

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

CRISTIANE LIMA

Análise funcional do regulador transcricional Rv3405c em *Mycobacterium bovis* BCG

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Biologia Celular e Molecular

Orientador (as): Prof. Dra. Leila de Mendonça Lima
Prof. Dra. Teca Calcagno Galvão

RIO DE JANEIRO

2013

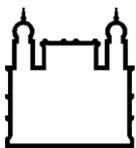
**FICHA CATALOGRÁFICA A SER ELABORADA PELA BIBLIOTECA CENTRAL DE MANGUINHOS PARA A VERSÃO FINAL DA TESE (A ser impressa no verso da primeira folha de rosto)
INDICAR APENAS AS PALAVRAS-CHAVE NA VERSÃO APRESENTADA PARA A DEFESA PÚBLICA DA TESE**

Lima, Cristiane

Análise funcional do regulador transcricional Rv3405c em *Mycobacterium bovis* BCG.
Cristiane Lima. – Rio de Janeiro: 2013.

Tese (Doutorado) – Instituto Oswaldo Cruz, Biologia Celular e
Molecular, 2013.

1. Gene *rv3405c*
2. BCG.
3. Gene *rv3406*.
4. Tet-R I. Título



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

CRISTIANE LIMA

Análise funcional do regulador transcricional Rv3405c em *Mycobacterium bovis* BCG

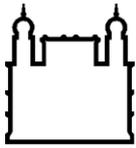
ORIENTADOR (AS): **Prof. Dra. Leila de Mendonça Lima**
 Prof. Dra. Teca Calcagno Galvão

Aprovada em: ____/____/____

EXAMINADORES:

Prof. Dr. Maria Cristina Vidal Pessolani - IOC/FIOCRUZ **Revisora e Presidente**
Prof. Dr. Beny Spira - USP
Prof. Dr. Luis Caetano Martha Antunes - ENSP/FIOCRUZ
Prof. Dr. Paulo Redner - ENSP/FIOCRUZ - suplente
Prof. Dr. Yara Maria Traub-Cseko - IOC/FIOCRUZ - suplente

Rio de Janeiro, 31 de julho de 2013.



Ministério da Saúde

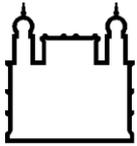
FIOCRUZ

Fundação Oswaldo Cruz

Análise funcional do regulador transcricional Rv3405c em *Mycobacterium bovis* BCG

RESUMO

Em 2011 foram notificados quase 9 milhões de novos casos de tuberculose (TB), culminando em aproximadamente 2 milhões de mortes. A vacina atual contra tuberculose (BCG) apresenta eficácia variável na proteção de adultos contra a tuberculose pulmonar (0% - 80%), com isso destacamos a necessidade de maiores investigações sobre a vacina. A cepa da vacina utilizada no Brasil é a BCG Moreau. Este trabalho incide sobre a caracterização de Rv3405c, um regulador transcricional (repressor) presente em todas as cepas de BCG, porém truncado em BCG Moreau devido à perda da região genômica RD16. Analisamos por RT-PCR a expressão de genes adjacentes a *rv3405c* em BCG Pasteur e Moreau, comprovando que sua perda funcional afeta a expressão de *rv3406*, que codifica uma proteína envolvida no metabolismo do enxofre. O aumento da expressão de *rv3406* em BCG Moreau pode representar um ganho funcional, devido à perda desse repressor transcricional (Rv3405). A fim de caracterizar melhor a sua função, *rv3405c* de BCG Pasteur e *rv3406* de BCG Moreau foram clonados e expressos em *E. coli*. O soro policlonal produzido em camundongos imunizados contra a proteína purificada rRv3406 foi utilizado para analisar sua expressão sob diferentes condições de crescimento. Rv3406 é encontrada na fração intracelular e na superfície de BCG Moreau, não tendo sido detectado em BCG Pasteur. Rv3405c recombinante purificada foi usada no ensaio de retardo em gel (EMSA), confirmando a sua capacidade para se ligar a um motivo de DNA conservado, identificado na região intergênica entre *rv3405c-rv3406*, sugerindo seu papel na regulação destes genes e justificando a expressão constitutiva de *rv3406* em BCG Moreau. Além disso, a adição de bile e seus componentes provoca o deslocamento da proteína desfazendo sua interação com DNA. A caracterização de *rv3405c* e suas funções é importante para uma análise de genômica funcional comparativa entre BCG Moreau e Pasteur. Este trabalho contribui para uma melhor compreensão da vacina brasileira contra a tuberculose.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Functional analysis of the transcriptional regulator Rv3405c in *Mycobacterium bovis* BCG

ABSTRACT

Tuberculosis (TB) was responsible for almost 9 million new cases in 2011, culminating in approximately 2 million deaths. The current approved TB vaccine (BCG) presents variable efficacy in adult protection against pulmonary TB (0%-80%), highlighting the need for further investigations on this vaccine. The vaccine strain used in Brazil is BCG Moreau. This work focuses on the characterization of Rv3405c, a transcriptional regulator (repressor) present in all BCG strains but truncated in BCG Moreau due to the loss of genomic region RD16. We have analysed the expression of adjacent genes in both BCG Moreau and Pasteur by RT-PCR finding that loss of Rv3405c impacts the expression of *rv3406*, encoding a protein involved in sulfur metabolism. Its increased expression in BCG Moreau may represent a functional gain through the loss of a transcription repressor. In order to better characterize their function, *rv3405c* from BCG Pasteur and *rv3406* from BCG Moreau were cloned and expressed in *E. coli*. The polyclonal sera produced in mice immunized with purified rRv3406 were used to follow the expression of this protein under different growth conditions. Rv3406 is found in the intracellular fraction and the surface of BCG Moreau, and was not detected in BCG Pasteur. Purified recombinant Rv3405c was used in electrophoretic mobility shift assays (EMSA), confirming its ability to bind a conserved DNA motif identified in the intergenic region between *rv3405c-rv3406*, suggesting its role in regulating these genes and justifying the constitutive expression of *rv3406* in BCG Moreau. Furthermore, the addition of bile and its components is able to displace Rv3405c from DNA. The characterization of *rv3405c* and its functions is important for the comparative functional genomics analysis of BCG Moreau and Pasteur. This work contributes to a better understanding of the Brazilian vaccine against TB.

Dedico esta tese à minha filha.

Um dia você Aprende...

Depois de algum tempo você aprende a diferença, a sutil diferença entre dar a mão e acorrentar uma alma. E você aprende que amar não significa apoiar-se, e que companhia nem sempre significa segurança. E começa a aprender que beijos não são contratos e presentes não são promessas. E começa a aceitar suas derrotas com a cabeça erguida e olhos adiante, com a graça de um adulto e não com a tristeza de uma criança. E aprende a construir todas as suas estradas no hoje, porque o terreno do amanhã é incerto demais para os planos, e o futuro tem o costume de cair em meio ao vão. Depois de um tempo você aprende que o sol queima se ficar exposto por muito tempo. E aprende que não importa o quanto você se importe, algumas pessoas simplesmente não se importam... E aceita que não importa quão boa seja uma pessoa, ela vai feri-lo de vez em quando e você precisa perdoá-la por isso. Aprende que falar pode aliviar dores emocionais. Descobre que leva-se anos para construir confiança e apenas segundos para destruí-la, e que você pode fazer coisas em um instante, das quais se arrepende pelo resto da vida. Aprende que verdadeiras amizades continuam a crescer mesmo a longas distâncias. E o que importa não é o que você tem na vida, mas quem você tem da vida. E que bons amigos são a família que nos permitiram escolher. Aprende que não temos que mudar de amigos se compreendemos que os amigos mudam, percebe que seu melhor amigo e você podem fazer qualquer coisa, ou nada, e terem bons momentos juntos. Descobre que as pessoas com quem você mais se importa na vida são tomadas de você muito depressa - por isso, sempre devemos deixar as pessoas que amamos com palavras amorosas, pode ser a última vez que as vejamos. Aprende que as circunstâncias e os ambientes têm influência sobre nós, mas nós somos responsáveis por nós mesmos. Começa a aprender que não se deve comparar com os outros, mas com o melhor que pode ser. Descobre que se leva muito tempo para se tornar a pessoa que quer ser, e que o tempo é curto. Aprende que não importa onde já chegou, mas onde está indo, e se você não sabe para onde está indo, qualquer lugar serve. Aprende que, ou você controla seus atos ou eles o controlarão, e que ser flexível não significa ser fraco ou não ter personalidade, pois não importa quão delicada e frágil seja uma situação, sempre existem dois lados. Aprende que heróis são pessoas que fizeram o que era necessário fazer, enfrentando as consequências. Aprende que paciência requer muita prática. Descobre que algumas vezes, a pessoa que você espera que o chute quando você cai, é uma das poucas que o ajudam a levantar-se. Aprende que maturidade tem mais a ver com os tipos de experiência que se teve e o que você aprendeu com elas, do que com quantos aniversários você celebrou. Aprende que há mais dos seus pais em você do que você supunha. Aprende que nunca se deve dizer a uma criança que sonhos são bobagens, poucas coisas são tão humilhantes e seria uma tragédia se ela acreditasse nisso. Aprende que quando está com raiva tem o direito de estar com raiva, mas isso não te dá o direito de ser cruel. Descobre que só porque alguém não o ama do jeito que você quer que ame, não significa que esse alguém não o ama com tudo o que pode, pois existem pessoas que nos amam, mas simplesmente não sabem como demonstrar ou viver isso. Aprende que nem sempre é suficiente ser perdoado por alguém, algumas vezes você tem que aprender a perdoar-se a si mesmo. Aprende que com a mesma severidade com que julga, você será em algum momento condenado. Aprende que não importa em quantos pedaços seu coração foi partido, o mundo não pára para que você o conserte. Aprende que o tempo não é algo que possa voltar para trás. Portanto, plante seu jardim e decore sua alma, ao invés de esperar que alguém lhe traga flores. E você aprende que realmente pode suportar... que realmente é forte, e que pode ir muito mais longe depois de pensar que não se pode mais. E que realmente a vida tem valor e que você tem valor diante da vida!"

William Shakespeare

Agradecimentos

A minha família, ao ECEPRAB (Espaço cultural de estudos e práticas religiosas afro-brasileiras) e amigos por sempre estarem presentes na minha vida dando suporte e apoio.

A Dra Leila de Mendonça Lima e Dra Teca Calcagno Galvão por toda dedicação, ensinamentos e apoio ao longo de todo meu doutorado.

Ao Dr Wim Degrave pela confiança em mim depositada.

A todos os amigos do Laboratório de Genômica Funcional e Bioinformática pelo auxílio nos experimentos cotidianos e por todo apoio dispensado especialmente nos momentos que as situações se tornaram mais difíceis.

A Plataforma de Sequenciamento de DNA da Fiocruz (PDTIS).

Ao órgão financiador da minha bolsa de estudos CNPq.

LISTA DE SIGLAS E ABREVIATURAS

°C - grau Celsius

µm - Micrômetro

Ag85 - Antígeno 85

BAAR - Bacilo álcool-ácido resistente

BCG - Bacilo de Calmette-Guérin

C - Citosina

CBB - *Coomassie brilliant blue*

cm - Centímetro

CRP - *c-AMP receptor protein*

Da - Dalton

DBD - *DNA binding domain*

DNA - Ácido desoxirribonucléico

DO - Densidade ótica

E. coli - *Escherichia coli*

EDTA - Ácido etileno diamino tetracético

EMSA - *Electrophoretic mobility shift assay*

FAP - Fundação Atauilho de Paiva

FC - Filtrado de Cultura

g - Aceleração da gravidade

G - Guanina

GC - Citosina e guanina

GCDC - Glicoquenodeoxicolato de sódio

h - Hora

HIV - *Human immuno deficiency virus*

IEF - Focagem isoelétrica (*Isoelectric focusing*)

IS6110 - Sequência de inserção 6110

kb - Quilobase

kDa - Quilodalton

kV - Quilovolts

LB - Luria Bertani

M - Molar

mg - Miligrama

min. - Minuto

mL - Mililitro

mm - Milímetro
mM - Milimolar
Mpb - Um milhão de pares de base
M. tuberculosis - *Mycobacterium tuberculosis*
nm - Nanômetro
NRAMP1 - *Natural resistance associated macrophage protein*
PAGE - Eletroforese em gel de poliacrilamida (*polyacrylamide gel electrophoresis*)
pb (s) - Par (es) de base
PBS - *Phosphate Buffered Saline*
PDIM - Dimicolato de Ftiocerol
PDTIS - Programa de Desenvolvimento Tecnológico em Insumos para Saúde
PE - *Proline-glutamic acid*
Pfam - Banco de dados de famílias de proteínas
PGL - Glicolídeo Fenólico
PGRS - Região polimórfica rica em G+C (*Polimorphic GC- Rich Sequence*)
pH - Potencial Hidrogeniônico
PPE - *Proline-proline-glutamic acid*
HPLC - Cromatografia líquida de Alta Eficiência (*High performance liquid chromatography*)
HTH - *Helix-turn-helix*
qsp - Quantidade suficiente para
RD - Região de diferença ou região de deleção
RFLP - *Restriction fragment length polymorphism*
RNA - Ácido ribonucleico
RNAP - RNA polymerase
RT-PCR - *Reverse transcription polymerase chain reaction*
rpm - Rotações por minuto
SDS - Dodecil sulfato de sódio
SF - Superfície
SNPs - Polimorfismos de base única (*single nucleotide polymorphisms*)
TauD - Taurina dioxigenase
TB - Tuberculose
TBE - *Tris-Borate-EDTA*
TBS - *Tris Buffered Saline*
TCA - Ácido tricloro acético
TCDC - Tauroquenodeoxicolato de sódio

TE - Tris-EDTA

TEA - Trietanolamina

TEMED - N,N,N,N' - Tetrametiletilenodiamina

TRIS - Tris-hidroxi-metil-amino-metano

TRITON X -100 - [(1,1,3,3-Tetrametilbutil)fenil-polietileno glicol]

UFC - Unidades formadoras de colônia

v - Volts

v/v - Volume/volume

WB - *Western blot*

WHO - *World Health Organization*

μF - MicroFaraday

ng - Nanograma

μg - Micrograma

μL - Microlitro

μm - Micrômetro

Ω - Ohm

σ - Sigma

LISTA DE FIGURAS

INTRODUÇÃO

1.1	Incidência de casos de tuberculose no mundo no ano de 2011	1
1.2	Processo de infecção por <i>M. tuberculosis</i>	3
1.3	Esquema representativo do ciclo da infecção por <i>M. tuberculosis</i>	4
1.4	Genealogia das vacinas BCG	8
1.5	Representação esquemática da região RD16 em <i>M. tuberculosis</i> H37Rv e do gene <i>rv3405c</i> (00-05) que aparece truncado em <i>M. bovis</i> BCG Moreau	9
1.6	Ligação de TetR no seu sítio operador	13
1.7	Vias metabólicas de taurina e hipotaurina, evidenciando a etapa catalizada pela TauD e sua ligação ao metabolismo de enxofre	14

RESULTADOS

4.1	Alinhamento das regiões intergênicas entre <i>rv3405c</i> e <i>rv3406</i> mostrando a presença de um motivo palindrômico típico de reguladores da família TetR	32
4.2	RT-PCR para o gene <i>rv3406</i> em BCG Moreau e BCG Pasteur	34
4.3	Clonagem em vetor pET28a dos genes <i>rv3405c</i> (pET-P3405c) e <i>rv3406</i> (pET-M3406) para expressão em BL21(DE3)	35
4.4	Super-expressão das proteínas recombinantes Rv3405c (A) e Rv3406 (B).	35
4.5	WB para confirmação do peso molecular das proteínas Rv3405c (A) e Rv3406 (B)	36
4.6	Teste de solubilidade das construções expressando as proteínas recombinantes Rv3405c (A) e Rv3406 (B).	36
4.7	Teste de solubilidade de Rv3406 (pET-M3406) em meio “overnight express” por 16 horas (~25°C)	37
4.8	Purificação das proteínas Rv3405c (A) e Rv3406 (B) por HPLC a partir dos corpos de inclusão	38
4.9	Proteínas Rv3405c e Rv3406 purificadas a partir de corpos de inclusão	38
4.10	Purificação das proteínas recombinantes solúveis Rv3405c (A) e Rv3406 (B)	39
4.11	WB para titulação do anticorpo policlonal anti-Rv3405c produzido em camundongos Balb-c	40
4.12	WB para titulação do anticorpo policlonal anti-Rv3406 produzido em camundongos Balb-c	40
4.13	WB para detectar a presença da proteína Rv3406 nas frações de proteínas	41

	intracelulares, de superfície e de filtrado de cultura de BCG Moreau e BCG Pasteur utilizando anticorpo primário anti-Rv3406	
4.14	RT-PCR para o gene <i>rv3405c</i> em BCG Moreau (gene quimérico <i>rv3405-00c</i>) e em BCG Pasteur (<i>rv3405c</i>)	41
4.15	RT-PCR para o gene quimérico <i>rv3400-05c</i> em ambas as direções, tanto a partir do promotor de <i>rv3400</i> quanto do promotor de <i>rv3405c</i>	42
4.16	RT PCR para <i>rv3400</i> e <i>rv3406</i> em BCG Moreau mostrando que não são co-transcritos	43
4.17	EMSA para demonstrar a interação entre a proteína Rv3405c e a região intergênica <i>rv3405-06</i>	44
4.18	EMSA para visualização da interação entre a região intergênica 05c-06 e a proteína rRv3405c purificada na presença de concentrações crescentes de taurina (painel A) ou bile bovina (painel B)	45
4.19	EMSA utilizando a região intergênica 05c-06 e a proteína rRv3405c para verificar a capacidade de componentes da bile bovina em desfazer essa ligação	46
4.20	RT-PCR para detecção de mRNA de <i>rv3406</i> em culturas de BCG Moreau e Pasteur cultivadas na presença de 4 e 2 mM Taurina (Tau), 0,05 e 0,1% bile e 100 uM tauroquenodeoxicolato (TCDC) e glicoquenodeoxicolato (GCDC)	48
4.21	WB em culturas de BCG Moreau e BCG Pasteur cultivadas em meio 7H9/ADC/Tween 80 na presença de bile, taurina, tauroquenodeoxicolato (TCDC) e glicoquenodeoxicolato (GCDC) utilizando o anticorpo primário anti-Rv3406	48
	DISCUSSÃO	49
5.1	Esquema representativo do gene truncado (<i>rv3400-rv3405</i>) com seus respectivos promotores funcionais (P_{rv3400} e $P_{rv3405c}$) e <i>rv3406</i> (P_{rv3406})	51

