg/ml. Culture supernatants were
assessed for IFN-γ secretion after 72h incubation at 37°C by ELISA. The remaining splenocytes were incubated for 4 days with the homologous peptides at 1(^g/ml in 25 cm3 flasks. The fusion partner BW5147 thymoma cells, lacking Hypoxanthine-Guanine PhosphoRibosyl Transferase activity, were cultured from stocks held at Pasteur Institut in standard medium and were centrifuged, washed twice, and counted. The re-stimulated splenocytes were enriched in live cells using 3 ml of Ficoll Hypaque solution (Lympholyte-M, Cedarlane, TEBU BIO, Le Perray-en-Yvelines, France), washed and counted. BW5147 cells and splenocytes were assembled at 1:1 ratio and washed twice at room temperature. 0.5 ml of polyethylene glycol 1500 were slowly added to the cell pellet, followed by a wash at 500 r.p.m. for 4 min. Subsequently, the mix was resuspended carefully by 8 ml of complemented RPMI before adding 32 ml of medium containing 20% FCS. Upon an incubation period of 4 h at 37°C, feeder splenocytes were obtained from naive C3H mice and added at 1 x 10^5 cells/ml and the suspension homogenized. Cells were distributed in 96 well plates at 100µl / well for incubation overnight. The next day, the HAT (Hypoxantine, Aminopterine, Thymidine) medium, blocking purine synthesis and thus eliminating the non-fused BW5147 cells, was added at IX, thereby allowing selection of only fused BW5147-splenocytes hybrids in which the purine synthesis pathway is restored. At day 7, 100µl of HT- complemented RPMI was added to each well and carefully monitored for appearing growth of hybridomas.

c. Detection, quantification, and sensitivity

The hybridomas were diluted using 100µl of complemented RPMI with 20% FCS and 1x HT (Hypoxantine, Thymidine) before transfer into 24 well plates containing 500µl of medium. To screen and select the hybridomas specific for the T-cell epitope, naive splenocytes were loaded with 10µg/ml of peptide in RPMI 20% FCS. The next day the hybridomas were added and the supernatant was examined for IL-2 after overnight incubation using ELISA. The positive hybridomas were then subjected for further analyses to specify at molecular level their sensitivity. To this end, naive splenocytes were incubated with the peptide as described above, in two-fold dilutions from 5µM to 3.9nM. After overnight incubation the hybridomas were added and IL-2 was quantified. The most
sensitive four hybridomes were named, amplified in 25m3 flasks and stored at -80°C. To determine the MHC restriction of the selected T-cell hybridomas, L fibroblasts, stably transfected with I-AK or I-EK were loaded with CFP-10:1.1-25 peptide or a negative control and co-cultured with the hybridomas. IL-2 was quantified in the co-culture supernatants after overnight incubation.

d. Dendritic cell infection assay

Bone marrow-derived dendritic cells (BM-DC) were obtained from femurs of female 6-week-old C57BL/6 or C3H mice depending on the restriction of the epitopes to be studied. Cells were cultured in petri dishes and differentiated into DC in the presence of 10ng/ml GM-CSF (prepared at Institut Pasteur) BM-DC were harvested, washed and distributed at 5 x 105 cells / well in 100μl in a 96 well plate. The different mycobacterial suspensions were added at two fold dilutions starting at an M.O.I. of 4 in RPMI without antibiotics. Upon 24 h incubation 1 x 105 cells of hybridomas per well in 100μl were added and stored at 37°C and 5% C02. The next day IL-2 was quantified in the supernatant using ELISA.

e. FRET assay and Flow Cytometry

Human pro-monocytic cell line THP-1 was cultured in RPMI at 10% FCS for one week. Cells were collected, washed and distributed in a 24 well plate at 5 x 105 cells / well in 2 ml. 20 ng/ml of PMA (Pharboil 12-myristate 13-acetate, Sigma Aldrich) was added to differentiate the cells into monocytes and incubated for 72h. Mycobacterial strains were grown until mid-log phase in Dubos Broth (Dico) and the OD600 was determined. Differentiated THP-1 were washed twice and infected with the bacterial suspensions at an M.O.I. of 1. After 2h of incubation the supernatant was removed and cells were washed three times with PBS. Fresh medium was added and cells were stored at 37°C 5% C02 for 4 days. Next, cells were incubated for 2h at RT in dark with the CCF-4 mix consisting of 50μM CCF-4 substrate (Invitrogen) in EM buffer (120 mM NaCl, 7 mM KC1, 1.8 mM CaCl2, 0.8 mM MgCl, 1.5 mM glucose and 25 mM Hepes at pH 7.3). Next, an appropriate dilution of anti-CD11b- APC (BD) was added to the wells and incubated at RT in dark for 30min. Marked cells were washed twice and fixed with 4% Paraformadehyde and kept at
4°C overnight. The following day the suspension was filtered using 100μm filters (BD) and subjected to Flowcytometry analyses using the CyAn® system and the Summit Software (Beckman Coulter, Villepinte, France). Phagosomal rupture was observed when the CCF-4 signal became blue and when CCF-4 was not cleaved the signal remained green. Data were analysed with FlowJo software (Treestar, OR).

f. Statistical Analyses

GraphPad Prism software (GraphPad Software, La Jolla, CA) was used to perform statistical analyses. The student's t and Mann-Whitney U tests were employed based on normal or skewed distribution of the data.

10 B. Results

Example 1: Complementation of BCG with the ESX-1 region of *M. marinum*

In order to create a recombinant BCG expressing the ESX-1 region of *M. marinum* the genetic material was introduced into BCG by electroporation. The RD1 region deletes nine coding sequences in the ESX-1 locus in BCG (Figure 6) [64]. The integrating cosmid vector pYUB412 [76] carried the entire ESX-1 locus of *M. marinum*. The vector sequence is provided in SEQ ID NO 2 and the inserted *M. marinum* sequence in SEQ ID NO :1. A map of the vector is provided in Figure 3. Figure 4 shows an alignment of ESX-1 loci from *Mycobacterium tuberculosis* and *Mycobacterium marinum* species. Figures 5 indicate the open reading frames encoded by the inserted fragment. To verify that the integrating cosmid harbours the desired ESX-1 locus a restriction profile analysis was performed using the Xhol restriction enzyme. In Figure 7A the PCR amplified digestion profile is shown and compared to in silico prediction and earlier restriction patterns done by the laboratory. Upon electroporation of electrocompetent BCG cells the cosmid vector integrates firmly into the BCG genome. The cosmid encodes a hygromycin cassette allowing to select resistant cones on 7H11 Hygromycin agars. Seventy clones were selected for PCR analyses. Three pairs of primers specific for genes within the ESX-1 region were used to ensure the presence of the *M. marinum* fragment in the BCG::ESX-1<sup>MM</sup> strain: MMAR_5441 (homologue of espGl), MMAR_5446 (homologue of eccCbl), and
MMAR_5452 (homologue of eccDl). Two clones, named A and B, were positive for all primers (Figure 7B), and were chosen for further analyses. One of the strains was designated ESX-1-Mar and deposited at the Collection Nationale de Cultures de Microorganisms (CNCM), 25, Rue de Docteur Roux, F-75724 Paris Cedex 15 FRANCE on June 3, 2014 under the Reference Number 1-4858.