LISTA DE TABELAS

METODOLOGIA

- | | | |
|-----|---|----|
| 3.1 | Reação de clonagem TOPO [®] TA | 23 |
| 3.2 | Razão entre DNA biotilado/DNA não biotilado | 30 |
| 3.3 | Variação da concentração de compostos nas reações de EMSA | 31 |

RESULTADOS

- | | | |
|-----|--|----|
| 4.1 | Concentração inibitória dos diferentes componentes da bile testados no ensaio de ligação de rRv3405c ao DNA da região intergênica 05c-06 | 47 |
|-----|--|----|

SUMÁRIO

	RESUMO	iv
	ABSTRACT	v
	LISTA DE SIGLAS E ABREVIATURAS	ix
	LISTA DE FIGURAS	xii
	LISTA DE TABELAS	xiv
1	INTRODUÇÃO	1
1.1	A tuberculose e seu contexto atual	1
1.2	Transmissão e patogenicidade	2
1.3	O complexo <i>Mycobacterium tuberculosis</i>	5
1.4	<i>M. tuberculosis</i> : agente etiológico da tuberculose	5
1.5	<i>M. bovis</i> BCG: origem e plasticidade genômica	7
1.6	<i>M. bovis</i> BCG Moreau: cepa vacinal brasileira	9
1.7	Regulação da expressão gênica em micobactérias	10
1.8	Reguladores do tipo TetR: <i>rv3405c</i>	11
1.9	Taurina dioxigenase e sua importância no metabolismo de enxofre	14
2	OBJETIVOS	16
3	METODOLOGIA	18
3.1	Condições de cultivo de cepas bacterianas	18
3.1.1	<i>Escherichia coli</i>	18
3.1.2	Cepas vacinais de <i>M. bovis</i> BCG Pasteur e Moreau	18
3.2	Obtenção de proteínas do filtrado de cultura de BCG Pasteur e Moreau	18
3.3	Obtenção das proteínas associadas à superfície de BCG Pasteur e Moreau	19
3.4	Obtenção do lisado micobacteriano	19
3.5	Precipitação das proteínas presentes nas diferentes frações de BCG	19
3.6	Determinação da concentração proteica	20
3.7	SDS-PAGE	20
3.8	Extração de DNA genômico das culturas de <i>M. bovis</i> BCG Pasteur e Moreau	20
3.9	Clonagem, expressão e purificação de proteínas recombinantes em <i>E. coli</i>	21
3.9.1	Preparo de células competentes de <i>E. coli</i> para eletroporação	21
3.9.2	Transformação por eletroporação em <i>E. coli</i>	22
3.9.3	Amplificação das regiões codificadoras de <i>rv3405c</i> e <i>rv3406</i> por PCR	22

3.9.4	Purificação de produtos de PCR	22
3.9.5	Extração e sequenciamento de DNA plasmidial	22
3.9.6	Subclonagem no vetor pBAD TOPO [®] TA	23
3.9.7	Digestão e ligação para clonagem em pET28a	23
3.9.8	Transformação da cepa BL21(DE3) e indução da expressão de proteínas recombinantes com IPTG	24
3.9.9	Teste de solubilidade das proteínas recombinantes	24
3.9.10	“Western blot”	25
3.9.11	Purificação das proteínas recombinantes a partir de corpos de inclusão	25
3.9.12	Purificação das proteínas recombinantes solúveis	26
3.10	Obtenção de soro policlonal contra as proteínas recombinantes	26
3.11	Titulação dos anticorpos	27
3.12	“Western blot” para as frações proteicas de <i>M. bovis</i> BCG Pasteur e <i>M. bovis</i> BCG Moreau	27
3.12.1	Preparo de lisado de <i>E. coli</i>	27
3.13	Extração de RNA total de culturas de <i>M. bovis</i> BCG Pasteur e Moreau	28
3.14	RT-PCR	29
3.15	Ensaio de retardo em gel (EMSA)	29
3.15.1	EMSA	29
3.15.2	EMSA com competição entre DNA biotilado e DNA não biotilado	30
3.15.3	EMSA na presença de pequenas moléculas	30
4	RESULTADOS	32
4.1	Alinhamento entre as regiões reguladoras de <i>rv3405c</i> mostra que o motivo palindrômico é conservado entre diversas espécies de micobactérias	32
4.2	RT-PCR revela que o gene <i>rv3406</i> é transcrito de forma constitutiva em BCG Moreau	33
4.3	Clonagem, expressão e purificação das proteínas recombinantes Rv3405c de <i>M. bovis</i> BCG Pasteur e Rv3406 de <i>M. bovis</i> BCG Moreau em <i>E. coli</i>	33
4.3.1	Amplificação dos genes <i>rv3405c</i> de <i>M. bovis</i> BCG Pasteur e <i>rv3406</i> de <i>M. bovis</i> BCG Moreau	33
4.3.2	Subclonagem no vetor pBAD TOPO TA e clonagem no vetor pET28a	34
4.3.3	Super-expressão e teste de solubilidade das proteínas recombinantes rRv3405c e rRv3406 em <i>E. coli</i> BL21(DE3)	35
4.3.4	Purificação das proteínas Rv3405 e Rv3406 recombinantes a partir de corpos de inclusão	37

4.3.5	Purificação das proteínas Rv3405c e Rv3406 recombinantes a partir da fração solúvel	39
4.3.6	Titulação dos anticorpos	39
4.4	A proteína Rv3406 é detectada por “western blot” em BCG Moreau, mas não em BCG Pasteur	40
4.5	RT-PCR para o gene <i>rv3405c</i> comprova que existe transcrição desse gene em BCG Pasteur e em BCG Moreau (gene quimérico <i>rv3405c-00</i>)	41
4.6	O gene quimérico <i>rv3400-05c</i> é transcrito em ambas as direções em BCG Moreau: do promotor de <i>rv3400</i> (<i>rv3400-05c</i>) e do promotor de <i>rv3405c</i> (<i>rv3405c-rv3400</i>)	42
4.7	Os genes <i>rv3400-05c</i> (BCG_M3438) e <i>rv3406</i> não são co-transcritos em BCG Moreau	43
4.8	EMSA mostra a interação entre Rv3405c e a região intergênica <i>rv3405c-06</i>	43
4.9	EMSA revela que a interação entre Rv3405c e a região intergênica <i>rv3405-06</i> pode ser modulada por compostos presentes em bile bovina	44
4.10	RT-PCR para o gene <i>rv3406</i> nos cultivos de BCG Pasteur e Moreau mostra que a presença de ligantes de Rv3405c induzem a transcrição de <i>rv3406</i>	47
4.11	“Western blot” para detecção de Rv3406 nos cultivos de BCG Pasteur e Moreau incubados com os possíveis moduladores confirma a acumulação desta proteína apenas nas culturas de BCG Moreau	48
5	DISCUSSÃO	49
6	CONCLUSÕES	56
7	REFERÊNCIAS BIBLIOGRÁFICAS	57
8	ANEXOS	65
	ANEXO I - Materiais utilizados neste estudo	65
	ANEXO II - Oligonucleotídeos utilizados nesse estudo	72
	ANEXO III - Condições de Ciclagem para PCRs	73
	ANEXO IV - Manuscrito submetido para publicação	75

1. INTRODUÇÃO

1.1. A tuberculose e seu contexto atual

A tuberculose (TB), causada pela bactéria *Mycobacterium tuberculosis* (*M. tuberculosis*), é uma das principais doenças em países em desenvolvimento, bem como um problema crescente em diversas áreas desenvolvidas do mundo. Em 1994, a Organização Mundial de Saúde (WHO, 1994) declarou a tuberculose uma emergência de saúde pública mundial, devido à ocorrência de cerca de 8 milhões de casos e 1,6 milhões de mortes. Os dados atuais de incidência de novos casos de TB em 2011 são vistos por área na **Figura 1.1**; estimam-se quase 9 milhões de casos novos e 1,4 milhões de mortes pela doença (WHO, 2012). Há pelo menos quase duas décadas, a TB ocupa o segundo lugar mundialmente em causa de morte por doença infecciosa, perdendo apenas para a síndrome de imunodeficiência adquirida (WHO, 1994; WHO, 2012). Apesar da mobilização a nível mundial para o controle da TB, o número de casos de pacientes infectados por *M. tuberculosis* multirresistente (tolerante a pelo menos duas das principais drogas usadas no tratamento, isoniazida e rifampicina) está em ascensão, dobrando nos 27 países com maior número de casos entre 2009 e 2011 (WHO, 2012).

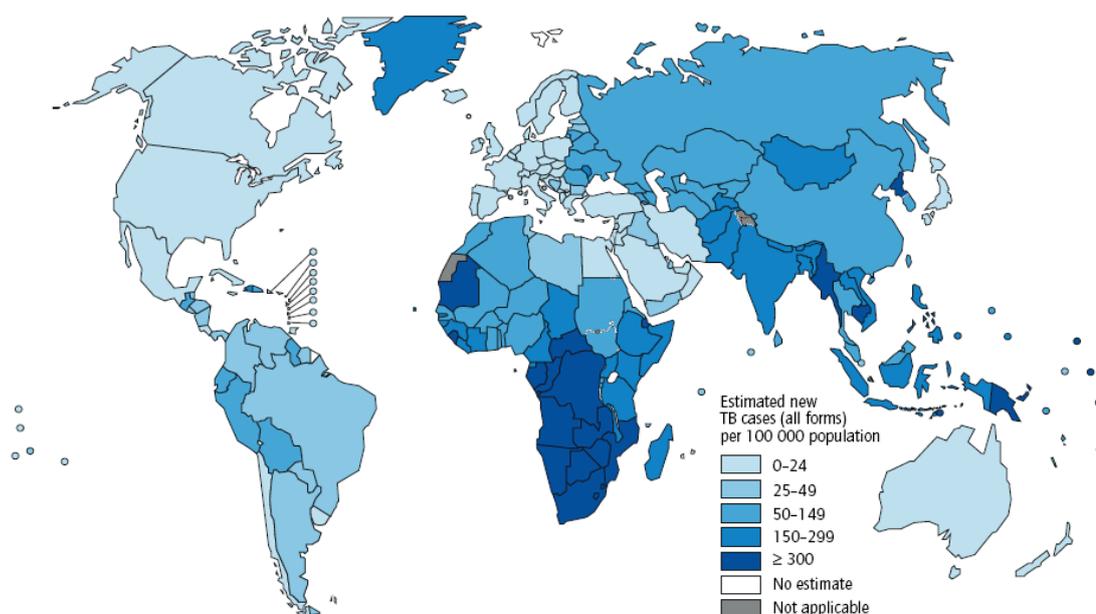


Figura 1.1. Incidência de casos de tuberculose no mundo no ano de 2011. (WHO, 2012)

No contexto atual de aumento no número de cepas de *M. tuberculosis* resistentes às drogas do tratamento, é urgente o desenvolvimento de vacinas mais eficazes, métodos diagnósticos mais rápidos e específicos, e novas drogas. Para abordar estas necessidades, estão em andamento estratégias governamentais e de pesquisa, que combinam as áreas de Genômica Funcional, Proteômica e Bioinformática, que apresentam um grande potencial para permitir o aprimoramento do tratamento da TB (Cole *et al.*, 1998).

1.2. Transmissão e patogenicidade

A transmissão da tuberculose é influenciada pelas condições sócio-econômicas. A TB se prolifera em áreas de grande concentração humana, com precários serviços de infra-estrutura urbana, onde coexistem a fome e a miséria. A tuberculose é disseminada por aerossóis como, por exemplo, a fala, o espirro e, principalmente, a tosse de um doente. Gotículas contendo no seu interior partículas de *M. tuberculosis* são liberadas a partir dos pulmões de pacientes com quadro clínico cavitário pulmonar (lesão cavitária) (**Figura 1.2**). A maioria dessas gotículas ficam retidas na mucosa do trato respiratório superior e são removidas dos brônquios pelo mecanismo mucociliar. Os bacilos removidos são deglutidos, inativados pelo suco gástrico e eliminados nas fezes. (Barreto *et al.*, 2002).

Daqueles que são expostos ao *M. tuberculosis*, somente 10-30% tornam-se infectados, indicando que muitas pessoas podem ser capazes de impedir que a infecção se estabeleça depois que a micobactéria atinge os pulmões. Daqueles que se tornam infectados, adquirindo a sensibilidade tardia, 90% ou mais não desenvolvem a tuberculose, significando que são capazes de resolver a infecção completamente, ou de reduzir a mesma a um nível incapaz de causar a doença. No último caso, a infecção é considerada latente e capaz de ser reativada em um estágio futuro. As pessoas com tuberculose latente representam um reservatório grande de *M. tuberculosis*. Assim, os seres humanos predispostos mostram um espectro de susceptibilidade (Dye *et al.*, 1999, Stewart *et al.*, 2003; Tufariello *et al.*, 2003). Atualmente, é difícil prever quais indivíduos infectados desenvolverão a doença. Logo, um dos grandes desafios na luta contra a tuberculose é ser capaz de diagnosticar, e tratar estes indivíduos antes de desenvolverem a forma ativa da doença.

A resistência/susceptibilidade à tuberculose envolve uma série de genes. Estudos de rastreamento de genoma de populações para determinar as regiões cromossômicas dos genes envolvidos na susceptibilidade à tuberculose (Cervino *et al.*, 2002), assim como estudos de casos controle de genes candidatos têm sido realizados (Marquet e Schurr, 2001; Cervino *et al.*, 2002). Polimorfismos nos genes ligados, por exemplo, à resistência de macrófagos

(NRAMP1), receptores de vitamina D, e proteínas de ligação a manose mostram que estas vias parecem estar associadas com a susceptibilidade à tuberculose. O aspecto central da susceptibilidade em humanos com relação aos fatores imunológicos envolve exclusivamente os pulmões em 85% dos casos (Kramnik *et al.*, 2000; Mitsos *et al.*, 2000; Casanova e Abel, 2002). Em contraste, nos indivíduos imunodeprimidos como os infectados pelo “Human Immunodeficiency Virus” (HIV), a tuberculose pode ser uma doença sistêmica envolvendo múltiplos órgãos (Nunn *et al.*, 2005). Isto sugere que a doença é confinada aos pulmões em indivíduos imunocompetentes devido aos mecanismos de defesa capazes de controlar a doença em todos os outros órgãos (Russell, 2001; North *et al.*, 2004).

O ciclo da infecção por *M. tuberculosis* se inicia quando uma pessoa infectada espirra ou tosse jogando no ar cerca de dois milhões de bacilos que permanecem em suspensão durante horas. Após a inalação, estes bacilos em suspensão no ar entram em contato com o primeiro sistema de defesa, o mucociliar. Contudo, alguns chegam aos alvéolos e são fagocitados por macrófagos residentes no pulmão, formando o sítio primário da infecção. No interior dos macrófagos a micobactéria se reproduz e à medida que a resposta imune progride, outros macrófagos e linfócitos se acumulam confinando o patógeno em um granuloma. O centro do granuloma torna-se caseoso devido a formação de áreas de necrose. O crescimento bacteriano é inibido pela baixa tensão de oxigênio e pH (**Figura 1.2**) (Stewart *et al.*, 2003).

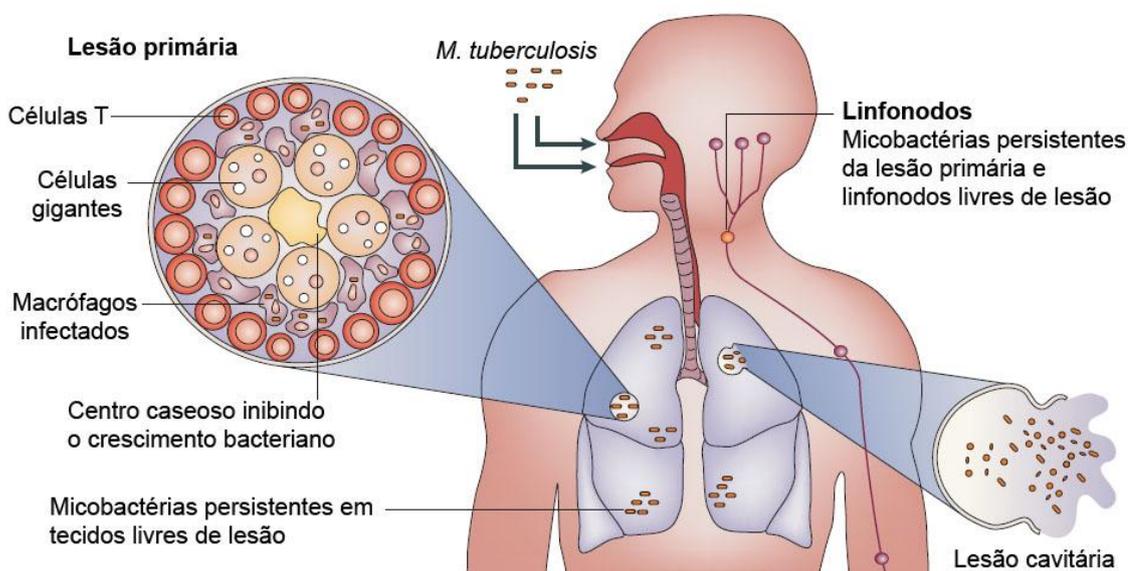


Figura 1.2. Processo de infecção por *M. tuberculosis*. Após inalação, partículas com bacilos em suspensão no ar entram em contato com o sistema de defesa mucociliar. Os bacilos que chegam aos alvéolos e são fagocitados por macrófagos residentes formam a lesão primária da infecção, onde ocorre a multiplicação dos bacilos. O granuloma provém de uma parede de macrófagos e linfócitos em torno das células infectadas de tal forma que assim que as células morrem, o centro do granuloma torna-se caseoso devido a formação de áreas de necrose, onde o crescimento bacteriano é inibido pela baixa tensão de oxigênio e pH. (Stewart *et al.*, 2003).

Dentro do granuloma, o bacilo fica retido e em estado de latência podendo permanecer assim durante anos na denominada tuberculose latente (Gomez & McKinney, 2004). Este granuloma consiste em um núcleo de macrófagos infectados, cercado por células do sistema imune e é delimitado por uma camada fibrosa composta de colágeno (Russell, 2001). O acúmulo de macrófagos e linfócitos no granuloma permite que os macrófagos fagocitem as células infectadas mortas (componentes celulares e bacilos liberados), prevenindo a disseminação da infecção (Saunders & Cooper, 2000).

Com a morte dos macrófagos infectados, no centro do granuloma ocorrem áreas de necrose. A amplificação da resposta imune intensifica o processo inflamatório e acaba por causar a destruição de tecidos vizinhos, promovendo necrose caseosa e formação de cavernas (cavitação). As cavernas são formadas pelo parênquima pulmonar destruído. Caso a lesão ocorra dentro ou próxima a um vaso, os bacilos são carregados pela corrente sanguínea (via linfo-hematogênica), disseminando a doença para outros órgãos. As reações são acompanhadas por fibrose e as lesões podem ser vistas em radiografias torácicas de pacientes afetados (Raviglione & O'Brien, 1998; Stewart *et al.*, 2003).