Example 2: Expression of ESX-1 proteins by recombinant BCG::ESX-1MM clone A

Next, the functionality of the ESX-1 secretion machinery was investigated. To this end a monoclonal antibody against ESAT-6 and a polyclonal serum against CFP-10 were used. As shown in figure 8, substantial homologies exist between Mtb and M. marinum ESX-1 genes. In particular, 91 and 97% sequence homologies are detected between the genes coding for ESAT-6 and CFP-10, respectively, leading to high sequence similarities at protein level. Therefore, antibodies specific to Mtb ESAT-6 or CFP-10 most likely will cross recognize their homologues in M. marinum [18]. To detect secreted ESX-1 proteins, cell lysates and supernatants from mid-logIO phase cultures were prepared and stained with the appropriate dilutions of antibodies. Figure 9 shows that only clone A expresses ESAT-6 and CFP-10 suggesting that the esxAB genes were heterologously expressed and exported via a functioning type VII secretion machinery. Clone B, tested positive on PCR screens, neither produces nor secretes these proteins, possibly due to incorrect integration of the cosmid into the BCG genome, or disintegration after some replication cycles. The integrated ESX-1 region contains more genes than are deleted in the RD1 region from BCG. Based on these results it cannot be inferred which proteins, apart from the RD1 deleted proteins, are BCG or M. marinum derived. Conventionally, an antibody staining the intracellular chaperone pro-tein GroEl-2 (Rv0440) is used to ensure that the presence of secreted proteins in the supernatant is due to actual secretion and not lysis of the cells. GroEl-2 was found in the super-natants suggesting that lysis took place and the integration of ESX-1 of M. marinum is toxic to the BCG cells after a number of replication cycles. Accordingly, different immunological tools were used to investigate the secretion of ESX-1 substrates from this clone in more detail.
Example 3: Search for Immunodominant Epitope in EspA

Apart from the virulence factors ESAT-6 and CFP-10, the ESX-1 region secretes other proteins encoded outside the ESX-1 locus, the ESX-1 secretion-associated proteins (EspS). In order to reveal the molecular mechanisms by which these proteins are involved in secretion and immunogenicity a powerful tool is to generate T-cell hybridomas which harbour T-cell receptors (TCR) specific to immunodominant epitopes of the secreted proteins. Upon secretion into the host cell phagosome these proteins are processed and presented by Major Histocompatibility Complex (MHC) class II molecules at the cell surface and recognized by the specific TCRs of the hybridoma. The secretion of even minute levels of proteins inside infected cells can be detected by quantifying the Interleukin-2 produced by the hybridoma using ELISA. The lab has previously generated T-cell hybridomas specific to the secreted proteins ESAT-6, EspC and to Ag85A. The latter is secreted via the conventional Sec pathway and presents therefore an internal positive control. The goal is to expand this library with hybridomas specific for EspA and CFP-10.

To this end the immunodominant regions of the EspA protein were identified by epitope mapping, while the immunodominant epitope of CFP-10 has been previously described [75]. For EspA, splenocytes from M×-infected mice were stimulated with each of the 77 individual 15-mer peptides scanning the entire sequence of EspA. Next, the IFN-γ response was quantified to map the immunodominant epitopes of EspA. Splenocytes from Mtb-immunized C57BL/6 (H-2b), BALB/C (H-2d) (data not shown) or C3H (H-2K) mice were screened and two immunodominant regions in C3H mice were identified (Figure 13). Only in C3H mice two different immunodominant region for EspA could be identified. C3H mice were subsequently immunized by a mixture of CFP-10:1 1-25, EspA:1-15 and EspA:324-340 synthetic peptides containing these immunodominant epitopes in order to isolate from the same mice, splenocytes harbouring T-cell receptors specific to the three different epitopes. T-cell hybridomas were generated by fusing the restimulated splenocytes with BW5147 thymoma cells, as described in Methods. In a first attempt, only T-cell hybridomas harbouring TCRs specific for CFP-10:11-25 appeared after fusion, the most probably due to the strong immunodominance of the latter over the EspA epitopes.
Example 4: Generation of XE12, an anti-CFP-10:11-25 T-cell Hybridoma

The anti-CFP-10:11-25 hybridomas were named according to their position on the 96 well plates. In a first screen, all hybridoma candidates were tested for their specificity for the CFP-10:11-25 peptide when co-cultured with naïve splenocytes that were presenting the homologous peptide. Of the 38 candidates that had fused 18 were selected for further analysis. To determine their specificity in more detail the 18 candidates were co-cultured with BM-DC either loaded with homologous peptide, a negative control peptide, or infected with Mtb H37Rv WT or AESX-1. Two hybridomas, XE12 and YH6 appeared to produce IL-2 in response to splenocytes infected with H37Rv WT or stimulated with the peptide but not to H37Rv Δ ESX-1 which is deleted for the ESX-1 region and hence for the CFP-10 gene (Figure 11). This is an important step as the majority of the hybridomas harbour TCRs specific for the epitope derived from the peptide but not the ordinary epitope displayed following infection with Mtb. Using this essay it was possible to select the two hyridomas, which recognize both the synthetic peptide and the native protein secreted following infection. The next step was to characterize the sensitivity of the hybridomas to determine the smallest concentration to which the hybridoma responds by IL-2 production. Here, BM-DC were cultured with different concentrations of the peptide and IL-2 was quantified after 24h incubation with the two hybridoma candidates. XE12 appeared to be highly sensitive already responding at peptide concentrations of 15 nM.

Example 5: Analysis of ESX-1 protein secretion and antigenic presentation in infected dentritic cells

The Western blot analysis (Figure 9) showed that cell lysates contain large amounts of ESAT-6 and CFP-10 indicating that the protein is properly expressed. Likewise, the proteins could be detected in the supernatant proving that the secretion machinery encoded in the ESX-1 region is functioning. To further characterize the secretion of virulence proteins bone marrow derived dentritic cells (BM-DC) from C3H mice were infected with the two BCG mutants, BCG::ESX-1MM A and B and detected levels of secreted ESX-1 proteins using the T-cell hybridomas. We were particularly interested in
the recombinant clone BCG::ESX-1MM B that was PCR positive but showed no response on Western blotting. Notably, also the NB1 I Anti-ESAT-6 T-cell hybridoma showed no secretion for this clone as opposed to the Western Blotting positive clone, BCG::ESX-1MM A and the positive control, H37Rv WT Mtb, confirming the correct integration and functioning of the ESX-1 region in the A clone (Figure 13). Likewise, using the XE12 anti-CFP-10 hybridoma it was possible to detect secretion in the A strain compared to the B strain which further corroborates functioning integration of the genetic addition in the A strain (data not shown). Next, the protein EspC which is encoded outside the ESX-1 locus was tested for secretion. This is therefore a BCG protein (Mb3645c) that needs to be secreted via the added M. marinum ESX-1 encoded Type VII secretion machinery. The amino acid match between the Mtb and BCG proteins was retrieved using Genolist (http://genolist.pasteur.fr/- BoviList/genome.cgi) which showed 100% homology for EspC:40-54, the immunogenic epitope recognized by I-Ab- restricted IF1 T-cell hybridoma. Substantial secretion of EspC by BCG::ESX-1 Mtb and varying levels of the different other strains was found. Notably the BCG::ESX-1MM A had a higher secretion than BCG::ESX-1MM B. Lastly, the DEI0 Anti-Ag85A hybridoma is used as a control for this assay as it detects a protein that is secreted via the conventional Sec pathway. Taken together these data prove the correct function and secretion, processing and antigen presentation of ESX-1 encoded proteins ESAT-6 and CFP-10 and confirms the results obtained by immunoblotting.

**Example 6: Immunogenicity**

The two substantial characteristics a vaccine candidate must possess are its safety and its immunogenicity. While we ascertain the safety of a vaccine in severely immunosuppressed mice (SCID) we evaluated the immunogenicity of clone A in immunocompetent C3H mice in a splenocyte stimulation assay. Proper secretion of proteins is a prerequisite to elicit strong and specific T-cell responses [64]. Hence this method is a validated tool to confirm expression and secretion of the virulent factors encoded in the ESX-1 region. To this end, groups of C3H mice (H-2K) were subcutaneously immunized with BCG::pYUB, as negative control, or BCG::ESX-1Mtb, as positive control, or the BCG::ESX-1MM A strain
(n = 2 / group). Three weeks later, T-cell immunity was assessed by comparing antigen-specific IFN-γ release by splenocytes. Concanavalin A was used to control the culturing and Purified Protein Derivative (PPD), a mix of Mtb proteins, was employed to ensure successful immunization. The results show that both the peptide CFP-10:1 1-25 as well as the recombinant protein CFP-10 cause IFN-γ release from splenocytes of immunized mice of the two recombinant strains (Fig 14). The BCG::ESX-1 Mtb and BCG::ESX-1MM A exert an immunogenic response at similar level. The fact that even BCG::pYUB immunized mice react to CFP-10 stimulation, which is not encoded in this strain, is due to presence of ESX-1 components in these bacteria. PCR analysis, subsequent to these results, confirmed that the BCG::pYUB strain used in this experiment was expressing ESX-1 genes. Hence, these experiments need to be repeated using a different negative control. The use of C3H allowed us to only study for peptides restricted by MHC-II of the H-2K haplotype, further experiments to study the immunogenic potential of ESAT-6 stimulation assays on C57BL/6 mice (H-2b restricted) are planned. Taken together these results support that the ESX-1 region has integrated and functions properly to secrete the encoded proteins, such as CFP-10.

**Example 7: Virulence**

The second substantial feature of a vaccine is its safety. BCG is the most used vaccine worldwide with an astonishing safety record [59]. A potential vaccine candidate needs to as safe as BCG in order to advance to the clinical evaluation. To investigate the virulence of the A strain we infected SCID mice (n = 5 / group) intravenously with BCG::ESX-1 MM A, BCG::ESX-1 Mtb and BCG::pYUB. The lungs and spleen were retrieved after 3 weeks and plated on agars to count the Colony Forming Units. While the bacteria injected were comparable for the A strain and BCG::pYUB (CFUs on day 1) the dose of BCG::ESX-1 Mtb appears to be too little (Figure 15). The recorded virulence of this strain, however, was confirmed as on day 21 most CFUs were counted in mice injected with this strain. Interestingly, the A strain showed lower CFUs on day 21 than BCG::pYUB, suggesting that the strain undergoes a self limiting process such as lysis. This would be in line with our results obtained by immunoblotting where we continued to detect the intracellular
chaperone GroEl-2 in the supernatant. It is possible that the genomic ESX-1 region from M. marinum, which is more distant than the respective region in Mtb, is not tolerated over a number of replication cycles in BCG and hence leads to lysis. While this could constitute a desirable feature of a vaccine, the exact process needs to be elucidated to be able to estimate the kinetics of antigenic material being presented to host cells to avoid exacerbated host immune responses [77,78].