O ciclo da infecção por *M. tuberculosis* está representado na **Figura 1.3**.

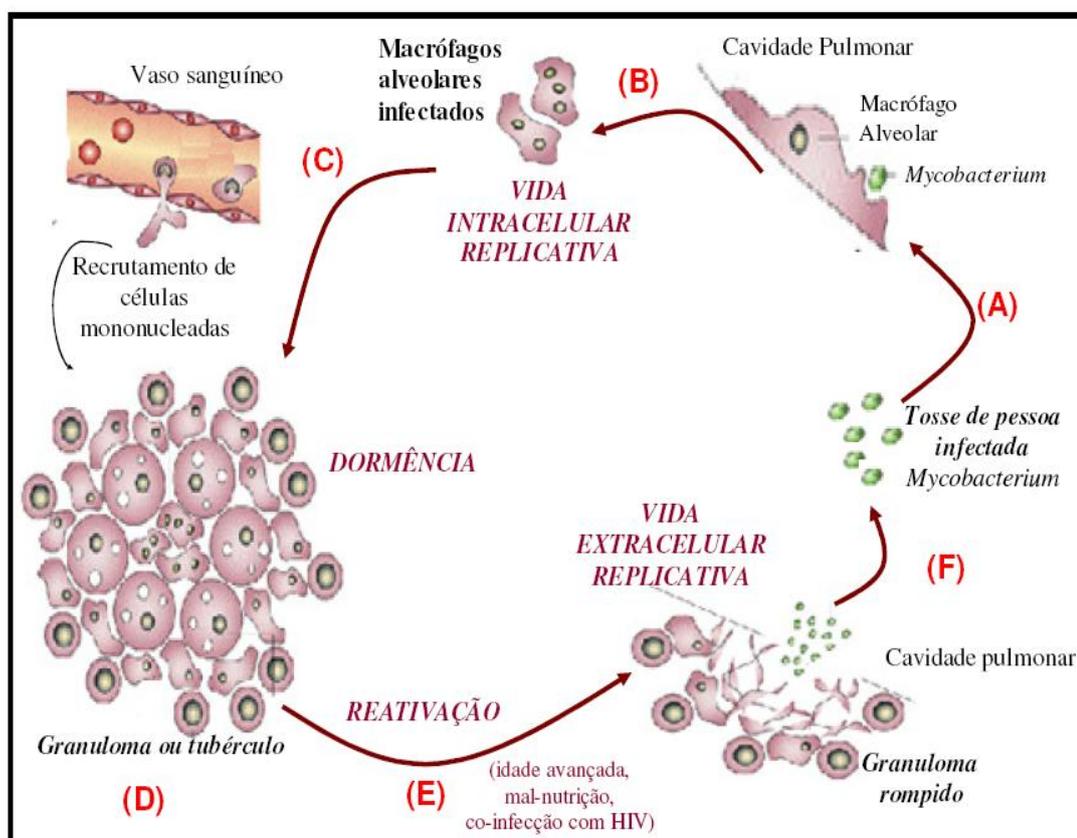


Figura 1.3. Esquema representativo do ciclo da infecção por *M. tuberculosis*. (Adaptado de Russel, 2001)

1.3. O complexo *Mycobacterium tuberculosis*

O *M. tuberculosis* é considerado uma bactéria Gram-positiva, da família Mycobacteriaceae, podendo apresentar uma forma ligeiramente curva ou reta (1 - 10 µm), sendo imóvel, aeróbio estrito, não esporulado nem encapsulado e tem como principal característica uma superfície celular complexa, rica em lipídeos, proteínas e açúcares, a qual lhe confere propriedades importantes como a resistência a alguns agentes químicos como álcool e ácido, sendo assim chamado de bacilo álcool-ácido resistente (BAAR), não sendo capaz de corar pelo método de Gram. Além disso, esta superfície celular confere a ele uma impermeabilidade à maioria dos antibióticos, propiciando um grau de resistência contra mecanismos micobactericidas, que é complementada pelas enzimas detoxificadoras produzidas pelo bacilo (Daffe *et al.*, 1999; Stewart *et al.*, 2003; Besra *et al.*, 1995). Os bacilos da tuberculose são também extremamente sensíveis ao calor e à radiação ultravioleta (Collins, 1971; Riley *et al.*, 1976).

O Complexo *M. tuberculosis* é composto por um grupo de micobactérias patogênicas altamente relacionadas entre si que compreende: *M. tuberculosis*, *M. africanum* (subtipo I e II), *M. bovis* (inclusive o *M. bovis* atenuado, bacilo de Calmette-Guérin, cepa vacinal), *M. bovis* subespécie *caprae*, e *M. microti* (Brosch *et al.*, 2002). As micobactérias do complexo *M. tuberculosis* são estritamente patogênicas, de crescimento lento e, antes de desencadear a doença no homem, encontram-se dissipadas no meio ambiente (Brennan e Nikaido, 1995). Elas diferem amplamente em termos de tropismos em relação ao hospedeiro, fenótipos e patogenicidade. Partindo do princípio de que todos eles são derivados de um ancestral comum, é intrigante que alguns são agentes patogênicos exclusivamente de humanos, como por exemplo *M. tuberculosis* e *Mycobacterium africanum*, ou de roedores (*Mycobacterium microti*), enquanto outros têm um espectro amplo de acolhimento (*Mycobacterium bovis*).

1.4. *M. tuberculosis*: agente etiológico da tuberculose

M. tuberculosis, identificado pelo médico alemão Robert Koch em 1882 como agente etiológico da tuberculose compartilha com outros membros do gênero *Mycobacterium* de uma parede celular de composição única devido à presença predominante de ácidos micólicos, responsáveis por mais de 50% de seu peso seco (Brennan, 2003). O tempo de geração estimado do *M. tuberculosis* em meio sintético ou em animais infectados é de 24 horas, o que contribui para a natureza crônica da doença, impondo esquemas terapêuticos longos e representando um obstáculo às pesquisas. O estado de dormência, no qual o bacilo permanece

quiescente no tecido infectado, pode refletir um “desligamento” metabólico resultante da ação da resposta imune celular do hospedeiro que pode conter, porém não eliminar a infecção (Chan e Kaufmann, 1994). As bases moleculares da dormência assim como da reativação do bacilo permanecem até os dias atuais não esclarecidas completamente. Todavia, acredita-se que sejam programadas geneticamente e que envolvam vias de sinalização intracelulares (Cole *et al.*, 1998).

As cepas de *M. tuberculosis* isoladas de pacientes podem ser distinguidas através da análise de polimorfismos, como, por exemplo, o tamanho de fragmentos de restrição (RFLP) do DNA. O método de “restriction fragment length polymorphism”(RFLP) empregado mais usualmente é baseado no polimorfismo numérico de cópias da sequência cromossômica da inserção *IS6110* (Brosch *et al.*, 2001) que se insere em locais diferentes ao longo do cromossomo. A utilidade do perfil *IS6110* é baseada em sua estabilidade relativa.

O sequenciamento do genoma de *M. tuberculosis* cepa H37Rv revelou 4.411.532 pb, com aproximadamente 4000 genes codificantes para proteínas, dos quais 1019 genes (correspondendo a cerca de 25% do total) foram categorizados como hipotéticos ou sem função atribuída (Lew *et al.*, 2011). Estudos de anotação funcional através de transferência de função por ortologia demonstraram que uma parte desses genes (600 genes) permanece sem função atribuída (Doerks *et al.*, 2012). Dos cerca de 4.4 Mpb, 65,9% são constituídos por guanina (G) e citosina (C). As consequências diretas do elevado conteúdo G + C é a alta frequência com que o códon de iniciação GTG é encontrado no genoma (35%) e a composição do proteoma, onde há a prevalência de aminoácidos codificados por códons ricos em G + C, como alanina, glicina, prolina e arginina (Cole *et al.*, 1998).

A partir da sequência do genoma, fica evidente que *M. tuberculosis* tem o potencial de alternar entre rotas metabólicas, incluindo crescimento aeróbico (por exemplo, a fosforilação oxidativa) e respiração anaeróbica (por exemplo, redução de nitrato). Outra característica do genoma é a presença de genes para a síntese e utilização de quase todos os tipos de lipídios, de ácidos graxos a moléculas complexas, tais como os ácidos graxos de cadeia longa denominados de ácidos micólicos. No total, existem 250 genes que codificam para enzimas envolvidas no metabolismo de diferentes ácidos graxos, em comparação com apenas 50 no genoma da *E. coli* (Cole *et al.*, 1998).

Entre as famílias gênicas mais interessantes, encontram-se as codificantes para as proteínas PE e PPE, que representam quase 10% do genoma. Acredita-se que estas proteínas desempenham um papel importante na sobrevivência e multiplicação de micobactérias em diferentes ambientes (Marri *et al.*, 2006).

A sequência genômica torna-se também crucial na identificação de alvos para quimioterapia. Salvo algumas exceções, existem alguns indícios de que isolados clínicos atuais de *M. tuberculosis* são mais virulentos para o homem do que as cepas isoladas há 100 anos atrás (Cole *et al.*, 1998).

1.5. *M. bovis* BCG: origem e plasticidade genômica

A vacina *M. bovis* BCG é a única vacina atualmente disponível contra a tuberculose. A eficácia dessa vacina varia em diferentes populações. Existem várias hipóteses para explicar essa variabilidade na proteção: variabilidade nas cepas de BCG (Brosch *et al.*, 2007; Behr, 2001), padrões genéticos distintos das populações, infecção prévia com *M. tuberculosis*, contato com micobactérias ambientais e viabilidade das bactérias nas diferentes preparações vacinais (revisto por Brandt *et al.*, 2002).

O BCG foi obtido a partir de um isolado clínico de *Mycobacterium bovis*, no início do século XX por Albert Calmette e Camille Guérin, no Instituto Pasteur de Lille na França, a partir do leite de uma vaca com mastite tuberculosa. Durante 13 anos (1908-1921) estes pesquisadores realizaram sub-cultivos da bactéria embebida em meio contendo bile bovina (como detergente), de forma a obter uma suspensão bacteriana sem grumos. Após 230 passagens sucessivas *in vitro* obtiveram culturas com alterações morfológicas das colônias e perda gradual da virulência. Esta amostra atenuada foi batizada como BCG (Bacilo de Calmette e Guérin), e foi utilizada com sucesso na imunização ativa de vacas, sem produzir a doença (Sakula, 1983). O mecanismo exato da atenuação permanece não esclarecido. Em julho de 1921, a vacina foi administrada pela primeira vez em humanos. Um bebê cuja mãe havia falecido de tuberculose poucas horas após o parto, recebeu três doses de 2 mg (total de 6 mg; $\sim 2,4 \times 10^8$ UFC) via oral (Bonah, 2005). Não foram observados efeitos colaterais deletérios e, mais importante, a criança não desenvolveu a tuberculose.

Pode-se dizer que as vacinas BCG atualmente disponíveis passaram por duas fases de atenuação. A primeira, entre 1908-1921, período em que ocorreu a atenuação da cepa de *M. bovis* após as 230 passagens *in vitro* realizadas por Calmette e Guérin. A segunda fase foi iniciada por volta de 1924, com a distribuição mundial da vacina BCG. Esta fase foi concluída várias décadas depois, e centenas de passagens mais tarde com a criação dos lotes sementes. Como consequência da fase inicial, as cepas de BCG compartilham mutações em comum, enquanto que a segunda fase resultou em mutações adicionais cepa-específicas. Em outras palavras, as cepas de BCG atualmente disponíveis apresentam mecanismos distintos de

atenuação. Assim, é essencial considerar a virulência, a eficácia e os mecanismos de atenuação dentro do contexto de cada cepa de BCG (Liu *et al.*, 2009).

Lotes vacinais deste “BCG” original foram distribuídos para diversos países no mundo a partir de 1921 e como na época não existiam técnicas como liofilização ou criopreservação, não era possível a preservação de lotes-semente, e a manutenção da vacina era feita através de cultivos pelos laboratórios produtores. Em 1961, o BCG (hoje conhecido como cepa Pasteur 1173) foi liofilizado na França. O BCG chegou ao Brasil na década de 1920, trazido por um médico uruguaio, Dr. Moreau (revisto por Benévolo-de-Andrade *et al.*, 2005). A vacina vem sendo produzida desde então e, atualmente, a Fundação Atilafho de Paiva (FAP) fornece cerca de 17 milhões de doses da vacina intradérmica (BCG Moreau) por ano (http://www.bcgfap.com.br/prod_bcg.htm). Embora já existissem relatos na literatura quanto as variações entre as diferentes cepas de BCG utilizadas nos diferentes países, a partir de 1999 foram publicadas análises mais detalhadas, demonstrando que vacinas BCG de diferentes origens passaram por alterações genéticas, sendo proposta uma “árvore genealógica” para as diferentes cepas (Behr *et al.*, 1999; Oettinger *et al.*, 1999; Behr, 2002; Mostowy *et al.*, 2003) (Figura 1.4).

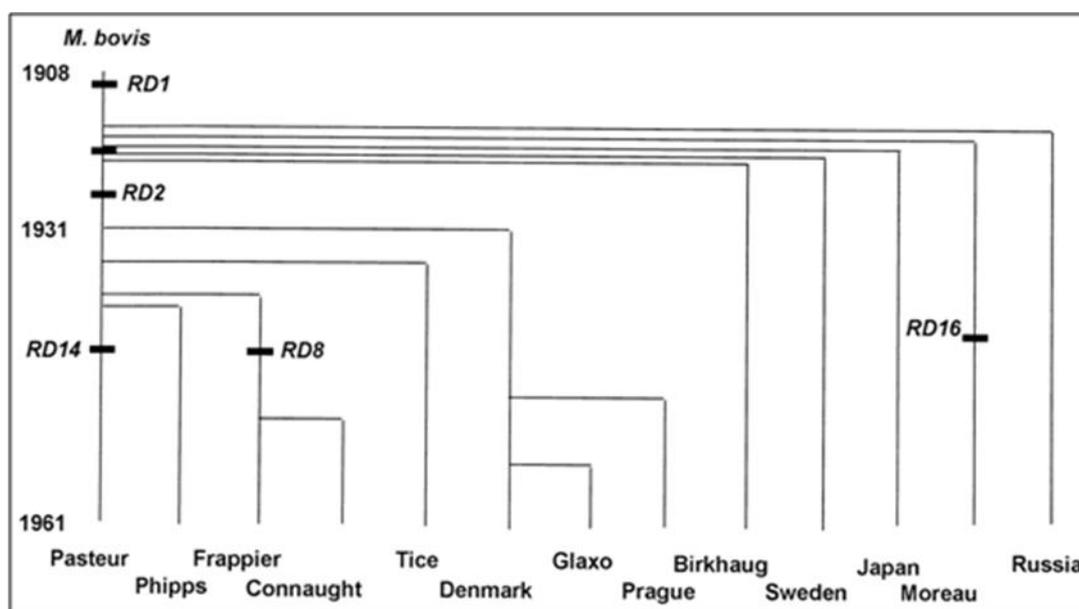


Figura 1.4. Genealogia das vacinas BCG: O esquema apresenta as regiões de diferença (RD) das vacinas (Behr *et al.*, 1999).

Exemplos da correlação entre diferenças genômicas e propriedades biológicas de cepas de BCG já existem. A perda da região RD1 (comum a todas as cepas de BCG) foi associada à diminuição de virulência (Pym *et al.*, 2002); estudos transcricionais também indicam que a perda dessa região tem implicações na expressão de vários genes não

relacionados (Mostowy *et al.*, 2004). O impacto de mutações pontuais (SNPs - *single nucleotide polymorphisms*) é evidenciado pela descrição de uma mutação pontual no gene *mma3* que resulta na perda de produção de metoximicolatos em cepas derivadas a partir de 1927 (Behr *et al.*, 2000) ou pela mutação pontual identificada em algumas cepas de BCG que afeta o domínio de ligação a DNA da proteína CRP, um regulador global de transcrição, podendo contribuir para os mecanismos de atenuação em BCG (Spreadbury *et al.*, 2005).

1.6. *M. bovis* BCG Moreau: cepa vacinal brasileira

A cepa BCG Moreau é a única que não possui a região de diferença chamada de RD16 em *M. tuberculosis* H37Rv (Behr & Small, 1999) (**Figura 1.5**). Além disso, essa cepa não possui algumas deleções encontradas em cepas mais recentes (como a deleção da região RD2, comum a muitas das cepas utilizadas na produção da vacina BCG). O sequenciamento genômico completo é o único método disponível capaz de identificar (numa escala global) pequenas deleções, duplicações e mutações pontuais, fontes importantes de variabilidade genômica que estão além do poder de detecção de outras técnicas empregadas no estudo da variabilidade. Assim, com o objetivo de obter o mapeamento detalhado de todas as inserções/deleções e mutações pontuais em BCG Moreau, nosso laboratório em convênio com a FAP, sequenciou o genoma completo da cepa vacinal brasileira (Gomes *et al.*, 2011). Neste trabalho foi confirmada a presença da duplicação “in tandem” DU2-I, a presença da deleção específica de Moreau (*fadD26-ppsA*), a região de diferença RD16 (de 7,608 pb) e a deleção em fase encontrada em *rv3887c* (*eccD2*). A análise comparativa entre as sequências genômicas de *M. bovis* BCG Moreau, BCG Pasteur e *M. bovis* vem permitindo também a identificação de diversas mutações não-descritas.

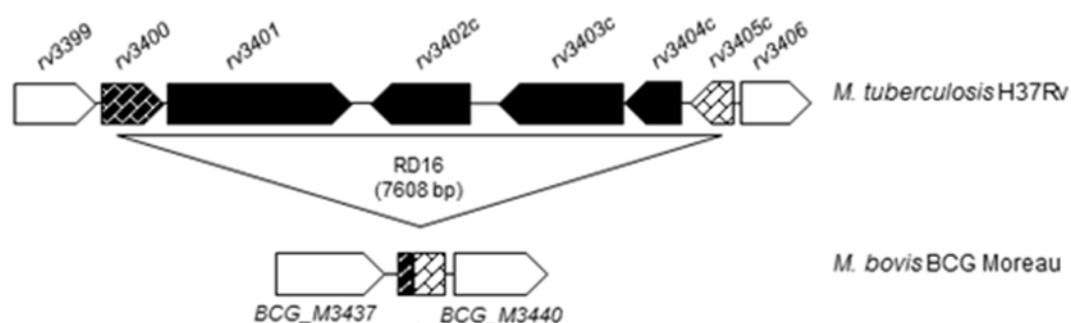


Figura 1.5: Representação esquemática da região RD16 em *M. tuberculosis* H37Rv e do gene *rv3405c* (00-05) que aparece truncado em *M. bovis* BCG Moreau como resultado da deleção da região RD16. A região RD16 de BCG Pasteur é idêntica à de *M. tuberculosis* H37Rv.