Example 8: Induction of phagosomal rupture

*Mtb* has long been thought to reside within the macrophage phagosome. This paradigm has only recently been challenged [79]. *Mtb*-mediated phagosomal rupture, possibly resulting in the translocation of the bacteria from the membrane-bound compartment into the host cell cytosol, correlates with virulence and pathogenicity [80]. Interestingly, this phagosomal rupture is dependant on ESAT-6 secretion and correct functioning of the ESX-1 T7SS [81]. To further characterize the pathogenic potential of the BCG:ESX-1-MM A strain in terms of its ability to induce phagosomal rupture we used a fluorescence-based method where the cytoplasm of THP-1 cells, a human monocytic cell line, is loaded with a chemical probe that is sensitive to fluorescence resonance energy transfer (FRET) measurements [82]. This chemical probe, CCF-4, can be cleaved by the BlaC mediated β-lactamase intrinsic activity of *Mtb* and BCG [83]. In THP-1 cells infected with the different mycobacterial strains, CCF-4 diffuses freely across the cytoplasmic membrane wherein it remains trapped. If the bacteria succeed in induction of phagosomal rupture, the CCF-4 is cleaved by the intrinsic β-lactamase activity of the bacteria which leads to a switch of the FRET signal from 535 nm (green) to 450 nm (blue) upon 405 nm excitation. Thus only bacteria with a functioning, ESAT-6 secreting ESX-1 system will show a FRET signal switch towards blue. The change in fluorescence can be detected using flowcytometry where gates on CD1lb positive cells are used to exclude bacteria and cell debris in the flow cytometry analysis. The results show that all strains owning a functioning ESX-1 region such as *Mtb* induce phagosomal rupture, as evident in the FRET signal shift towards blue (Figure 16A). Interestingly, the BCG::ESX-1MM A strain similarly induces a strong shift towards blue, confirming its ability to export the ESX-1 proteins and therefore the phagosomal rupture
We also tested whether the signal for the BCG::ESX-1MM B strain, that was PCR positive for ESX-1 gene primers and showed no ESAT-6/CFP-10 expression on immunoblotting, would cause a FRET signal shift towards blue. Indeed, the signal remained green similar to the negative controls BCG::pYUB and non-infected cells (Fig 16B). These results may underline the pathogenic potential of the BCG::ESX-1MM A strain in addition to the correct and integrative functioning of the added genetic ESX-1 region.

Example 9: Evaluation of the virulence of strain *Mycobacterium bovis* BCG::ESX1M*^*marinum (also named BCG::RD1M*^*marinum or BCG::RD1-marinum) in comparison with *Mycobacterium tuberculosis* H37Rv, *Mycobacterium bovis* BCG Pasteur, and *Mycobacterium bovis* BCG::ESX-1-mtb (also named BCG::RD1-2F9) in severe combined immune-deficient (SCID) mice.

Two experiments of virulence test were undertaken for which the results are shown on figures 17 - 19. In both experiments, injection of groups of mice were done intravenously (iv) and the infectious dose was determined at day 1 (mice were euthanized at day 1). The infectious dose recovered from the lungs and spleens corresponds to the mean CFU from 2 mice (a total of 12 mice was infected).

Cfu values in mice of Experiment 1 (Figure 17):

<table>
<thead>
<tr>
<th>Strains</th>
<th>cfu/lungs</th>
<th>cfu/spleen</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCG::RD1-marinum</td>
<td>9375</td>
<td>31250</td>
</tr>
<tr>
<td>BCG Pasteur</td>
<td>6250</td>
<td>25000</td>
</tr>
<tr>
<td>H37Rv</td>
<td>1250</td>
<td>6250</td>
</tr>
</tbody>
</table>
Cfu values in mice of Experiment 2 (Figure 18):

<table>
<thead>
<tr>
<th></th>
<th>CFU lungs at day 1</th>
<th>CFU spleen at day 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCG Pasteur</td>
<td>5.31E+04</td>
<td>1.17E+05</td>
</tr>
<tr>
<td>BCG::RD1-Marinum</td>
<td>4.06E+04</td>
<td>2.08E+05</td>
</tr>
<tr>
<td>BCG::RD1-2F9</td>
<td>2.50E+04</td>
<td>9.35E+04</td>
</tr>
<tr>
<td>H37Rv</td>
<td>2.92E+04</td>
<td>9.37E+04</td>
</tr>
</tbody>
</table>

*M. microti* ATCC 35782 was purchased from the American Type Culture Collection (designation TMC 1608)

*M. microti* OV254 was originally isolated from voles in the United Kingdom in the 1930s (Wells, A. Q. 1937. Tuberculosis in wild voles. Lancet i:1221).


The *M. microti* strains were tested in the same model because some of them were meant to be vaccine strains and some of them were isolated from infected humans. These strains were control strains that were used in the same experimental setting, but they do not intervene with the testing of the BCG and recombinant BCG strains. Only the strains marked with an arrow, ie *M. microti* 2272-94, BCG::RD1* M.marmum* and BCG Pasteur (Figure 17), are the main strains of interest for the vaccine testing purpose.
The virulence test corresponds to a survival experiment (time to reach the humane endpoint, which corresponds to a 20% weight loss for the mice).

For survival experiments, groups of 10 female SCID mice were subjected to intravenous challenge with a single dose of bacteria (nominally $10^6$ CFU/mouse) of the different strains and weight was controlled regularly. Upon weight loss of 20% mice were killed and the date reported in the graph.

Virulence data are shown in Figure 17 (experiment 1) and Figures 18 & 19 (experiment 2).

As a conclusion, SCID mice infected with the BCG::ESX1$_{M.~marinum}$ strain survived much longer than BCG::ESX1-l-mtb infected mice (until reaching the humane endpoint of loss of 20% bodyweight). Survival was very similar to mice infected with the BCG Pasteur parental strain.

**Example 10:** Virulence tests of BCG::RD1$_{M.~marinum}$ in guinea pigs infected by aerosol, compared to BCG::RD1$_{M.~tuberculosis}$ and $M.~tuberculosis$.

**Material & Methods :**

Animals used in this experiment are Hartley guinea pigs (female) purchased from Charles River, 69592 LArbresle, France, up to 200g at arrival in the animal house.

Twelve days after arrival, the animals were infected via aerosol using commercially available sterile nebulizers (atomizers) and bacterial solutions made from aliquots of a frozen (-80°) stock culture of Mtb H37Rv (concentration used 2.5x10$^{5}$ cfu/ml), BCG::RD1-marinum (also named BCG::ESX1$_{M.~marinum}$) (concentration used 3.5x10$^{5}$ cfu/ml), and BCG::RD1-Mtb (concentration used 2.75x10$^{5}$ cfu/ml). The concentration of live bacteria was verified a second time before infection, by plating out various dilutions of an aliquot of the stock culture.
Before the infection, the weight of the animals is determined (~ in average 250g). Then, animals are transferred into 500 ml plastic containers (bottles) with their nose towards the 1.5 mm opening of the container. Four such containers are then placed into each of the two aerosol chambers (340 mm long, 170 mm high and 180 mm wide), which are located in an airtight isolator that is kept under negative pressure (and which can be connected to the isolator containing the animal cages).