Em paralelo, o laboratório determinou também o perfil proteômico de superfície celular (Pagani *et al.* manuscrito em preparação) e de proteínas secretadas da cepa BCG Moreau (Berrêdo-Pinho *et al.*, 2011) comparando com a cepa BCG Pasteur (cujo genoma já havia sido sequenciado anteriormente, no Instituto Pasteur, Paris; Brosch *et al.*, 2007). Os dados analisados demonstram uma grande conservação do perfil proteômico da fração de proteínas associadas à superfície entre as cepas de BCG Moreau (mais “antiga”) e Pasteur (mais “recente”), com poucas diferenças significativas detectáveis através da metodologia que foi padronizada no laboratório. O estudo confirmou algumas diferenças já relatadas na literatura, como as proteínas Mpt64 e Mpt70 e algumas ainda não descritas como diferenciais entre as cepas de BCG (Pagani *et al.*, manuscrito em preparação). Dentre as proteínas diferenciais visualizadas na análise da fração de superfície, foi identificada a presença da proteína Rv3406 em BCG Moreau e sua ausência em Pasteur. Berrêdo-Pinho *et al.*, 2011 demonstraram que o Ag85B (FbpB) foi encontrado mais abundante no filtrado de cultura de BCG Moreau do que em BCG Pasteur. Esta proteína induz proteção parcial contra TB em modelo animal e é considerada como um importante antígeno imunodominante e candidato a vacina. As proteínas MPT64 (Rv1980) e CFP21 (Rv1984c) foram somente encontradas em BCG Moreau e ambas são descritas na literatura como antígenos imunogênicos e potenciais alvos para o desenvolvimento de novas vacinas contra TB. Os genes que codificam essas proteínas estão localizados na RD2, uma região de diferença deletada em um número de cepas de BCG mais recentes incluindo BCG Pasteur, mas presente em BCG Moreau. Os baixos níveis de MPB70 e MPB83 em BCG Pasteur foram também confirmados neste estudo e a expressão reduzida deles deve-se a uma mutação pontual no códon de iniciação do gene *sigK* (Charlet *et al.*, 2005), observado em muitas cepas de BCG, mas ausente em BCG Moreau. MPB63, proteína altamente imunogênica, apresenta expressão aumentada em BCG Moreau em relação a Pasteur. Outras características diferenciais tanto a nível de expressão e/ou localização de proteínas podem ter impactos na interação com o sistema imune tendo como resultado a variação na eficácia das vacinas.

1.7. Regulação da expressão gênica em micobactérias

Em procariotos, o controle da expressão gênica é exercido primariamente ao nível de iniciação de transcrição, mediada pela enzima RNA Polimerase (RNAP). Essa enzima contém 5 subunidades principais e uma subunidade dissociável chamada de fator sigma (σ), que é responsável por conferir à RNAP capacidade de reconhecer sequências promotoras. Assim, através de mudanças no tipo de fator sigma disponível a célula bacteriana consegue mudar

globalmente seu perfil de expressão gênica, possibilitando respostas rápidas e específicas a diferentes condições ambientais. Em *M. tuberculosis* a expressão temporal de diferentes regulons em resposta, por exemplo, a diferentes estágios da interação patógeno-hospedeiro é controlada por diversos fatores sigma (revisto por Sachdeva *et al.*, 2010)

O gênero *Mycobacterium* inclui mais de 150 espécies (Dai *et al.*, 2011) com diferentes características fenotípicas, desde saprófitas ambientais de crescimento rápido a patógenos intracelulares de crescimento lento, como *M. tuberculosis*. A variação na expressão gênica certamente contribui para essas características, permitindo que as bactérias respondam de forma diferenciada às mudanças nas condições ambientais (Newton-Foot & van Pittius, 2012). Apesar das semelhanças com *E. coli*, os componentes dos promotores micobacterianos são bem mais variáveis em sequência; como exemplo, não há conservação da região -35. Isto pode ser consequência do elevado conteúdo de G + C dos genomas micobacterianos, ou do maior número de fatores sigma (σ) presentes em micobactérias enquanto apenas 7 fatores sigma (σ) são conhecidos em *E. coli*, capazes de conferir especificidade a diferentes sequências de promotores (revisto por Newton-Foot & van Pittius, 2012).

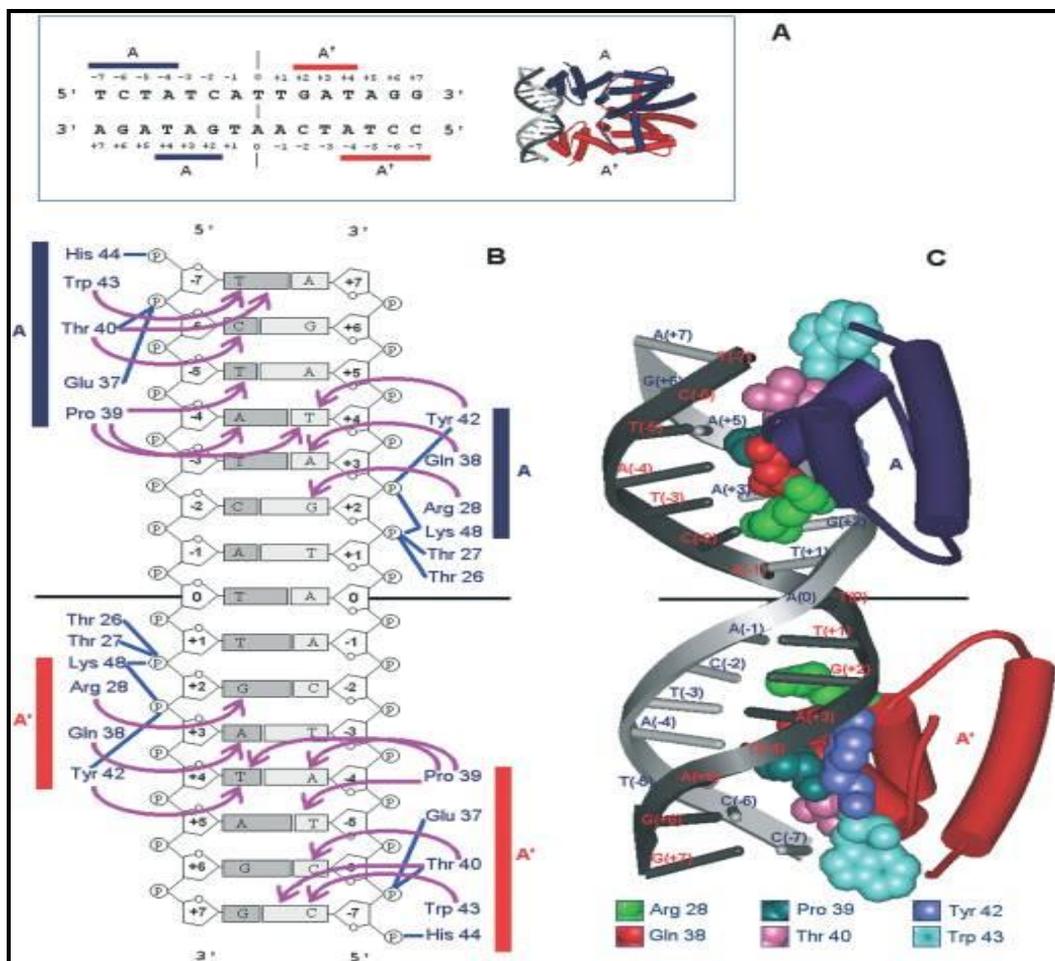
1.8. Reguladores do tipo TetR: *rv3405c*

O gene *rv3405c* é conservado em micobactérias; a análise da sua sequência proteica deduzida revela que contém uma região “helix-turn-helix” (HTH) do tipo encontrado em reguladores transcricionais da família TetR, que comumente atuam como repressores.

O foco deste projeto é a caracterização do produto do gene *rv3405c*, um regulador transcricional putativo, que se encontra truncado em BCG Moreau como resultado da deleção da região RD16. A região RD16 compreende os genes *rv3400* a *rv3405c*, e está representada no esquema da **Figura 1.5**. Estes dois genes, que flanqueiam a deleção, são truncados ao ter uma porção deletada pela perda da região RD16. O gene *rv3405c* intacto tem 567 pb, codificando uma proteína de 188 aminoácidos (**Figura 1.6**). Sua truncagem ocorre na posição 294 pb (98 aminoácidos), sendo que mais 34 aminoácidos são incorporados provenientes do gene *rv3400* formando o gene quimérico *rv3405c* (00-05). A proteína resultante tem seu domínio de ligação ao DNA (DBD - “DNA binding domain”) intacto (aminoácidos 23-69, de acordo com dados do Pfam), mas o domínio C-terminal, necessário para dimerização e ligação ao ligando (Yu *et al.*, 2010) está ausente. Imediatamente a jusante da deleção de RD16 (Behr *et al.*, 1999; Gomes *et al.*, 2011) está *rv3406*, codificando uma taurina dioxigenase putativa. Ainda que a RD16 inclua outros quatro genes, ausentes em BCG

Moreau, a alteração em um regulador transcricional é notável visto seu potencial em alterar uma série de outros processos por meio de mudanças no perfil de expressão gênica.

Esta classe de reguladores TetR pode possuir regulons compostos de muitos genes, podendo também autoregular-se. Assim, é comum que a região intergênica entre o regulador e o gene adjacente inclua a região promotora de ambos, onde se encontra o operador, normalmente composto por sequências palindrômicas. O regulador se une ao operador através do seu domínio N-terminal, na forma de um homodímero, geralmente bloqueando a interação da RNA polimerase com a região promotora (**Figura 1.6**). O complexo proteína-DNA se desfaz quando ocorre a ligação do indutor ao domínio C-terminal, levando a uma mudança conformacional no regulador e liberando a transcrição (Ramos *et al.*, 2005, Yu *et al.*, 2010).



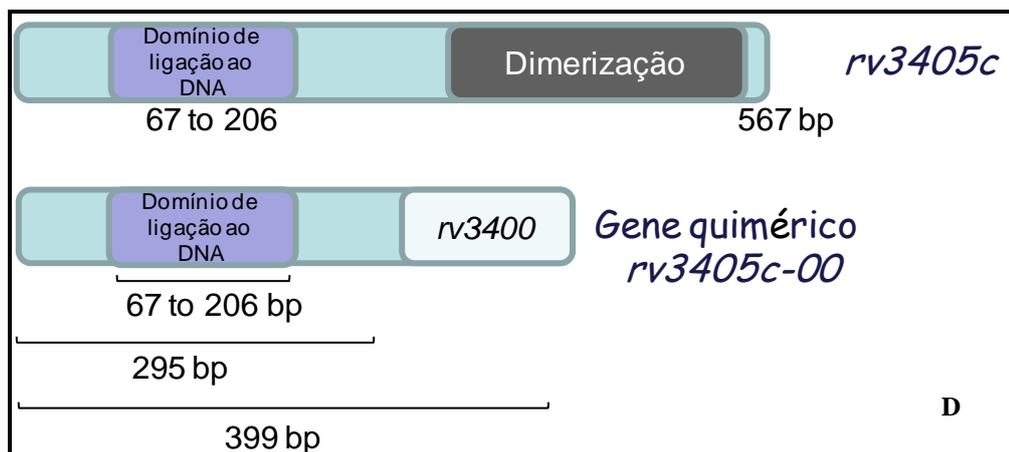


Figura 1.6: Ligação de TetR no seu sítio operador. **A-** operador tetR e regiões de contato. O operador tetR é uma sequência palindrômica. Barras horizontais mostram os nucleotídeos em contato com cada monômero do dímero TetR. **B-** A interação dos resíduos TetR com nucleotídeos específicos (setas) e esqueleto de fosfato (linhas azuis) na região de operador. Os aminoácidos envolvidos na ligação de DNA ficam entre os resíduos 27-48. **C-** Representação de cada homodímero ligado ao operador TetR numa representação de dupla hélice (Ramos *et al.*, 2005). **D-** Representação do gene *rv3405c* em BCG Pasteur e do gene quimérico *rv3405c-00* em BCG Moreau.

Exemplo de um regulador importante, do tipo TetR, é o KstR que possui um regulon grande necessário ao processo de patogênese em *M. tuberculosis* (Kendall *et al.*, 2007; Kendall *et al.*, 2010; Uhía *et al.*, 2011). Outro exemplo envolve um operon relevante ao metabolismo de carbono: é regulado pelo repressor *rv1773* (Alexander *et al.*, 2007), que encontra-se truncado em BCG Pasteur devido à deleção da RD14, tal como ocorre com *rv3405c* na RD16 em BCG Moreau. Repressores da família TetR atuam sobre a expressão de produtos gênicos envolvidos na resistência a múltiplas drogas, enzimas envolvidas em vias catabólicas diversas, biossíntese de antibióticos, *stress* osmótico e patogenicidade de bactérias gram-positivas e gram-negativas (Ramos *et al.*, 2005).

Há evidência de um papel de *rv3405c* na morfologia e proliferação *in vivo*, onde existe correlação entre uma deleção de 22 pb no gene *rv3405c* da cepa BCG Japão e maior eficácia de crescimento/colonização em órgãos de camundongos em contraste com a cepa BCG Japão contendo *rv3405c* intacta. Portanto, a deleção parcial de *rv3405c* pode implicar em alteração na expressão dos genes por ele regulados com subsequente mudança na fisiologia de BCG Moreau (Honda *et al.*, 2006).

Recentemente foi demonstrado que as cepas Moreau, Japão e Glaxo não produzem dimicolatos de ftiocerol (PDIMs) ou glicolipídios fenólicos (PGLs), lipídios de parede celular com importância reconhecida na virulência de *M. bovis* e *M. tuberculosis*, sugerindo que essas cepas são mais atenuadas do que outras. Foi sugerido que o *rv3405c* está envolvido na síntese de PGLs e que a deleção desse gene causa deficiência nessa biossíntese (Chen *et al.*, 2007).

Visamos empregar abordagens de genética molecular bacteriana para caracterizar a proteína codificada por *rv3405c*, definindo sua união ao DNA e seu papel funcional como regulador do gene *rv3406* adjacente, onde foi comprovada a presença da proteína Rv3406 na parede celular de BCG Moreau e sua ausência em BCG Pasteur (Pagani *et al.*, manuscrito em preparação), reforçando a hipótese do papel funcional do gene *rv3405* na regulação da expressão de Rv3406.

2. OBJETIVOS

2.1. Objetivo geral:

- Realizar uma análise funcional do gene *rv3405c* (fator de transcrição putativo) através de uma análise genômica comparativa entre *M. bovis* BCG Pasteur e *M. bovis* BCG Moreau e investigar o possível impacto da perda funcional deste gene em BCG Moreau.

2.2. Objetivos específicos:

- Analisar a conservação da região intergênica entre *rv3405c* e *rv3406* em espécies de micobactérias.
- Determinar o perfil transcricional do gene *rv3405c* e do gene adjacente (*rv3406*) nas cepas BCG Moreau e Pasteur sob diferentes condições de cultivo, e se ligantes o alteram.
- Determinar por RT-PCR se o gene quimérico *rv3400-05c* é transcrito em ambas as direções em BCG Moreau: do promotor de *rv3400* (*rv3400-05c*) e do promotor de *rv3405c* (*rv3405c-rv3400*).
- Determinar por RT-PCR se os genes *rv3400-05c* e *rv3406* são co-transcritos em BCG Moreau.
- Determinar por RT-PCR e ensaio de retardo em gel se Rv3405c é um repressor da transcrição e se sua união ao DNA é modulada na presença de possíveis indutores (compostos presentes na bile bovina).
- Caracterizar funcionalmente as proteínas Rv3405c e Rv3406:
 - Expressar Rv3405c (de BCG Pasteur) e Rv3406 (de BCG Moreau) sob forma recombinante em *E. coli*;
 - Purificar as proteínas recombinantes;
 - Obter anticorpos policlonais em camundongos Balb/C contra as proteínas recombinantes purificadas;
 - Analisar a presença/ausência de Rv3406 em frações proteicas de BCG Moreau e Pasteur, sob diferentes condições de cultivo (com ou sem a presença dos

possíveis moduladores), utilizando o anticorpo anti-rRv3406 em ensaios de *western blot*;

3. METODOLOGIA

Os materiais utilizados estão descritos no Anexo I.

3.1. Condições de cultivo de cepas bacterianas

3.1.1. *Escherichia coli*

As cepas de *E. coli* TOP 10 e BL21(DE3) (Anexo I) foram utilizadas neste estudo para propagação e clonagem em vetores de expressão e de sequenciamento, sendo cultivadas a 37°C por 16 horas sob agitação, em meio Luria Bertani (LB) suplementado com 100 µg/mL de ampicilina ou 50 µg/mL de canamicina, quando necessário.

3.1.2. Cepas vacinais de *M. bovis* BCG Pasteur e Moreau

A cepa de *Mycobacterium bovis* BCG Pasteur foi obtida da coleção de referência do Instituto Pasteur, Paris, França. No laboratório de Genômica Funcional e Bioinformática da FIOCRUZ, foi mantida em 20% glicerol a -70°C. A cepa de *M. bovis* BCG Moreau foi fornecida em meio Sauton/batata pela Fundação Ataufo de Paiva (FAP). Essas cepas de BCG foram semeadas na superfície de uma batata, em meio Sauton, em tubo de Roux, e incubadas por 15 dias sem agitação a 37°C. Após esse período foi feito um repique para a superfície de meio Sauton líquido, seguido de incubação a 37°C sem agitação por mais 15 dias, para obtenção da “nata” bacteriana. Além disso, essas cepas (Anexo I) também foram cultivadas em 30 mL de meio líquido Bacto® Middlebrook 7H9 (DIFCO) contendo 0,05% (v/v) Tween 80 (DIFCO) suplementado com ADC (Anexo I). As culturas ficaram sob agitação a 37°C até uma DO_{600nm} de 0,8 para BCG Pasteur e DO_{600nm} de 1,0 para BCG Moreau. A partir desse cultivo inicial em 7H9 (ADC) foi feito o lote-semente contendo 1 mL de cultura em 15% glicerol (Sigma). Quando necessário, o lote semente foi cultivado em meio 7H9 (1:10) e acrescido dos seguintes componentes: 2 mM taurina, 0,05% ou 0,1% bile bovina, 100 µM TCDC (tauroquenodeoxicolato de sódio) e 100 µM GCDC (glicoquenodeoxicolato de sódio).

3.2. Obtenção de proteínas do filtrado de cultura de BCG Pasteur e Moreau

A fração de proteínas presentes no filtrado de cultura foi preparada a partir de culturas de 15 dias, estáticas, em meio Sauton. O sobrenadante de cultura foi separado da nata bacteriana e filtrado em membrana de 0,22 µm de baixa ligação de proteínas (0,22 µm GP

Express PLUS, Milipore). Foram feitas alíquotas de 10 mL e o material estocado a -70°C até sua utilização.

3.3. Obtenção das proteínas associadas à superfície de BCG Pasteur e Moreau

Após a remoção do sobrenadante de cultura, os componentes associados à superfície celular foram obtidos por agitação manual da película bacteriana com esferas de vidro de 2 mm de diâmetro por 40 segundos, como descrito por Ortalo-Magné e colaboradores (1995). Em seguida, os componentes de superfície foram recuperados em água milli-Q, submetidos à centrifugação por duas vezes (10 minutos a 2.500 xg , 4°C) para retirar as células bacterianas. Alíquotas de 1,5 mL deste material foram estocadas a -70°C até sua utilização.

3.4. Obtenção do lisado micobacteriano

O sedimento bacteriano resultante após remoção da fração de superfície foi ressuspense em 1 mL de tampão de lise bacteriana (Anexo I) e adicionado de coquetel de inibidores de proteases para uso em extratos bacterianos (diluição 1:100; Sigma/P8465). As células foram completamente rompidas por lise mecânica com 5 ciclos de 1 minuto no equipamento Bead Beater (Biospec Products), com 1 minuto de repouso entre cada ciclo. Em seguida, foram submetidas à centrifugação (1 minuto a 16.000 xg), o sobrenadante recolhido e armazenado a -70°C .

3.5. Precipitação das proteínas presentes nas diferentes frações de BCG

As proteínas foram precipitadas pela adição de TCA na concentração final de 17% seguida de mistura e incubação a -20°C por 10 minutos e centrifugação por 10 minutos a 16.000 xg . O sobrenadante foi descartado e o sedimento lavado sequencialmente com 300 μL de acetona gelada contendo 1% trietanolamina (TEA) e duas vezes com 300 μL de acetona 100% gelada. O sedimento final foi ressuspense em 20 μL de tampão de focalização isoeletrica (tampão IEF - Anexo I).

3.6. Determinação da concentração proteica

A concentração das amostras proteicas foi determinada utilizando o kit RCDC (BioRad) segundo as especificações do fabricante. Este método é uma adaptação do protocolo de Lowry (1951), que é compatível tanto com agentes redutores (RC) quanto detergentes (DC).

3.7. SDS-PAGE

Uma quantidade de 5 µg de proteína foi misturada ao tampão de amostra de proteína 1x (Anexo I) para um volume final de 20 µL. A mistura foi em seguida aquecida a 100°C por 5 minutos. As amostras foram submetidas à eletroforese em géis 15% SDS-PAGE (Anexo I) de 7 centímetros em tampão de corrida (Laemmli, 1970) a 200 volts empregando o sistema MiniProtean (Bio-Rad). Posteriormente, os géis passaram em sua maioria por coloração por Coomassie Brilliant Blue R-250; em alguns casos foi feita uma coloração pela prata (Anexo I). O marcador de peso molecular utilizado foi o SDS-PAGE Standard Low Range (Bio-Rad) ou Standard Broad Range (Bio-Rad). Para a captura de imagens dos géis corados foi utilizado o densitômetro GS-800 (BioRad).

3.8. Extração de DNA genômico das culturas de *M. bovis* BCG Pasteur e Moreau

A partir de 5 mL de BCG Pasteur e BCG Moreau de culturas estáticas foram transferidos 1 mL de cada cultura para tubos eppendorf e em seguida centrifugados por 10 minutos a 1000 g a temperatura ambiente. As células foram então ressuspensas em tampão PBS 1x (Anexo I) e novamente centrifugadas. O sedimento bacteriano foi ressuspenso em tampão de lise (25% sacarose, 50 mM Tris-HCl pH 8,0 e 10 mM EDTA) e transferido para um recipiente de lise contendo pérolas de vidro de 0,1 mm. As células foram lisadas em equipamento Bead Beater, com 3 pulsos de 1 minuto. O lisado obtido, após decantação das pérolas, foi então transferido para novos tubos e incubado com lisozima a uma concentração final de 0,5 µg/µL. Após incubação a 37°C por 1 h, dodecil sulfato de sódio (SDS) e protease K, respectivamente, com concentrações finais de 1% e 0,5 µg/µL, foram adicionados e o material incubado a 55°C por 1h. Os sobrenadantes foram transferidos para microtubos de 1,5 mL e submetidos ao método de extração e purificação por fenol-clorofórmio saturado em Tris-HCl 0,1 M pH 8,0 seguido por uma etapa de mistura e centrifugação à temperatura ambiente por 5 min a 2,250 xg. O processo foi repetido até que não houvesse mais presença de material proteico na interface. Este procedimento foi repetido por mais uma vez utilizando apenas extração com clorofórmio/álcool isoamílico (24:1). Ao sobrenadante coletado foi

então adicionado cloreto de sódio 5 M para uma concentração final de 1 M e 2 volumes de etanol absoluto para precipitação do DNA. Após homogeneização, o material foi incubado a 4°C por 16h e centrifugado por 15 min a 12.000 xg. Após essa etapa o precipitado foi lavado 3 vezes com etanol 70% gelado. O DNA precipitado foi ressuspensão em 100 µL de TE (Anexo I). A concentração foi determinada pela leitura da absorbância a 260 nm.

3.9. Clonagem, expressão e purificação de proteínas recombinantes em *Escherichia coli*

Duas proteínas (Rv3405c de BCG Pasteur e Rv3406 de BCG Moreau) foram selecionadas para estudos mais detalhados e expressas, sob forma recombinante, em *E. coli*. Para tal, plasmídeos para subclonagem e clonagem foram selecionados (Anexo I), e os oligonucleotídeos (Anexo II) foram desenhados e sintetizados para amplificação dos genes de interesse.

3.9.1. Preparo de células competentes de *E. coli* para eletroporação

Células da cepa TOP10 e da cepa BL21(DE3) de *E. coli* criopreservadas em nitrogênio líquido foram inicialmente estriadas em meio LB/ágar e incubadas a 37°C por 16 horas de modo a obter colônias isoladas. Para cada cepa uma colônia isolada foi inoculada em 5 mL de meio LB e incubada a 37°C, sob agitação, por uma noite. No dia seguinte, o pré-inóculo foi utilizado para inocular 250 mL de meio LB numa diluição de 1:100 e a cultura incubada a 37°C, sob agitação, até atingir uma DO_{600nm} entre 0,4 e 0,6. A cultura foi então mantida em banho de gelo por no mínimo 30 min e, em seguida, centrifugada a 1,000 xg por 15-20 min a 4°C. O sobrenadante foi descartado e o sedimento lavado três vezes com 250 mL e 125 mL de água Milli-Q estéril e com 125 mL de glicerol 10% estéril (todas mantidas a 4°C), respectivamente, nas mesmas condições de centrifugação anteriormente descritas. Após a última centrifugação o sedimento foi ressuspensão em 2,5 mL de glicerol 10% estéril. A suspensão bacteriana foi distribuída em alíquotas de 50 µL, com congelamento imediato em banho de gelo seco/etanol e posteriormente foram mantidas a -80°C até o momento do uso.

3.9.2. Transformação por eletroporação em *E. coli*

Após o tempo de incubação necessário para a ligação do inserto ao vetor, 2 µL da reação foram utilizados para transformar 50 µL de células eletrocompetentes, em cuvetas de 0,2 cm utilizando o eletroporador Gene Pulser (BioRad). Os parâmetros utilizados para eletroporação foram: 2,5 kV, 25µF e 200Ω. Após a eletroporação, a mistura foi imediatamente transferida para um tubo contendo 1 mL de LB e incubada por uma hora a 37°C sob agitação lenta. Posteriormente, as células foram plaqueadas em LB/ágar, contendo o antibiótico apropriado para seleção de transformantes, e incubadas a 37°C por 16h.

3.9.3. Amplificação das regiões codificadoras de *rv3405c* e *rv3406* por PCR

As reações de PCR foram realizadas em um volume final de 50 µL, utilizando o Master Mix 2x (Promega), 200 ng de cada oligonucleotídeo específico (Senso e Antisenso) e 5 µL de uma diluição 1:100 de DNA genômico de *M. bovis* BCG Pasteur ou BCG Moreau. A ciclagem utilizada para as reações de PCR estão especificadas no Anexo III.

3.9.4. Purificação de produtos de PCR

Os fragmentos de PCR obtidos correspondentes aos genes *rv3405c* de *M. bovis* BCG Pasteur e *rv3406* de *M. bovis* BCG Moreau foram purificados através da utilização do Wizard PCR Preps DNA Purification System (Promega) segundo recomendações do fabricante.

3.9.5. Extração e sequenciamento de DNA plasmidial

DNA plasmidial foi extraído a partir de uma cultura saturada de 5 mL de *E. coli* TOP 10 portadora da construção no vetor pBAD TOPO[®]TA, utilizando-se o Kit Miniprep DNA Purification System (Promega[®]), segundo especificações do fabricante. O DNA plasmidial obtido foi utilizado para sequenciamento de DNA, com o kit de sequenciamento BigDye 3.1 (Applied Biosystems[®]), sendo as reações analisadas em sequenciador capilar ABI3730 pela Plataforma de Sequenciamento de DNA PDTIS/FIOCRUZ.

3.9.6. Subclonagem no vetor pBAD TOPO[®] TA

Para a clonagem das regiões codificantes (*rv3405c* e *rv3406*) no vetor de expressão optamos por primeiro subclonar o fragmento de PCR utilizando o “kit” “TOPO[®]TA Cloning” (Invitrogen) e o vetor pBAD TOPO[®]TA (Invitrogen), seguindo as especificações do fabricante (Tabela 3.1). A clonagem TOPO[®]TA é uma estratégia de clonagem rápida que dispensa o uso de T4 DNA ligase, não necessitando de longos períodos de incubação uma vez que é realizada à temperatura ambiente por no máximo 30 min.

Tabela 3.1: Reação de clonagem TOPO[®]TA

REAGENTE	VOLUME
H ₂ O milli-Q autoclavada	2 µL
Solução de sal diluído ¼	1 µL
Produto do PCR purificado	2 µL
Vetor pBAD TOPO [®] TA	0,5 µL

3.9.7. Digestão e ligação para clonagem em pET28a

Os oligonucleotídeos desenhados para essa clonagem conforme descrito no item 3.9 (Anexo II) possuem em suas pontas sítios de restrição para digestão com as enzimas *Nco*I e *Xho*I. A digestão do DNA do vetor de expressão pET28a e dos vetores pBAD05/pBAD06 com estas enzimas gera pontas coesivas que podem ser ligadas pela T4 DNA ligase (Promega), permitindo a construção dos plasmídeos pET-P3405c (pET28a + *rv3405c*) e pET-M3406 (pET28a + *rv3406*). Conforme descrito acima, para realização dessa clonagem optamos por fazer uma subclonagem prévia dos produtos de PCR para os genes *rv3405c* e *rv3406* no vetor pBAD TOPO[®]TA, facilitando a digestão com as enzimas especificadas acima. Seguimos as especificações do fabricante para cada enzima para realização das digestões e após purificação do vetor e dos insertos utilizando o kit de purificação para produto de PCR (Promega), fizemos a ligação utilizando o protocolo padrão da T4 DNA ligase (Promega) respeitando a proporção de 5:1 de inserto e vetor analisados em gel de agarose pela visualização das bandas. Clones contendo insertos foram identificados por PCR, minipreparação de DNA, digestão e sequenciamento.

3.9.8. Transformação da cepa BL21(DE3) e indução da expressão de proteínas recombinantes com IPTG

DNA plasmidial das construções obtidas e analisadas no item 3.9.8 acima foi utilizado para transformar a cepa de expressão de *E. coli* BL21(DE3). Colônias isoladas após a transformação foram inoculadas em 3 mL de meio LB/Kanamicina e incubadas a 37°C por 16 horas, sob agitação. Para indução da expressão 200 µL de cada cultura foi adicionada em 20 mL de meio LB/Kanamicina. Após atingir uma DO_{600nm} de 0,5, 10 mL foram separados para indução (I) e o resto da cultura utilizado como controle não induzido (NI). Após adição do indutor IPTG numa concentração de 1 mM, as culturas foram incubadas por mais 3 horas a 37°C sob agitação. Essas condições de concentração, temperatura e tempo de indução podem variar na tentativa de obter melhor expressão da proteína de interesse. Neste caso, testamos também as alternativas de concentração de 0,1 mM IPTG, incubação à temperatura ambiente e tempo de indução de 16h. Ao final da incubação medimos a DO_{600nm} e 1 mL de cada cultura foi transferida para microtubo e centrifugada a 16.000 xg por 5 minutos. Cada amostra foi ressuspensa num volume final de tampão de amostra de modo a obter o equivalente a 0,1 DO_{600nm} da cultura original por 10 µL de tampão para SDS-PAGE. As amostras foram fervidas por 5min, centrifugadas e 10 µL aplicados em gel SDS-PAGE.

3.9.9. Teste de solubilidade das proteínas recombinantes

Os sedimentos obtidos após centrifugação das amostras de 1 mL de cultura (induzida; ver item 3.9.8, acima) foram ressuspensos em 1 mL de tampão de lise bacteriana (Anexo I) e transferidos para um microtubo contendo 500 µL de pérolas de vidro de 0,1 mm. As células foram lisadas mecanicamente em equipamento Bead Beater (Biospec Products INC.), com 3 pulsos de 1 minuto. Do lisado obtido, após decantação das pérolas de vidro, foram separados 200 µL o qual foi chamado de lisado total (condição A); o restante (800 µL) foi clarificado por centrifugação por 10 min (16.000 xg), obtendo-se assim uma fração contendo proteínas solúveis no sobrenadante e outra fração contendo os resíduos celulares e proteínas insolúveis (em corpos de inclusão) no sedimento. As proteínas contidas no lisado total e na fração solúvel após clarificação foram precipitadas pela adição de TCA (final 17% v/v), conforme já descrito acima (item 3.5) e estas amostras, assim como o sedimento obtido após clarificação do lisado, ressuspensos em tampão SDS-PAGE para posterior análise por SDS-PAGE, mantendo-se a relação de 0,1 DO_{600nm} da cultura original para cada 10 µL de tampão.

3.9.10. “Western blot” (WB)

As proteínas resolvidas em géis SDS-PAGE (15%) foram eletro-transferidas para membrana de nitrocelulose (Hybond-C, GE) em equipamento Mini Transblot (BioRad), a 100 V durante uma hora, segundo especificações do fabricante. Após a transferência, a membrana foi corada com Ponceau S (Sigma) ou MemCode (Pierce) para visualização e verificação da qualidade da transferência, e em seguida bloqueada em 10% leite desnatado/TBS 1x (Anexo I) por uma noite a 4 °C com agitação. Após o bloqueio a membrana foi lavada com TBS 1x (3x por 5 min) e, em seguida, incubada durante 2 horas, com agitação suave, com o anticorpo primário na diluição apropriada, conforme especificado em resultados, em 5% leite desnatado/TBS 1x.

A membrana foi então lavada com TBS 1x/Tween 0,05% (3x por 5 min) e TBS 1x (3x por 5 min) antes da adição do anticorpo secundário anti-IgG de camundongo conjugado à peroxidase (Pierce), diluído 1:10.000 em 5% leite desnatado/TBS 1x, incubada por mais 1 hora, com agitação suave e lavada como descrito acima. Para revelar utilizamos o Kit Super Signal[®] (Pierce) conforme as instruções do fabricante.

3.9.11. Purificação das proteínas recombinantes a partir de corpos de inclusão

A partir de um pré-inóculo de 3 mL de uma cultura sob agitação a 37°C foi medida a DO_{600nm} e quando atingiu 0,5 uma alíquota foi retirada para correr em gel SDS-PAGE (cultura não induzida) e, posteriormente, a cultura foi induzida por 3 horas a 37°C sob agitação utilizando uma concentração final de 1 mM de IPTG. Após indução foi medida a DO_{600nm} e uma alíquota foi retirada para corrida em gel, 200 μ L da cultura foi adicionada em 20 mL de meio LB/Kanamicina em erlenmeyer. Após atingir uma DO_{600nm} de 0,5 foi retirado 5 mL dessa cultura como pré-inóculo para um volume de 500 mL de cultura em meio LB/Kanamicina. Posteriormente, a cultura foi induzida nas mesmas condições descritas acima. Após centrifugação a 5.000 rpm (rotor JA-14) por 15 minutos a 4°C, o sedimento foi ressuspenso em tampão de lise bacteriana contendo lisozima (concentração final 1 mg/mL) (Anexo I) e as células rompidas por lise mecânica em equipamento Bead-Beater (1 a 3 minutos) no pote de 30 mL (10 mL de beads e 20 mL de suspensão bacteriana a cada utilização do equipamento). Após a lise foi realizada uma clarificação do lisado por centrifugação a 5.000 rpm (rotor JA-14) por 15 minutos a 4°C. O sobrenadante (proteínas solúveis) foi armazenado a -20°C para posterior purificação da proteína recombinante solúvel (descrito abaixo). O sedimento, contendo os corpos de inclusão, foi ressuspenso em 10 mL de tampão de lavagem contendo Triton X-100 (Anexo I) com ajuda de Dounce. Posteriormente,

a amostra foi centrifugada a 10.000 rpm (rotor JA-20) por 15 minutos a 4°C. O processo de lavagem foi repetido por um total de 3 vezes. O sedimento da última lavagem foi ressuspensão em 10 mL de solução desnaturante (Anexo I) e colocado a 4°C sob agitação durante a noite. No dia seguinte o material foi centrifugado a 10.000 rpm por 15 minutos a 4°C e o sobrenadante mantido a 4°C até purificação. Esta purificação ocorreu utilizando o equipamento HPLC com coluna Histrap HP (GE) de 1 mL e eluição em gradiente de imidazol (20mM – 500 mM). Foram utilizados dois tampões: tampão de ligação e tampão de eluição (Anexo I) de forma a obter as diversas frações coletadas através de um gradiente de imidazol que compete com o Ni⁺² adsorvido na coluna. Após verificação das frações coletadas em géis SDS-PAGE, foi feito um *pool* das frações contendo a proteína purificada e este material foi submetido ao procedimento de diálise contra PBS 1x (durante 16h a 4°C). Feito isso, as amostras foram alíquotadas e congeladas a -70°C.