For each round of infection, 6ml at the above mentioned concentration (dilution in PBS) are used to be aerosolized at 1.6 bar into each of the aerosol chambers. This procedure lasts 15 minutes until the liquid is completely aerosolized.

After infection, animals are transferred back to their cages, and the procedure is repeated until all animals have been aerosolized.

Two control animals are killed at day 1 (using C02) after the aerosolization and their lungs are removed, homogenized and plated in different dilutions onto Middlebrook 7H11 plates containing the supplement Middlebrook OADC (Oleic Albumin Dextrose Catalase) in duplicate (1 set without antibiotics and 1 set with ampicilline as lungs of animals are often containing other bacteria) to estimate the infectious dose obtained in the lungs.

The four remaining animals per group are then kept for 6 weeks, and then killed by using increasing concentrations of C02 (according to the recommendasions of the ethics committee).

Then lungs and spleens are removed, and transferred into specific containers used for grinding the organs in the BSL3 laboratory to obtain a homogenate that is used for preparing dilutions that are plated onto 7H11 plates containing the supplement Middlebrook OADC (Oleic Albumin Dextrose Catalase) in duplicate (1 set without antibiotics and 1 set with appropriate antibiotics mix as lungs of animals are often containing other bacteria). Homogenates are kept at -80° until CFU counts have been done.
Results:

Six guinea pigs per group were infected with the different mycobacterium strains:
BCG :: RD\textsubscript{1\textit{M. marinum}} (also named BCG//RD1-Marinum), BCG :: RD\textsubscript{1\textit{M. tuberculosis}} (also named BCG :: RD1-2F9) and \textit{M. tuberculosis} H37Rv.

The infectious doses at day 1 and 6 weeks after the aerosol infection are shown in Table below and in Figure 23A (day 1) and Figure 23B (6 weeks).

<table>
<thead>
<tr>
<th></th>
<th>CFU/lung at day 1 after infection</th>
<th>CFU/lung at 6 weeks post infection</th>
<th>CFU/spleen at 6 weeks post infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCG::RD1-Marinum</td>
<td>208</td>
<td>below detection level</td>
<td>below detection level</td>
</tr>
<tr>
<td>BCG::RD1-2F9</td>
<td>212</td>
<td>1.25E+05</td>
<td>6.89E+04</td>
</tr>
<tr>
<td>H37Rv</td>
<td>293</td>
<td>2.63E+06</td>
<td>7.59E+05</td>
</tr>
</tbody>
</table>

From previous experiments it is known that the above described aerosol procedure results in about 100 - 500 CFU per lung at day 1.

As shown in Figure 23B and Table above BCG :: RD\textsubscript{1\textit{M. marinum}} was found below detection level at 6 weeks, suggesting that this recombinant strain is strongly attenuated in this sensitive animal model.

From a general aspect of the organs of guinea pigs infected with BCG::ESX-marinum, we conclude that they are not affected by pathology.
Example 11: Induction of CD4+ T-cell responses specific to ESX-1 or ESX-5 substrates in C57BL/6 (H-2b) and in C3H (H-2k) mice immunized with BCG::RD1 <em>M. marinum</em>

Mice (<em>n</em> = 3/group) were immunized subcutaneously by 1 x 10^6 CFU of mycobacterial strains of vaccine interest. Three weeks post immunization, splenocytes were stimulated in vitro with different antigens to measure the CD4+ T-cell responses.

These data showed that BCG::ESX-1<sub>M. marinum</sub> induced strong T cell responses against ESAT-6 and CFP-10 (Rv3875 and Rv3874), similar to <em>M. tuberculosis</em> and BCG::ESX-1-mtb, and the same was true for the ESX-1 associated antigen EspC (Rv3615c), which was also strongly induced in the BCG::ESX-1 complemented BCG strains. Results are presented in Figures 20 and 21.

In contrast, no response was obtained for EspA (Rv3616c), which already in <em>M. tuberculosis</em> produced a weaker response than the gene product of its neighboring - gene EspC (results on Figures 20 and 21).

Example 12: New vaccination experiment in which C57BL/6 mice were vaccinated with different vaccine candidates and BCG Pasteur-vector as the comparator strain.

Mice (<em>n</em> = 5/group) were immunized subcutaneously by 1 x 10^6 CFU of mycobacterial strains of vaccine interest. Four weeks post immunization, mice were challenged by aerosol with <em>M. tuberculosis</em> H37Rv virulent strain in order to deliver approximately 150 CFU/mouse.

Mycobacterial loads in the spleen and lungs were determined at four weeks post challenge. As seen from these experiments, BCG::ESX-1<sub>M. marinum</sub> similarly to BCG::ESX-1-mtb and an attenuated <em>M. tuberculosis</em> ESX-5-ko vaccine (named MtbAppe25-pel9) induced
significantly better protection in lungs and spleens against an aerosol challenge with virulent \textit{M. tuberculosis} H37Rv. Results are shown in Figure 22.

Taken together, these data show that the BCG: \textit{ESX-1} \textit{M. marinum} strain has a very good vaccine profile.

**C. Discussion**

In this study a TB vaccine candidate was created based on a recombinant BCG encoding the ESX-1 secretion system of \textit{M. marinum} (BCG: \textit{ESX-1}MM, also named BCG: \textit{ESX-1} \textit{M. marinum}, or BCG: \textit{RD1}-Marinum). \textit{In vitro} and \textit{in vivo} screens suggest that the BCG: \textit{ESX-1}MM A strain is immunogenic due to export of the potent T-cell antigens ESAT-6 and CFP-10.

The BCG Pasteur strain was first complemented with the ESX-1 region of \textit{M. marinum}. While the RD1 deletion in BCG includes only some of the core ESX-1 proteins, the genetic addition spanned the entire ESX-1 locus of \textit{M. marinum}.