3.9.12. Purificação das proteínas recombinantes solúveis

A partir do sobrenadante obtido no item 3.9.12 acima, demos prosseguimento a precipitação das proteínas solúveis com 85% de sulfato de amônia e posterior purificação em coluna de Sepharose - Ni⁺². O sulfato de amônia foi adicionado aos poucos no sobrenadante, em banho de gelo, com agitação por 2 horas. O material foi centrifugado a 10.000 rpm por 15 minutos a 4°C e o sedimento armazenado a -20°C. Este sedimento foi posteriormente ressuspensão em 10 mL de tampão 10 mM Tris-HCl pH 7,5 /10 mM imidazol /200 mM NaCl /0,05% Tween 20. Este procedimento foi realizado vagarosamente em gelo evitando formação de bolhas. O material foi aplicado em coluna Ni⁺²-IMAC neste mesmo tampão com eluição através de concentrações crescentes de imidazol (50 mM, 100 mM, 150 mM e 200 mM). Dez frações de 500 µL foram coletadas a cada passo e armazenadas em gelo, onde alíquotas de 10 µL foram retiradas para corrida em gel de SDS-PAGE. As frações positivas, visualizadas por gel corado por Coomassie Blue R-250, contendo a proteína purificada foram reunidas formando um pool. Esse pool foi dialisado contra 2 L de tampão 10 mM Tris-HCl pH 7,5/1 mM EDTA/50 mM NaCl/2,5% glicerol/0,05% Tween 20. Após 16h de diálise com agitação a 4°C, glicerol foi adicionado às amostras (16% v/v concentração final) e estas foram então congeladas em alíquotas e armazenadas a -70°C.

3.10. Obtenção de soro policlonal contra as proteínas recombinantes

Camundongos Balb/c (6 fêmeas de 7 semanas) foram imunizados com 10 µg de proteína recombinante (Rv3405c ou Rv3406) em adjuvante incompleto de Freund por via intraperitoneal. Foram realizadas 3 imunizações com intervalos de 15 dias. Ao final do

protocolo de imunização, o sangue total dos animais foi obtido por punção cardíaca, incubado por 30 min a 37 °C e centrifugado para obtenção do soro. Este procedimento foi realizado em conformidade com os princípios que regem o uso científico de animais com protocolo licenciado pela CEUA-FIOCRUZ (LW46/11).

3.11. Titulação dos anticorpos

O soro individual de cada camundongo foi testado por ensaio de WB contra a proteína específica utilizada na imunização. A partir dessa análise foi feito um “pool” de soros que apresentavam eficiência na resposta contra a proteína inoculada para posterior titulação de anticorpos policlonais. Foi feito um gel preparativo de 15% SDS-PAGE para cada proteína a ser analisada, onde foram aplicados cerca de 20 µg da proteína pura dialisada. Posteriormente, essa proteína foi eletro-transferida para membrana de nitrocelulose (Hybond-C, GE) em equipamento Mini Transblot (BioRad), a 100V durante uma hora, segundo especificações do fabricante. Após bloqueio da membrana em 10% leite desnatado/TBS por uma noite e lavagem em TBS, a membrana foi cortada em tiras verticais e cada tira incubada com diluições do soro policlonal (1:1.000, 1:2.000, 1:4.000, 1:8.000, 1:16.000, 1:32.000 e 1:64.000 para o soro anti-Rv3406 e 1:500 e 1:1.000 para o soro anti-Rv3405c).

3.12. “Western blot” para as frações proteicas de *M. bovis* BCG Pasteur e *M. bovis* BCG Moreau

As frações de proteínas associadas à superfície, proteínas intracelulares solúveis e filtrados de culturas obtidas a partir de *M. bovis* BCG Pasteur e Moreau foram testadas em ensaio de WB contra o anticorpo primário anti-Rv3406 (1:1.000) contendo lisado total de *E. coli* a fim de reduzir ligações inespecíficas, e anticorpo secundário anti-IgG de camundongo conjugado à peroxidase (1:50.000). Amostras de aproximadamente 25 µg de proteínas de cada fração e o controle 1 µg da proteína recombinante Rv3406 purificada foram resolvidas em gel 15% SDS-PAGE e eletro-transferidas para membrana de nitrocelulose (Hybond-C, GE) em equipamento Mini Transblot (BioRad) a 100V durante uma hora. O protocolo de revelação utilizado foi o descrito no ítem 3.9.11 acima.

3.12.1. Preparo de lisado de *E. coli*

Utilizando 5 mL de meio LB a 37 °C, foi crescido um pré-inóculo da cepa de expressão de *E. coli* BL21(DE3). Após atingir DO_{600nm} igual a 0,4 foi diluído 1:100 em 500 mL de meio LB. Foi induzida com 1 mM IPTG por 2h após atingir a DO_{600nm} de 0,4, centrifugada a

30.100 xg (Rotor JA-14) por 15 minutos a 4°C e o sedimento foi ressuspenso em 20 mL de tampão de lise bacteriana (Anexo I). Logo em seguida, foi lisado em equipamento Bead Beater (Biospec Products INC.), com 3 pulsos de 1 minuto no pote de 30 mL (10 mL de beads e 20 mL de suspensão bacteriana). Após decantados os “beads”, o lisado foi transferido para tubo falcon e alíquotas de 500 µL foram armazenadas a - 20°C.

3.13. Extração de RNA total de culturas de *M. bovis* BCG Pasteur e Moreau

Utilizando alças estéreis, uma pequena amostra de cultura de *M. bovis* BCG Pasteur e Moreau cultivados sob a forma de “nata” em meio Sauton sem agitação foi coletada e congelada em 1 mL de Trizol® (Invitrogen) a - 70°C para posterior extração de RNA para ensaios de RT PCR. Da mesma forma também foram coletadas amostras de 2 mL cultura de *M. bovis* BCG Pasteur e Moreau cultivadas em meio 7H9/ADC sob agitação. Após descongelamento foram adicionadas pérolas de vidro de 0,1 mm e as células foram lisadas em equipamento Bead Beater (Biospec Products INC.), com 3 pulsos de 1 minuto. A fase aquosa foi retirada para novos tubos eppendorf, foram adicionados às amostras 250 µL de clorofórmio e os tubos agitados em vórtex por 15 segundos e centrifugados a 16.000 xg por 15 min a 4°C. A fase aquosa, contendo o RNA, foi transferida para um novo tubo, e o RNA precipitado pela adição de 70% (v/v) isopropanol seguido de incubação por 16h a -70°C. Posteriormente, o material foi centrifugado a 16.000 xg por 15 min a 4°C e o sobrenadante descartado. O sedimento foi lavado com 500 µL de etanol 75% gelado e novamente centrifugado a 16.000 xg por 15 min a 4°C. Após remoção do sobrenadante, o sedimento foi seco à temperatura ambiente e finalmente ressuspenso em 50 µL de água livre de RNase (H₂O DEPC) (Anexo I) e adicionado de 0,25 µL de RNasin - inibidor de RNase (Promega®). A concentração de RNA total foi estimada a partir da leitura da DO a 260 nm no espectrofotômetro Bio Photometer (Eppendorf) (1 DO_{260nm} equivale a 40 µg/mL de RNA total ou 400 ng/µL). Foram aliqotados 2 µg de RNA total de cada amostra e tratados com RQ-DNase (Promega®) ou com a DNase Turbo (Ambion) conforme indicação do fabricante, de modo a remover qualquer contaminação com DNA genômico. A fim de confirmar se realmente não havia contaminação com DNA genômico, foi feita uma reação de PCR diretamente sobre a amostra de RNA tratado sem utilizar a transcriptase reversa para produção do cDNA. Um resultado negativo no PCR nestas amostras indica que todo DNA foi removido. O RNA total extraído e tratado foi mantido a -70°C para posterior avaliação por RT-PCR semi-quantitativo.

3.14. RT-PCR

Para obtenção do cDNA, as amostras de RNA total foram divididas em dois tubos (+RT e -RT) contendo cada uma 1 µg de RNA. Um tubo (-RT) foi utilizado como controle negativo, onde a adição da transcriptase reversa foi omitida, possibilitando a identificação de contaminação com DNA genômico. A cada tubo foi adicionado 1 µL de iniciador antisense (reverso) específico a 200 ng/µL (Anexo II) e o protocolo original foi seguido conforme instruções do fabricante da transcriptase reversa (ImPromII™ Promega ou Superscript™ III, Invitrogen). Nas reações de RT-PCR foram utilizados 2 µL de cDNA em cada reação de PCR contendo os oligonucleotídeos específicos (200 ng/µL por reação). A ciclagem de PCR varia conforme o transcrito que está sendo analisado (Anexo III).

Para avaliar uma possível contaminação das amostras com DNA genômico, foi feita uma reação de PCR para cada amostra, com visualização dos produtos amplificados em gel de agarose. Somente as amostras +RT poderiam apresentar produtos amplificados. Esta etapa foi utilizada para a análise da expressão de alguns genes envolvidos neste estudo, realizada com RT-PCR semiquantitativo.

3.15. Ensaio de retardo em gel (EMSA)

3.15.1. EMSA

Este ensaio avalia se existe uma interação entre a proteína estudada (Rv3405c) e uma determinada porção do DNA (região intergênica *rv3405-rv3406*). Os oligonucleotídeos iniciadores para esta região foram desenhados de maneira que o senso estivesse biotilado (Anexo II) e o produto de PCR tivesse 101 pb. Foi preparado um gel PAGE 5% (poliacrilamida não desnaturante) em tampão 0,5X TBE (Anexo I). Este tipo de gel mantém intacta a estrutura da proteína nativa. O gel foi pré-corrído por 30 minutos a 100 V. Utilizamos um kit LightShift™ EMSA Optimization & Control (Pierce) para preparo das reações de ligação. Para um volume final de 15 µL foram utilizados: 1,5 µL 10x Binding Buffer, 0,75 µL 50% glicerol, 0,75 µL poli dI-dC, 0,1 µL de DNA biotilado a 0,4 ng/µL, H₂O Milli-Q estéril e proteína recombinante purificada Rv3405c em diferentes concentrações (4 µM / 2 µM / 200 nM / 20 nM / 2 nM / 0,4 nM). O tempo de incubação da reação foi de 45 minutos a temperatura ambiente. Posteriormente adicionamos Loading Buffer (para 1x concentração final), as amostras foram aplicadas no gel e corridas em 0,5x TBE a 100V a temperatura ambiente, até que o bromofenol azul chegasse a metade do gel. O material foi então transferido para membrana Hybond N+ (GE Healthcare) em equipamento Mini

Transblot (BioRad) a 180 mA durante 1 hora. A membrana foi submetida à luz ultravioleta (5 minutos de cada lado) para fixar o DNA na membrana por “crosslinking”. Nesta etapa podemos guardar a membrana a seco a 4°C para revelação posterior. Utilizamos o Kit LightShift™ Chemiluminescent EMSA, com conjugador HRP-streptavidina, para revelação da membrana, seguindo as instruções do fabricante.

3.15.2. EMSA com competição entre DNA biotilado e DNA não biotilado

Para um volume final de 10 µL a reação de ligação continha 1 µL de 10 x Binding Buffer, 0,5 µL 50% glicerol, 0,5 µL poli dI-dC, água Milli Q estéril, 1 µL de proteína purificada nativa Rv3405c (concentração final 40 nM, exceto na reação 1, sem proteína), e DNA biotilado e não biotilado.

Tabela 3.2: Razão entre DNA biotilado/DNA não biotilado.

	Razão DNA biotilado/DNA não biotilado						
	-	-	0,1:1	1:1	10:1	100:1	1000:1
	1	2	3	4	5	6	7
DNA biotilado							
0,1 uL a x ng/uL	5	5	5	5	5	5	0,5
ng/uL na reação	0,05	0,05	0,05	0,05	0,05	0,05	0,005
DNA não biotilado							
0,1 uL a x ng/uL			0,5	5	-	-	-
1 uL a x ng/uL			-	-	5	50	50
ng/uL na reação			0,005	0,05	0,5	5	5

3.15.3. EMSA na presença de pequenas moléculas

Este ensaio avalia se possíveis moduladores deslocam a proteína Rv3405c da interação com o DNA. As reações de 10 µL continham: 1 µL 10 x *Binding Buffer*/0,5 µL 50% glicerol /0,5 µL poli dI-dC/0,1 µL DNA biotilado 5 ng/µL/água Milli-Q estéril e 1 µL de proteína purificada nativa Rv3405c para concentração final de 40 nM, exceto na reação 1 (sem proteína). A concentração de molécula moduladora foi variada como descrito na Tabela 3.3, modificando o volume adicionado à reação de 10 µL. Cada composto foi diluído em água na concentração listada.

Tabela 3.3: Variação da concentração de compostos nas reações de EMSA.

Composto (concentração do estoque)	Concentrações utilizadas nos ensaios							
	0,19 mM	0,56 mM	1,7 mM	5 mM	10 mM			
taurina (20 mM)	0,19 mM	0,56 mM	1,7 mM	5 mM	10 mM			
bile bovina (0.5%)	0,015%	0,02%	0,025%	0,03%	0,035%	0,04%	0,045%	0,05%
taurodeoxicolato (1%)	0,005%	0,01%	0,02%	0,04%	0,08%	0,16%	0,32%	0,64%
taurocolato (1%)	0,005%	0,01%	0,02%	0,04%	0,08%	0,16%	0,32%	
tauroquenoodeoxicolato (1%)	0,005%	0,01%	0,02%	0,04%	0,08%	0,16%	0,32%	
glicoquenoodeoxicolato (1%)	0,005%	0,01%	0,02%	0,04%	0,08%	0,16%	0,32%	
ácido glicocólico (1%)	0,005%	0,01%	0,02%	0,04%	0,08%	0,16%	0,32%	
glicocolato (1%)	0,005%	0,01%	0,02%	0,04%	0,08%	0,16%	0,32%	

4. RESULTADOS

4.1. Alinhamento entre as regiões reguladoras de *rv3405c* mostra que o motivo palindrômico é conservado entre diversas espécies de micobactérias

A fim de analisar a conservação da região intergênica entre *rv3405c* e *rv3406*, foi feita uma busca por homólogos sintênicos dessa região intergênica incluindo tanto o gene *rv3405c* como o gene *rv3406* através do programa BLAST (Altschul *et al.*, 1990) utilizando o banco de dados NCBI. Somente as sequências de espécies distintas e que continham os genes *rv3405c* e *rv3406* e a região intergênica foram posteriormente analisadas. A partir destas sequências, foi feito um alinhamento global utilizando apenas a região contida entre o códon final do *rv3405c* e o códon inicial do *rv3406*, utilizando a ferramenta “T-Coffee - A novel method for multiple sequence alignments” (Notredame *et al.*, 2000). O resultado desse alinhamento mostra que essa região intergênica é bastante conservada e contém um motivo palindrômico, tal qual um operador típico de reguladores da família TetR. A única espécie não pertencente ao gênero *Mycobacterium* que apresenta homologia a essa região é *Segniliparus rugosus* (Figura 4.1).

```

COL  CATCTGACTACAC-TACC-----AAATGTAGTCAGTTGACTACACC--C-GTTT-----TGA-GCCAAGGAAA-C-GACAATG
INT  CATCTGACTACAC-TACC-----AAACGTAGTCACTTGACTACACC--C-GTTT-----TGA-CTCGAGGAAC-A-CGACATG
AVI  CATCTGACTACAG-TAGC-----ATTGTAGTCAGTTGACTACACC--C-GTTTCGGGTCGCCCT--C-----TTCAAGCCAAGGACG-ACGACAATG
SEG  CACATGGCTACAA-TAGCGGACAA-C-ATTGTAGTCAGTTGACTAAR-----TTG----AAAGGAAC-G-GACCATG
ABS  CATGTGTGTACAG-TAGCGGATGACCAAAGTGTAGTCAGTTGACTACACC--G-GATT-A-----T-----CGGACGGAAGGAAC-G-GGTCATG
MAS  CATGTGTGTACAG-TAGCGGATGACCAAAGTGTAGTCAGTTGACTACACC--G-GATT-T-----T-G-----TCGGACGGAAGGAAC-G-GGTCATG
PAR  CATACGACTACAG-TACT-----CGATGTAGTCAGTTGACTACGCC-----CCTTGA-CACTAGGATTGATGACAATG
MC2  CACTCGACTACAT-TAGC-----GTGTAGTCATTCGACTACAAA--C-G-----T-----GAGGAGC-GCGCAGATG
KAN  CACTCGACTACGC-TACC-----CCATGTAGTCAACCGACTACACC--C-G-----TCTC--CGAAGGACC-ATG-CAATG
ULC  CACACGACTACAC-TAGC-----CGATGTAGTCAACTGACTACACC--C-TACG-----TTTG-CCTAAGGAA--C-AGCAATG
MAR  CACACGACTACAC-TAGC-----CGATGTAGTCAACTGACTACACC--C-TACG-----TTTG-CCTAAGGAA--C-AGCAATG
MOR  CATCCGGCTACAG-TAAC-----CGATGTAGTCATCTGACTACACT--A-ACCA-----TTCA-T-TGAGGACG-CCAGCAATG
XEN  CACCCGGCCTCGTTTACCTGTTGACACCGGTGTAGTCATATGACTACAGTAGCGTATGT-AGTCAATTGACTACATACGCTAC-GCCAAGGAA---T-CCGATG
**  *  *  **  *****  *****  ****  ***

```

Figura 4.1: Alinhamento das regiões intergênicas entre *rv3405c* e *rv3406* mostrando conservação das sequências e a presença de um motivo palindrômico (em cinza), típico de reguladores da família TetR. Os códons de início de tradução de *rv3405c* e *rv3406* estão indicados em vermelho e azul, respectivamente. A sequência sublinhada corresponde à sequência de “Shine-Dalgarno”. Todos os homólogos sintênicos descritos abaixo foram identificados após análise pelo Blast utilizando as regiões intergênicas entre os genes *rv3405c* e *rv3406*. (Entre parênteses está indicado o código de acesso da sequência e as coordenadas da região incluída no alinhamento). *A região de BCG Moreau é idêntica em todos os membros do complexo *M. tuberculosis*.

COL: *M. colombiense* CECT 3035 (AFVW01000010.1 ; 191991-192056); **INT:** *M. intracellulare* ATCC 13950 (NC_016946.1; 4448020-4448085); **AVI:** *M. avium* 104 (NC_008595.1; 4491798-4491857); **SEG:** *Segniliparus rugosus* ATCC BAA-974 (ACZI01000111.1; 25640-25703); **ABS:** *M. abscessus* ATCC 19977 (NC_010397.1; 82011-82088); **MAS:** *M. massiliense* CCUG 48898 MMAS_1 (AHAR01000001.1; 33626-33705); **PAR:** *M. parascrofulaceum* ATCC BAA-614 (ADNV01000278.1; 1019-954); **MAR:** *M. marinum* M (NC_010612.1; 1370781-1370716); **KAN:** *M. kansasii* ATCC 12478 (ACBV01000001.1; 171738-171676); **ULC:** *M. ulcerans* Agy99 (NC_008611.1; 947398-947333); **MC2:** *M. smegmatis* str. MC2 155 (NC_008596.1; 204171-204227) **MOR*:** *M. bovis* BCG str. Moreau RDJ (AM412059.2; 3779925-3779991); **XEN:** *M. xenopi* RIVM700367 (AJFI01000021.1; 109596-109693).