While the genomes of BCG and Mycobacterium tuberculosis (\textit{Mtb}) are highly conserved (>99.9 \%) as reflected in the good tolerance of the \textit{Mtb} ESX-1 genetic addition into BCG we here introduced a genomic region from a more distantly related mycobacterium, \textit{M. marinum}, that encodes many more non-synonymous Single Nucleotide Polymorphisms and hence amino acid differences than its \textit{Mtb} homologues. We could ascertain the establishment and functioning of the genetic addition in the BCG genome using different immunoblotting and antigen presenting assays. Both Western blot and T-cell hybridoma assays detected secretion of ESX-1 encoded proteins. Notably, these results indicate that the \textit{M. marinum} genes share enough similarities with their \textit{Mtb} counterparts to be tolerated in the BCG genome and recognized by the BCG promoters for transcription, despite sequence differences of up to 50\% (Rv3876 and its homologue MMAR_5451). This creates a range of possibilities to complement more virulence factors, such as the ESX-5 system of \textit{M. marinum} [36] to achieve the optimal level of immunogenicity while maintaining low levels of virulence for the needed vaccine properties.
To get a detailed picture of the induced immune response and immunogenicity properties of the BCG::ESX-1MM A clone a major step would be the identification of further T7SS substrates and their function. Little is known about most of the substrates but the role of the ESAT-6/CFP-10 pair has recently been reviewed [17]. They are associated with suppression of a pro-inflammatory response, necrosis, apoptosis, membrano-lysis and cytolysis. It remains questionable whether ESAT-6 and CFP-10 actually are the effectors causing these effects or rather building blocks in the process of secretion to facilitate the way for other effectors [84]. This is an intense area of research as it is unclear if other substrates apart from the ESAT-6/CFP-10 family exist. The PE/PPE proteins are heavily secreted by the T7SS, mainly ESX-5, while having many paralogues in the genome (up to 10 % of Mtb genome encodes proteins of the PE/PPE family), suggesting that there are many more substrates than initially thought [85]. The ESX-1 region, for instance, is required to arrest phagosome maturation in macrophages yet the known ESX-1 substrates are not involved in this process proposing other secreted proteins via this pathway [86].

Likewise, the fact that T7SS are found in other gram-positive bacteria, among which some non-pathogenic and non-virulent, implies that the ESX-1 system fulfils other functions apart from virulence. As such the T7SS in Listeria monocytogenes is not required for growth [87], suggesting that also in Mtb, there could be other T7SS substrates not yet identified that are responsible for virulence. In support of this the C-terminal sequence of CFP-10 was shown to be portable when attached to yeast ubiquitin underpinning the notion that other substrates can be transported via the ESX-1 secretion machinery [53]. Also, the ESX-1 substrate EspA lacks this specific C-terminal sequence proposing that it could bind to another protein, such as EspC, and be transported together. Based on this the range of potential substrates widens even further.

In this study we made use of a newly developed assay using FRET to investigate the potential of BCG::ESX-1MM A to induce phagosomal rupture [80]. Our vaccine candidate clone demonstrated its ability of phagosomal escape compared to its control partner, the Western blot negative clone B. Arising evidence that mycobacteria are able to leave the phagosome has challenged the long standing paradigm that mycobacteria reside inside the
phagosome postulating that they are able to induce vacuolar rupture to access the cytosol [79]. Presence of mycobacterial antigens in the cytosol would offer an explanation for the strong CD-8+ T-cell response that *Mtb* antigens are able to elicit, as antigens could get access to the cytosolic MHC-I presentation machinery only upon translocation [88]. It is well established that the membrane bound phagosome of the macrophage is the *in vivo* niche of *Mtb* [20] where they avoid phagolysosome fusion to circumvent the host immune response [9]. How exactly the bacilli interacts with the macrophage remains to be elucidated [20] however it was shown to be a very dynamic relationship [89]. While electron microscopy studies showed that the translocation into the cytosol is ESX-1 dependent [79] this explains that the negative control in our study, BCG::pYUB did not cause a shift in the FRET signal. A potential explanation is the described role of ESAT-6 in pore forming [65,90] so it could be a crucial player in offering other proteins access into the cytosol. More is known about the host-pathogen interactions for *M. marinum*. Stamm and colleagues [91] showed that *M. marinum* recruits host actin upon phagosomal rupture to be able to translocate within the cytosol. Yet *Mtb*, in contrast to *M. marinum*, does not recruit actin [92] suggesting that *Mtb* might have lost its ability to move around in favour of its persistence under harsh host conditions [20].

An interesting aspect we observed was the tolerance of the *M. marinum* genetic addition in BCG. While the ESX-1 region of *Mtb* is well tolerated and functional in BCG [64] our data imply that the distance of the *M. marinum* genes may cause cellular lysis of BCG. To verify secretion of the ESX-1 encoded proteins we employed T-cell hybridomas, which are sophisticated immunological tools that sense even minute levels of secreted bacterial proteins presented by Bone Marrow-Dendritic Cells. This confirmed that the ESX-1 substrates ESAT-6, CFP-10 and EspC are indeed secreted, even though at a much lower level than H37Rv WT. To control that the presence of ESX-1 proteins in the culture supernatant is actually due to secretion and not to lysis of the cells we employed anti-GroEL2-antibody, binding to an intracellular chaperone that is not found in culture supernatants. For the BCG::ESX-1MM clones we continued to detect this chaperone in the supernatant from mid-log phase grown cultures starting at day 3. The suggested lysis after
3-4 days may therefore represent the intolerance of BCG towards the *M. marinum* proteins, and function as self-limiting process.

This phenomenon has multiple repercussions on pathogenicity and virulence of this recombinant BCG strain. In a recent study, Simeone and colleagues (manuscript in preparation) showed that induction of early phase phagosomal rupture, a pathogenicity determinant, is independent of replication. Our data are in accordance with these observations as the BCG::ESX-1MM A strain produces a blue shift in our FRET assay indicating cytosolic presence of mycobacteria. This underpins the pathogenic potential of this strain in the light of the observed lysis. Indeed, this feature could exemplify a new vaccine strategy with initial antigen presentation by secretion of virulence factors followed by lysis and hence self-limiting infection, reducing the risk for BCG disease [93] in immunocompromised patients.

Taken together, the immunogenicity and virulence data obtained for the BCG::ESX-1MM A strain make it a promising candidate. Eliciting strong T-cell immunity while being safe are essential features of a vaccine. In a next step the protective efficacy of this strain have been evaluated in a murine model demonstrating that it can protect mice from developing TB upon aerosol challenge with *Mtb*.

While the present invention has been described with reference to the specific embodiments thereof, it should be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the true spirit and scope of the invention. In addition, many modifications may be made to adapt a particular situation, material, composition of matter, process, process step or steps, to the objective, spirit and scope of the present invention. All such modifications are intended to be within the scope of the claims appended hereto.
REFERENCES


Wells, A. Q. 1937. Tuberculosis in wild voles. Lancet 1:1221

CLAIMS

1. A recombinant strain of *Mycobacterium bovis* bacille Calmette-Guerin (*M. bovis* BCG) comprising a heterologous nucleic acid sequence comprising a plurality of open reading frames, wherein the plurality of open reading frames comprise open reading frames that encode proteins each at least 95% homologous to the *Mycobacterium marinum* (*M. marinum*) proteins MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455.

2. The recombinant strain of *M. bovis* BCG according to claim 1, wherein the plurality of open reading frames further comprise an open reading frame that encodes a protein at least 95% homologous to the *M. marinum* proteins MMAR5443, MMAR5444, and MMAR5457.

3. The recombinant strain of *M. bovis* BCG according to claim 2, wherein the plurality of open reading frames further comprise open reading frames that encode proteins at least 95% homologous to the *M. marinum* proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5444, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455.

4. The recombinant strain of *M. bovis* BCG according to any one of claims 1-3, wherein the plurality of open reading frames comprise open reading frames that encode proteins at least 97% homologous to the listed *M. marinum* proteins.

5. The recombinant strain of *M. bovis* BCG according to any one of claims 1-3, wherein the plurality of open reading frames comprise open reading frames that encode proteins at least 99% homologous to the listed *M. marinum* proteins.
6. A recombinant strain of *M. bovis* BCG comprising a heterologous nucleic acid sequence comprising a plurality of open reading frames, wherein the plurality of open reading frames comprise open reading frames that encode *M. marinum* proteins MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455.

7. The recombinant strain of *M. bovis* BCG according to claim 6, wherein the plurality of open reading frames further comprise an open reading frame that encodes the *M. marinum* proteins MMAR5443, MMAR5444, and MMAR5457.

8. The recombinant strain of *M. bovis* BCG according to claim 7, wherein the plurality of open reading frames further comprise open reading frames that encode the *M. marinum* proteins MMAR5429, MMAR5430, MMAR5431, MMAR5432, MMAR5433, MMAR5434, MMAR5435, MMAR5436, MMAR5437, MMAR5438, MMAR5439, MMAR5440, MMAR5441, MMAR5442, MMAR5443, MMAR5445, MMAR5446, MMAR5447, MMAR5448, MMAR5449, MMAR5450, MMAR5451, MMAR5452, MMAR5453, and MMAR5455.

9. The recombinant strain of *M. bovis* BCG according to any one of claims 6-8, wherein the open reading frames that encode the listed proteins have the sequence of the corresponding open reading frames listed in SEQ ID NO:35-67.

10. A recombinant strain of *M. bovis* BCG comprising a heterologous nucleic acid sequence comprising the nucleic acid sequence of SEQ ID NO:1.

11. The recombinant strain of *M. bovis* BCG according to any one of claims 1-10, wherein the heterologous nucleic acid sequence is present on a plasmid.

12. The recombinant strain of *M. bovis* BCG according to any one of claims 1-10, wherein the heterologous nucleic acid sequence is integrated into the *M. bovis* BCG chromosome.
13. The recombinant strain of *M. bovis* BCG according to any one of claims 1-10, wherein the recombinant strain secretes the CFP-10 and ESAT-6 proteins of *M. marinum*.

14. The recombinant strain of *M. bovis* BCG according to any one of claims 1-13, wherein the recombinant strain induces a protective immune response greater than the parent *M. bovis* BCG when introduced into a subject, and wherein the virulence of the recombinant strain is equal to or lower than the virulence of the parent *M. bovis* BCG.

15. A pharmaceutical composition comprising a recombinant strain of *M. bovis* BCG according to any one of claims 1-14 and a carrier.

16. A kit comprising a recombinant strain of *M. bovis* BCG according to any one of claims 1-14 in a container.

17. The kit of claim 16, further comprising at least one protein or peptide antigen of a mycobacterium.

18. The kit of claim 17, wherein the mycobacterium is a strain selected from *M. bovis* BCG, *M. marinum*, and *Mycobacterium tuberculosis* (*M. tuberculosis*).

19. The kit according to any one of claims 17 and 18, wherein the protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof.

20. A method for inducing a protective immune response against *M. tuberculosis* in a subject, comprising administering an effective dose of a pharmaceutical composition according to claim 15 to a subject and inducing an immune response in the subject that is protective against *M. tuberculosis*.

79
21. The method of claim 20, further comprising administering at least one isolated recombinant protein or peptide antigen of a *mycobacterium* to the subject.

22. The method of claim 21, wherein the at least one isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof.

23. The method of claim 20, further comprising administering at least one subunit vaccine to the subject.

24. The method of claim 20, further comprising administering at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, H1-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and *M. vaccae* to the subject.

25. A method for treating an *M. tuberculosis* infection in a subject, comprising administering an effective dose of a pharmaceutical composition according to claim 15 to a subject and inducing an immune response in the subject that is protective against *M. tuberculosis*.

26. The method of claim 25, further comprising administering at least one isolated recombinant protein or peptide antigen of a *mycobacterium* to the subject.

27. The method of claim 26, wherein the at least one isolated recombinant protein or peptide antigen is selected from CFP-10 protein, ESAT-6 protein, and peptides thereof.

28. The method of claim 25, further comprising administering at least one subunit vaccine to the subject.
29. The method of claim 25, further comprising administering at least one vaccine selected from MVA85A, rBCG30, AERAS-402, AdAg85A, M72, H1-IC31, HI-CAF01, H4-IC31 (AERAS-404), rBCGdeltaUreC:Hly (VPM1002), RUTI, and M. vaccae to the subject.

30. A method of making a recombinant strain of M. bovis BCG according to any one of claims 1-14, comprising providing a vector comprising the heterologous nucleic acid sequence comprising the plurality of open reading frames, introducing the vector into M. bovis BCG cells, and selecting M. bovis BCG cells that stably maintain the heterologous nucleic acid sequence comprising the plurality of open reading frames.

31. The method of claim 30, wherein the vector is an integrating vector and the method further comprises selecting M. bovis BCG cells in which the heterologous nucleic acid sequence comprising the plurality of open reading frames has integrated into the host cell chromosome.

32. A recombinant strain of M. bovis BCG according to claim 10, wherein the heterologous nucleic acid sequence comprises the M. marinum nucleic acid sequence inserted in the recombinant pYUB412 vector carried by the bacteria deposited at the CNCM under the reference number 1-4858 on June 3, 2014.