4.2. RT-PCR revela que o gene *rv3406* é transcrito em BCG Moreau, mas não em BCG Pasteur

A fim de corroborar os dados obtidos de proteômica (Pagani *et al.*, manuscrito em preparação) que demonstra a presença da proteína Rv3406 apenas em BCG Moreau e não em Pasteur, utilizamos a ferramenta de RT-PCR para avaliar a expressão do gene *rv3406*.

Visualizamos por RT-PCR, realizado com RNA extraído de triplicatas biológicas de culturas de 15 dias em meio Sauton, que *rv3406* é transcrito em BCG Moreau, mas não em BCG Pasteur (**Figura 4.2**). Os controles positivos e negativos confirmam as condições experimentais adequadas. Portanto, confirmamos a presença do transcrito do gene *rv3406* somente em BCG Moreau, em conformidade com os dados anteriores do laboratório (Pagani *et al.*, manuscrito em preparação).



Figura 4.2: RT-PCR revela que o gene *rv3406* é transcrito em BCG Moreau, mas não em BCG Pasteur. +RT, PCR para *rv3406* sobre preparação de cDNA utilizando o primer reverso de *rv3406*. Controles negativo (-), PCR diretamente sobre a preparação de RNA tratada com *DNase* e, positivo (+), PCR utilizando DNA genômico. M, marcador de peso molecular de 100 pb.

4.3. Clonagem, expressão e purificação das proteínas recombinantes Rv3405c de *M. bovis* BCG Pasteur e Rv3406 de *M. bovis* BCG Moreau em *E. coli*

4.3.1 Amplificação dos genes *rv3405c* de *M. bovis* BCG Pasteur e *rv3406* de *M. bovis* BCG Moreau

De modo a obter a região codificante para as proteínas de interesse visando sua expressão heteróloga, desenhamos os oligonucleotídeos listados no Anexo II. Esses oligonucleotídeos possuem sítios de restrição específicos para *NcoI/XhoI* formando pontas coesivas para ligação com o vetor digerido com as mesmas enzimas. Eles foram utilizados em reações de PCR contendo o DNA genômico de *M. bovis* BCG Pasteur ou *M. bovis* BCG Moreau (como descrito na metodologia), resultando na amplificação de fragmentos de 580 pb (*rv3405c*) e 901 pb (*rv3406*), respectivamente.

4.3.2 Subclonagem no vetor pBAD TOPO TA e clonagem no vetor pET28a

Para facilitar a clonagem de *rv3405c* e *rv3406* no vetor de expressão pET28a utilizamos a estratégia de primeiro subclonar os produtos de PCR no vetor pBAD TOPO TA, segundo instruções do fabricante. Feita a subclonagem, o DNA plasmidial dos clones em pBAD foi extraído e utilizado para digestão com as enzimas *NcoI/XhoI*. Após a purificação dos fragmentos e do DNA do vetor pET28a digerido com as mesmas enzimas (**Figura 4.3**), deu-se prosseguimento à ligação utilizando T4 DNA ligase. As ligações foram eletroporadas em *E.coli* Top10 e clones positivos identificados por PCR de colônia. Fizemos a miniprep para obtenção dos plasmídeos e a confirmação da presença dos insertos foi feita por digestão com as enzimas de clonagem. Verificamos por sequenciamento de DNA (Plataforma de sequenciamento PDTIS/Fiocruz) que as sequências obtidas dos fragmentos estão livres de mutações. Com isso, prosseguimos com a eletroporação na cepa *E. coli* BL21 (DE3), utilizada para expressão das proteínas recombinantes Rv3405c e Rv3406 (construções agora denominadas pET-P3405c e pET-M3406, respectivamente). É importante ressaltar que a clonagem dos genes *rv3405c* e *rv3406* no vetor pET28a resultou em construções que geram proteínas com fusão C terminal com cauda de 6-histidinas, sendo que a massa molecular esperada para cada uma é de aproximadamente 20 e 32 kDa, respectivamente.

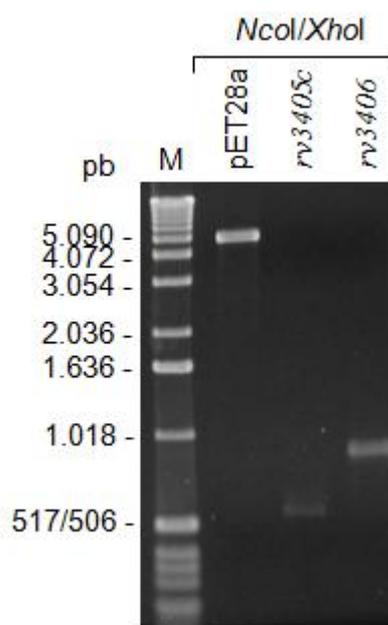


Figura 4.3: Clonagem em vetor pET28a dos genes *rv3405c* (pET-P3405c) e *rv3406* (pET-M3406) para expressão em BL21(DE3). Vetor pET28a digerido - *NcoI/XhoI* (5.369 pb) e os genes *rv3405c* de Pasteur (580 pb) e *rv3406* de Moreau (901 pb) digeridos - *NcoI/XhoI* da subclonagem no vetor pBAD TOPO. M, marcador de tamanho molecular de 1 Kb.

4.3.3 Super-expressão e teste de solubilidade das proteínas recombinantes rRv3405c e rRv3406 em *E. coli* BL21(DE3)

A fim de melhorar a expressão das proteínas recombinantes, várias condições de temperatura, tempo de indução e concentração do indutor foram testadas para definir a condição ideal de indução da expressão de Rv3405c e Rv3406. A melhor condição encontrada foi indução com 1 mM IPTG adicionado a cultura na OD₆₀₀ de 0,5 e incubação durante 3 horas a 37°C (dados não mostrados). Como demonstra a **Figura 4.4 (A e B)**, os clones apresentam expressão basal da proteína recombinante antes da indução, aumentando consideravelmente após a adição de IPTG (I).

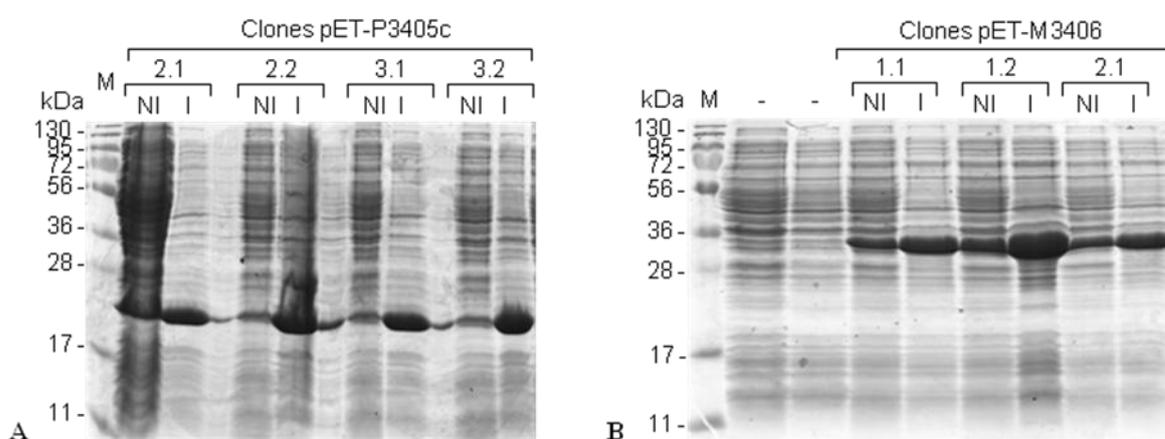


Figura 4.4: Super-expressão das proteínas recombinantes Rv3405c (A) e Rv3406 (B). As culturas de *E. coli* BL21(DE3) transformadas com o vetor pET-P3405c e vetor pET-M3406 foram crescidas até OD₆₀₀ 0.5 sendo induzidas com 1 mM IPTG por 3 h. O equivalente a 0,1 DO_{600nm} da cultura original foi aplicado em cada canal de géis SDS-PAGE 15% para visualização das proteínas presentes. NI, não induzido. I, induzido. M, marcador de tamanho molecular.

Selecionamos alguns clones para ensaio de WB para confirmação do especificado peso molecular da proteína recombinante e presença da cauda de histidinas na porção C-terminal, utilizando para isso anticorpo primário anti-His. Como mostra a **Figura 4.5**, confirmamos o tamanho esperado das bandas como descrito acima.

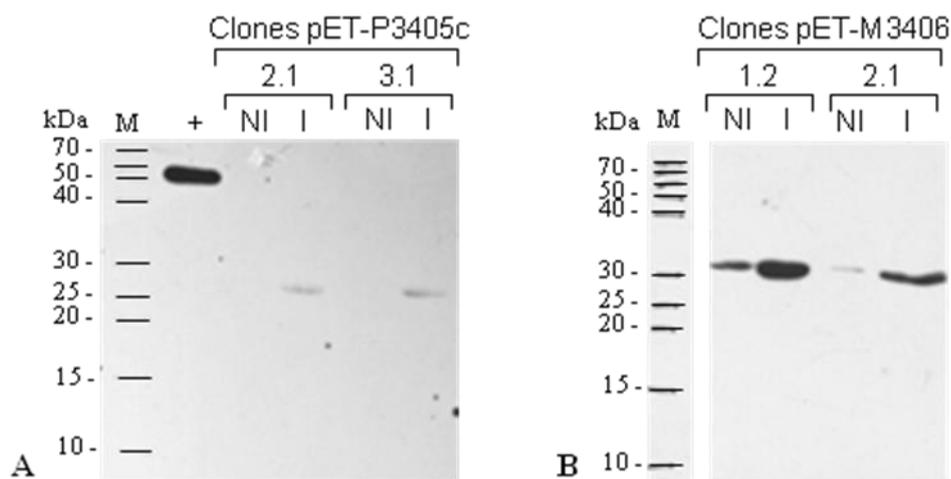


Figura 4.5: WB para confirmação do peso molecular das proteínas Rv3405c (A) e Rv3406 (B). NI, não induzido. I, induzido. M, marcador de tamanho molecular. +, controle positivo (proteína Erp de *Mycobacterium leprae*). WBs foram revelados com anti-his.

Avaliamos a solubilidade das proteínas recombinantes após indução com 1 mM IPTG e incubação a 37°C, com agitação por 3 ou 16 horas. Após lise bacteriana separamos o sedimento (contendo corpos de inclusão) do sobrenadante (proteínas solúveis), e comprovamos na corrida em SDS-PAGE que o tempo de 3 h leva a uma maior produção das proteínas recombinantes na forma solúvel (**Figura 4.6 A e B**).

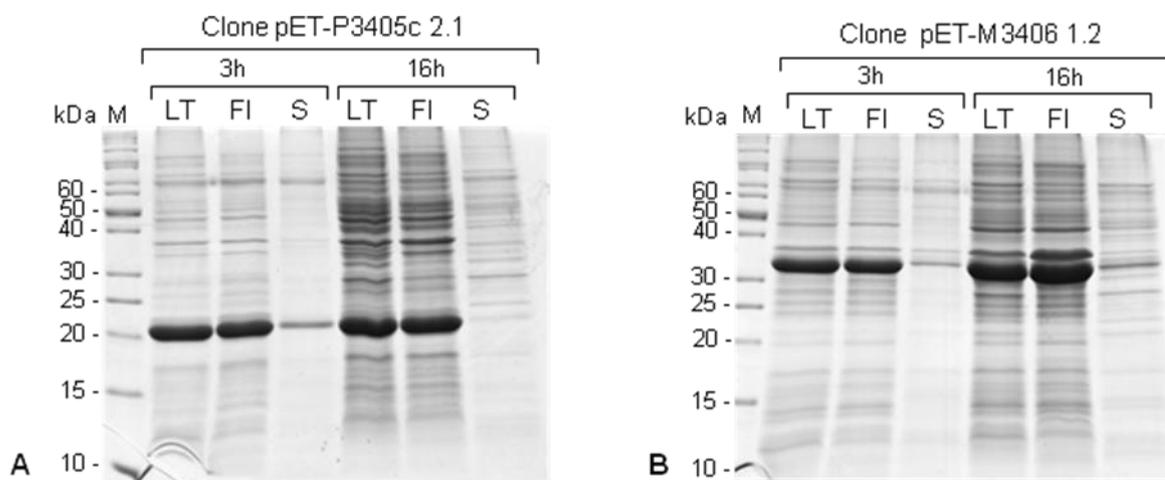


Figura 4.6: Solubilidade das proteínas recombinantes. Teste de solubilidade das construções expressando as proteínas recombinantes Rv3405c (A) e Rv3406 (B). As culturas foram induzidas com 1mM IPTG durante 3 h ou 16 h e o equivalente a 0,1 DO_{600nm} da cultura original foi aplicado em cada canal de géis SDS-PAGE 15% para visualização das proteínas presentes no lisado total (LT), fração insolúvel (FI) e solúvel (S) após lise e clarificação. M, marcador de peso molecular indicado à esquerda dos géis, em kDa.

Na tentativa de melhorar ainda mais a expressão da proteína Rv3406 na forma solúvel, para realização de ensaios enzimáticos com a proteína nativa, utilizamos o meio “overnight express” conforme descrito na metodologia. Verificamos que houve uma pequena melhora na expressão da proteína assim como um aumento na quantidade de proteína solúvel (**Figura 4.7**). Testamos se existe alguma diferença entre o meio “overnight express” vendido comercialmente ou fabricado no próprio laboratório e não observamos diferença na expressão da proteína Rv3406 (dados não mostrados).

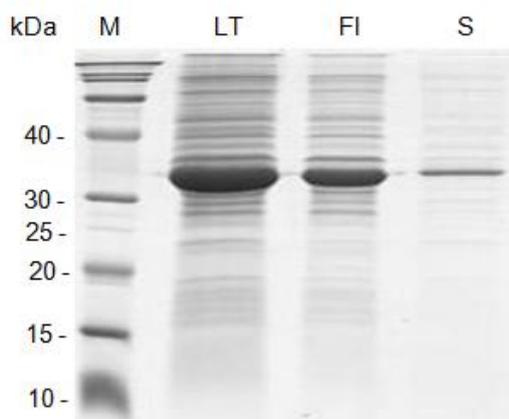


Figura 4.7: Teste de solubilidade de Rv3406 (pET-M3406) em meio “overnight express” por 16 horas (~25°C). O equivalente a 0,1 OD_{600nm} da cultura original foi aplicado em cada canal do gel SDS-PAGE 15% para visualização das proteínas presentes no lisado total (LT), fração insolúvel (FI) e solúvel (S) após lise e clarificação. M, marcador de peso molecular indicado à esquerda, em kDa.

4.3.4 Purificação das proteínas Rv3405 e Rv3406 recombinantes a partir de corpos de inclusão

Seguindo o protocolo de lise e clarificação descrito na metodologia, a purificação das proteínas a partir dos corpos de inclusão foi feita em coluna HisTrap de afinidade a níquel após solubilização das proteínas em solução desnaturante contendo 8 M ureia. As frações coletadas foram analisadas em SDS-PAGE como mostrado na **Figura 4.8 (A e B)**.

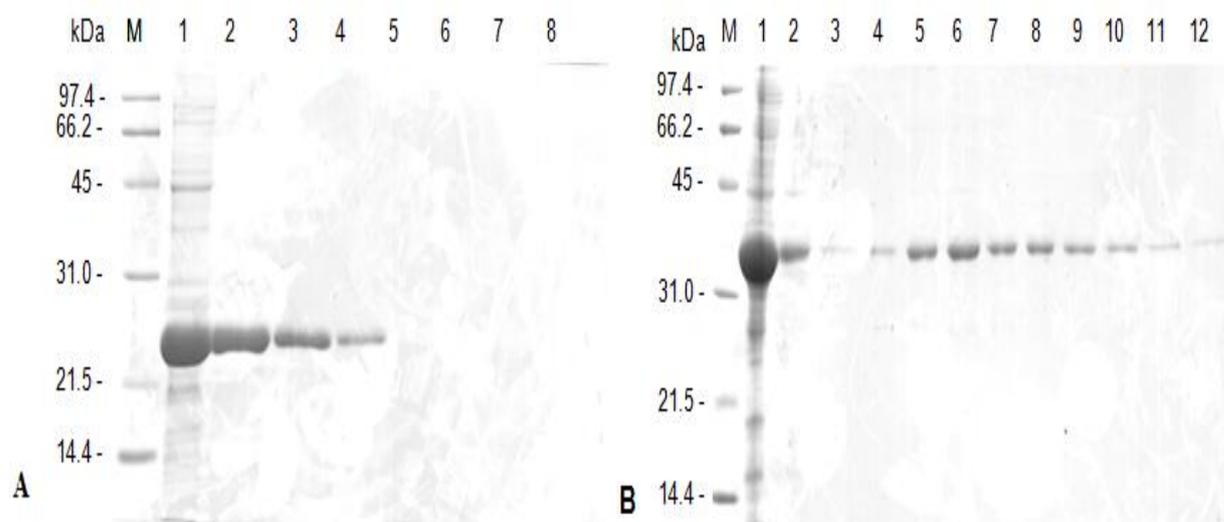


Figura 4.8: Purificação das proteínas Rv3405c (A) e Rv3406 (B) por FPLC a partir dos corpos de inclusão. M, marcador de tamanho molecular; 1, fração de corpos de inclusão ressuspensos em 8 M uréia; 2 a 12, frações coletadas durante eluição da coluna de Ni⁺² com gradiente de imidazol.

As frações mais puras foram reunidas em *pool* e dialisadas contra PBS 1x para retirar a uréia (**Figura 4.9**). Utilizamos essas proteínas para imunização de camundongos Balb/c para obtenção de soros policlonais.