33. A method according to claim 30, wherein the vector is the recombinant pYUB412 vector carried by the bacteria deposited at the CNCM under the reference number 1-4858 on June 3, 2014.
Figure 2
Figure 5
Figure 7

A
ESX-1
*M. mar.*
(38.5 kb)

B

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>-</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6 kb</td>
<td>MMAR_5441</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6 kb</td>
<td>MMAR_5446</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6 kb</td>
<td>MMAR_5452</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reference Cosmid used for transformation
Figure 9

**BCG::pYUB**
- CL
- SN
  - α-CFP-10

**BCG::ESX-1**
- CL
- SN
  - α-ESAT-6

**BCG::RD1 2fg**
- CL
- SN
  - α-CFP-10

**M. marinum USA**
- CL
- SN
  - α-ESAT-6
Figure 13

IF1 Anti-EspC:40-54

NB11 Anti - ESAT 6:1-20

DE10 Ag85A:241-260
Figure 14

[Bar chart showing IFN-γ (pg/ml) response to different antigens: Milieu, rMalE, CFP-10:11-25, rCFP-10, PPD, Con A. Legends indicate responses to BCG::ESX-1 MM Clone A, BCG::ESX-1 Mtb, BCG::PYUB.]
Figure 16 (continued)
Figure 19

[Graph showing percent survival over time post-infection (days)]

- BCG
- BCG-AD139
- M. Tuberculosis
Figure 20 (continued)
Figure 21 (continued)
Figure 21 (continued)
Figure 23 (continued)

B

![Graph showing CFU counts over 6 weeks for BCG::RD1-Marinum, BCG::RD1-279, and H37RV](image)

- **CFU/Lung**
- **CFU/spleen**
<table>
<thead>
<tr>
<th>Box No. I</th>
<th>Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:</td>
</tr>
<tr>
<td></td>
<td>a. □ [ ] forming part of the international application as filed:</td>
</tr>
<tr>
<td></td>
<td>□ in the form of an Annex C/ST.25 text file.</td>
</tr>
<tr>
<td></td>
<td>□ on paper or in the form of an image file.</td>
</tr>
<tr>
<td></td>
<td>b. □ furnished together with the international application under PCT Rule 13fer1 (a) for the purposes of international search only in the form of an Annex C/ST.25 text file.</td>
</tr>
<tr>
<td></td>
<td>c. □ furnished subsequent to the international filing date for the purposes of international search only:</td>
</tr>
<tr>
<td></td>
<td>□ in the form of an Annex C/ST.25 text file (Rule 13fer1 (a)).</td>
</tr>
<tr>
<td></td>
<td>□ on paper or in the form of an image file (Rule 13fer1 (b) and Administrative Instructions, Section 7:13).</td>
</tr>
<tr>
<td>2.</td>
<td>□ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.</td>
</tr>
<tr>
<td>3.</td>
<td>□ Additional comments:</td>
</tr>
</tbody>
</table>
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/35 A61K39/04 C12N1/36

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, Sequence Search, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>wo 03/085098 A2 (PASTEUR INSTITUT [FR]; COLE STEWART [FR]; PYM ALEXANDER S [GB]; BROSCH) 16 October 2003 (2003-10-16) the whole document</td>
<td>1-5, 11, 12, 15-31, 33</td>
</tr>
<tr>
<td>X</td>
<td>PYM A S ET AL: &quot;Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis osi s&quot;, NATURE MEDICINE, vol. 9, no. 5, 14 April 1 2003 (2003-04-14), pages 533-539, XP002402781, ISSN: 1078-8956 cited in the application on the whole document</td>
<td>1-5, 11, 12, 15-31, 33</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:
- A: document defining the general state of the art which is not considered to be of particular relevance
- E: earlier application or patent but published on or after the international filing date
- L: document which may throw doubts on priority claim(s) on which the application is based
- O: document referring to an oral disclosure, use, exhibition or other means
- P: document published prior to the international filing date but later than the international priority date

T: later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle of theory underlying the invention
X: document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
Y: document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

Date of the actual completion of the international search: 27 July 2015
Date of mailing of the international search report: 03/08/2015

Name and mailing address of the ISA/Authorized officer:

van de Kamp, Mart

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel.: (+31-70) 340-2040, Fax: (+31-70) 340-3018
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>BRODIN PRISCILL ET AL: &quot;Di ssecti on of ESAT-6 system 1 of Mycobacterium smegmatis and impact on immunogenicity and virulence&quot;, INFECTION AND IMMUNITY, vol. 74, no. 1, 1 January 2006 (2006-01-01), pages 88-98, XP002474771</td>
<td>1-5, 11, 12, 15-31, 33</td>
</tr>
<tr>
<td>X</td>
<td>TOM H. M. OTTENHOFF ET AL: &quot;Vaccines against Tuberculosis: Where Are We and Where Do We Need to Go?&quot;, PLOS PATHOGENS, vol. 8, no. 5, 10 May 2012 (2012-05-10), page e002607, XP055204235, DOI: 10.1371/journal.ppat.1002607</td>
<td>1-5, 11, 12, 15-31, 33</td>
</tr>
<tr>
<td>A</td>
<td>SIMEONE R ET AL: &quot;ESX/type VII secretion systems and their role in host-pathogen interaction&quot;, CURRENT OPINION IN MICROBIOLOGY, vol. 12, no. 1, 18 January 2009 (2009-01-18), pages 4-10, XP026194629</td>
<td>1, 6, 10, 11, 003, 11, 003, 41, 007</td>
</tr>
</tbody>
</table>

Form PCT/ISA/210 (continuation of second sheet) (April 2005)
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO 03085098 A2</td>
<td>16-10-2003</td>
<td>AT 469963 T</td>
<td>15-06-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AT 535600 T</td>
<td>15-12-2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2003223039 AI</td>
<td>20-10-2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2481318 AI</td>
<td>16-10-2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CY 1110764 TI</td>
<td>10-06-2015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DK 1350839 T3</td>
<td>12-07-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1350839 AI</td>
<td>08-10-2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1492867 A2</td>
<td>05-01-2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1914299 A2</td>
<td>23-04-2008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 2302033 A2</td>
<td>30-03-2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES 2346043 T3</td>
<td>08-10-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HK 1059803 AI</td>
<td>22-10-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PT 1350839 E</td>
<td>12-08-2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2005220811 AI</td>
<td>06-10-2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2012189662 AI</td>
<td>26-07-2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 03085098 A2</td>
<td>16-10-2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2007010413 A2</td>
<td>25-01-2007</td>
</tr>
</tbody>
</table>