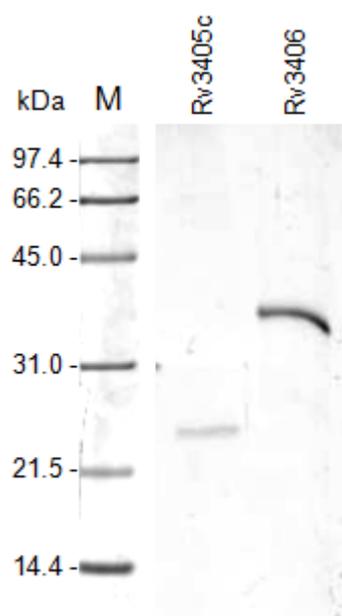


Figura 4.9: Proteínas Rv3405c e Rv3406 purificadas a partir de corpos de inclusão. M, marcador de tamanho molecular. Gel SDS-PAGE 15% aplicados 10 µl de cada proteína após purificação e diálise em PBS 1x.

4.3.5 Purificação das proteínas Rv3405c e Rv3406 recombinantes a partir da fração solúvel

A fim de preservar a estrutura nativa fizemos a purificação das proteínas Rv3405c e Rv3406 recombinantes a partir da fração solúvel obtida de culturas crescidas em 500 mL de meio LB das cepas de *E. coli* BL21(DE3) contendo as construções pET-3405c e pET-3406 sob as condições ideais de indução (indução com 1mM IPTG durante 3h a 37°C com agitação). Após a lise, o material foi centrifugado e as proteínas contidas na fração solúvel precipitadas com adição de 85% de sulfato de amônio, conforme descrito na metodologia. Após ressuspensão, as proteínas recombinantes foram purificadas em coluna Sepharose-Ni²⁺, e as frações eluídas em gradiente de imidazol analisadas por SDS-PAGE como vemos na **Figura 4.10 (A e B)**. Todas as frações contendo cada proteína pura foram reunidas em *pool* e dialisadas contra tampão 10 mM Tris-Cl pH 7,5/50 mM NaCl contendo glicerol 2,5% e Tween 20 0,05%. A dosagem foi feita pelo método DC da BioRad utilizando como curva padrão BSA com fator de calibração 0,1 mg/mL. O resultado da dosagem das proteínas solúveis Rv3405c e Rv3406 mostrou que estão a 0,12 mg/mL e 0,13 mg/mL, respectivamente.

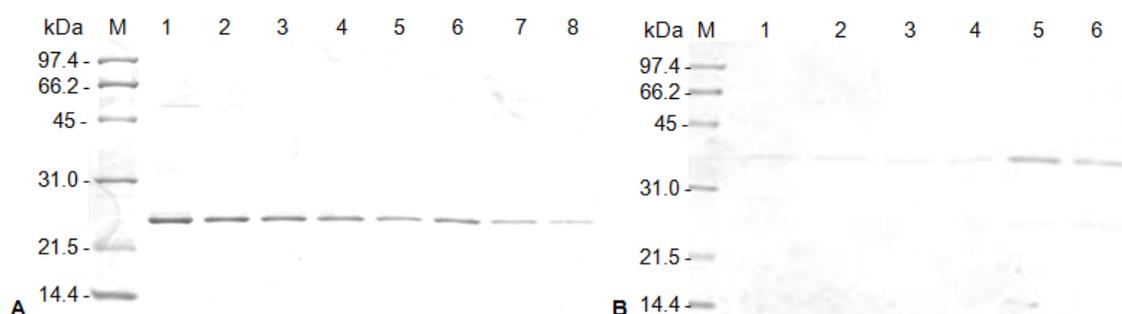


Figura 4.10: Purificação das proteínas recombinantes solúveis Rv3405c (A) e Rv3406 (B). As proteínas presentes na fração solúvel após lise e clarificação foram precipitadas com 85% de sulfato de amônio e, após ressuspensão, purificadas em coluna de Ni²⁺ - Sepharose. As frações eluídas em gradiente imidazol foram visualizadas em géis SDS-PAGE 15%. M, Marcador de peso molecular indicado à esquerda em kDa.

4.3.6 Titulação dos anticorpos

Para definir a diluição adequada para uso dos soros policlonais anti-Rv3405c e anti-Rv3406 em WB, os anticorpos foram titulados variando a diluição empregada contra a proteína recombinante utilizada na imunização. Cada proteína foi submetida a separação em gel SDS-PAGE preparativo e transferida para membrana de nitrocelulose. Após bloqueio, a membrana foi cortada em tiras e cada tira incubada com diluições do soro policlonal. O soro anti-Rv3405c foi testado nas diluições 1:500 e 1:1000, e ambas apresentaram título satisfatório (**Figura 4.11**).

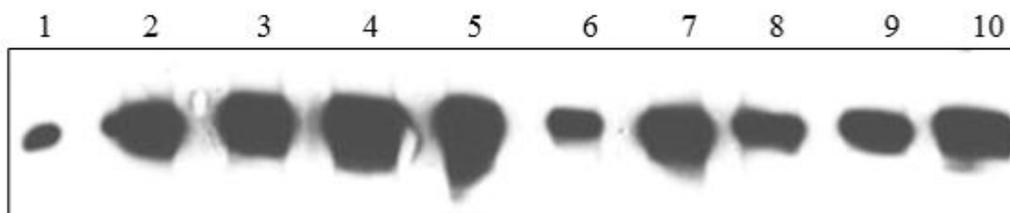


Figura 4.11: WB para titulação do anticorpo policlonal anti-Rv3405c produzido em camundongos Balb-c. Detecção de Rv3405c recombinante pelo soro obtido de 5 camundongos diferentes nas diluições 1:500 (canais 1-5) e 1:1000 (canais 6-10).

Já o soro anti-Rv3406 apresentou um título de até 1:16000 (**Figura 4.12**).

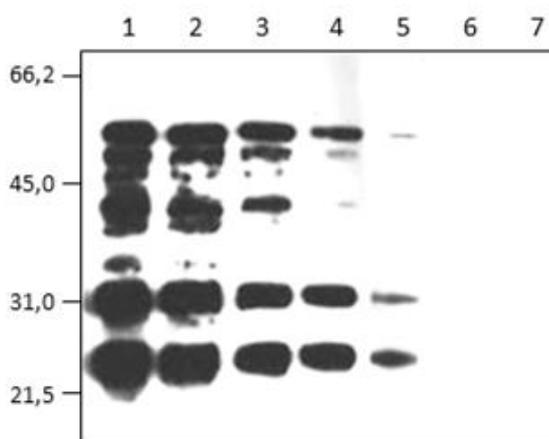


Figura 4.12: WB para titulação do anticorpo policlonal anti-Rv3406 produzido em camundongos Balb-c. Detecção de Rv3406 recombinante pelo soro policlonal diluído (1) 1:1.000; (2) 1:2.000; (3) 1:4.000; (4) 1:8.000; (5) 1:16.000; (6) 1:32.000; (7) 1:64.000. Marcador de peso molecular indicado à esquerda, em kDa.

4.4. A proteína Rv3406 é detectada por “western blot” em BCG Moreau, mas não em BCG Pasteur

Na tentativa de reforçar os dados de proteômica que mostra a proteína Rv3406 presente somente em BCG Moreau e não em Pasteur, este experimento de WB especifica além da presença da proteína também a sua localização nas frações de BCG Moreau e não em Pasteur e, portanto, está incluído no artigo Pagani *et al.* (manuscrito em preparação). A proteína Rv3406 está presente na fração intracelular e na de superfície de BCG Moreau (**Figura 4.13; canais 1 e 3**), porém ausente nas de BCG Pasteur (**canais 2 e 4**). Já no filtrado de cultura Rv3406 está ausente tanto em BCG Moreau como em BCG Pasteur (**canais 5 e 6**). O lisado de *E. coli* presente na incubação com o soro anti-Rv3406 (1:1.000) foi utilizado para reduzir as bandas inespecíficas (**canais 1 a 4**), demonstrando que a proteína no momento da injeção nos camundongos não estava totalmente pura, ou seja, existiam proteínas de *E. coli*

misturadas e que no momento de formação dos anticorpos policlonais estes também passam a reconhecer alguns epítomos inespecíficos, provenientes dessas proteínas de *E. coli* contaminantes. A massa molecular predita das proteínas Rv3406 endógena e recombinante são, respectivamente, 32579 e 33833 Da.

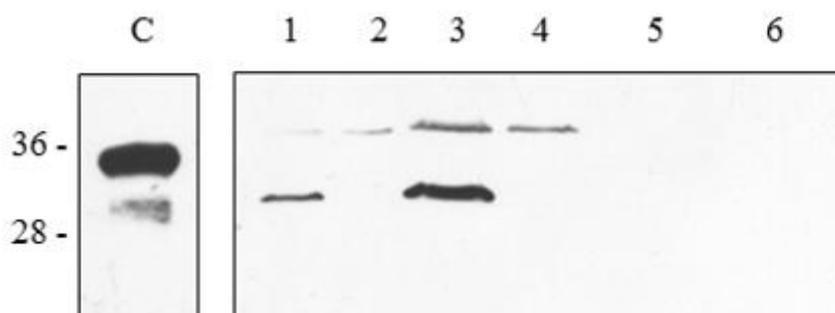


Figura 4.13: A proteína Rv3406 é detectada em BCG Moreau (canais ímpares) mas não em BCG Pasteur (canais pares). WB utilizando como anticorpo primário anti-Rv3406 (1:1000) para detectar a presença da proteína Rv3406 nas frações de proteínas intracelulares (canais 1 e 2), de superfície (canais 3 e 4) e de filtrado de cultura (canais 5 e 6) de BCG Moreau (canais 1, 3 e 5) e Pasteur (canais 2, 4 e 6). C, controle da proteína Rv3406 recombinante. Marcadores de peso molecular indicados à esquerda, em kDa.

4.5. RT-PCR para o gene *rv3405c* comprova que existe transcrição desse gene em BCG Pasteur e em BCG Moreau (gene quimérico *rv3405c-00*)

A fim de verificar se existe a presença do transcrito do gene *rv3405c* em BCG Pasteur, assim como do gene quimérico (*rv3405c-00*) em BCG Moreau realizamos o experimento de RT-PCR para detecção de mRNA extraído de culturas crescidas em meio Sauton. Como podemos visualizar na **Figura 4.14** existe transcrito tanto do gene *rv3405c* em BCG Pasteur (**canal 6**) como do gene quimérico *rv3405c-00* em BCG Moreau (**canal 3**).

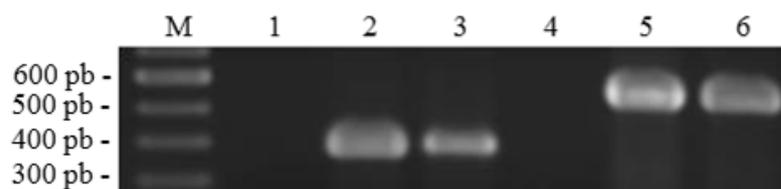


Figura 4.14: RT-PCR revela que o gene *rv3405c* é transcrito em BCG Moreau (gene quimérico *rv3405-00c*) e em BCG Pasteur (*rv3405c*).; 1-3, BCG Moreau; 4-6 BCG Pasteur. 1 e 4, -RT* de *rv3405c*. 2 e 5, C+ M (392 pb) e P (567 pb). 3 e 6, +RT de *rv3405c*. M, Marcador de tamanho molecular de 100 pb. *-RT: PCR para *rv3405c* sobre preparação de RNA submetida a tratamento com *DNase I*, sem adição da transcriptase reversa. C+: PCR utilizando DNA genômico; +RT, PCR para *rv3405c* sobre preparação de cDNA utilizando o primer reverso de *rv3405c*.

4.6. O gene quimérico *rv3400-05c* é transcrito em ambas as direções em BCG Moreau: do promotor de *rv3400* (*rv3400-05c*) e do promotor de *rv3405c* (*rv3405c-rv3400*)

Em BCG Moreau a deleção RD16 gera uma nova organização cromossômica na qual a sequência de *rv3405c* pode estar sendo transcrita tanto a partir de um promotor adjacente (a montante - *upstream*) a *rv3400* quanto pelo promotor do gene *rv3405c* (**Figura 4.15 A**). Para avaliar se há transcrição a partir destes promotores, usamos RT-PCR do gene truncado *rv3405c* (BCG_M3438) em BCG Moreau (**Figura 4.15 B**). Os resultados mostram que *rv3405c* é transcrito tanto no sentido 05c-00 (promotor de *rv3405c*; **canal 3**) como no sentido 00-05 (promotor de *rv3400*; **canal 4**).

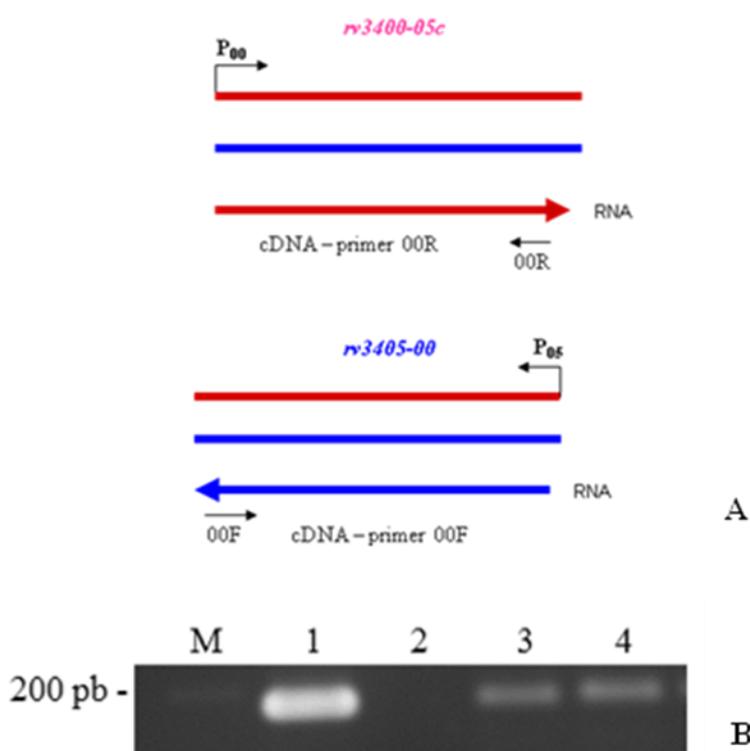


Figura 4.15: RT-PCR revela que o gene quimérico *rv3400-05c* é transcrito em BCG Moreau em ambas as direções, tanto a partir do promotor de *rv3400* quanto do promotor de *rv3405c*. **(A)** Esquema representativo do desenho experimental. **(B)** 1, C+ (198 pb); 2, -RT de *rv3400-05c*; 3, +RT de *rv3405-00c* (cDNA Forward); 4, +RT de *rv3400-05c* (cDNA Reverso). -RT, PCR para 00-05 sobre preparação de RNA submetida a tratamento com *DNase I*. +RT, PCR para 00-05 sobre preparação de cDNA utilizando ou o primer reverso de 00-05 verificando a direção do promotor de 00 ou o primer forward de 00-05 verificando a direção do promotor de 05. M, marcador de tamanho molecular.

4.7. Os genes *rv3400-05c* (BCG_M3438) e *rv3406* não são co-transcritos em BCG Moreau

Uma vez que conhecidamente em bactérias muitos genes são co-transcritos em mRNAs policistrônicos, formando operons, este experimento de RT-PCR verificou se a deleção RD16 poderia gerar um acoplamento transcricional nesta região entre os genes *rv3400-05c* (BCG_M3438) e *rv3406*. Contudo, conforme demonstrado na **Figura 4.16** os genes *rv3400-05c* e *rv3406* não são co-transcritos em BCG Moreau (**canal 5**). Este resultado sugere a hipótese que o motivo palindrômico presente na região intergênica 05-06 (**Figura 4.1**) poderia funcionar como um grampo de parada para a RNA polimerase.

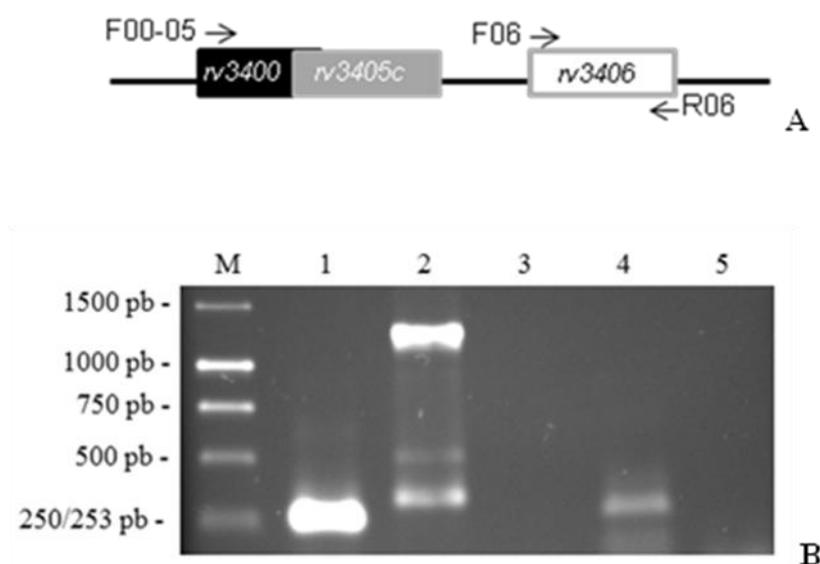


Figura 4.16: RT-PCR revela que *rv3400* e *rv3406* não são co-transcritos em BCG Moreau. (A) Esquema representativo do desenho experimental. (B) 1, C+ PCR de *rv3406* sobre DNA (218 pb); 2, C+ PCR de *rv3400-06* sobre DNA (1128 pb); 3, -RT de *rv3400-06*; 4, +RT de *rv3406*; 5, +RT de *rv3400-06*. -RT, PCR para *rv3400-06* sobre preparação de RNA submetida a tratamento com *DNase I*. +RT, PCR para *rv3406* ou *rv3400-06* sobre preparação de cDNA utilizando o primer reverso de *rv3406*. M, marcador de tamanho molecular de 1 kb.

4.8. Ensaio de retardo em gel (EMSA) mostra a interação entre Rv3405c e a região intergênica *rv3405c-06*

Para se comprovar a interação entre a proteína Rv3405c e a região intergênica *rv3405c-06*, que apresenta um motivo palindrômico (**Figura 4.1**), fizemos um ensaio de retardo em gel (EMSA). Neste ensaio a revelação é feita através do fragmento de PCR do DNA biotilado (101 pb). Um retardo na corrida representa uma interação entre a proteína e a porção de DNA analisada.

Como podemos visualizar na **Figura 4.17**, o ensaio de retardo em gel comprova que existe um retardo no gel referente a banda de DNA à medida que aumentamos a quantidade da proteína Rv3405c. A interação de Rv3405c começa a ocorrer quando a proteína está na concentração de 400 pM. Com isso, comprovamos que a região intergênica *rv3405c-06* interage com o fator de transcrição Rv3405c, inclusive podendo formar possivelmente dímeros ou tetrâmeros devido a presença de mais de uma banda no retardo (**canais 5 a 8**).

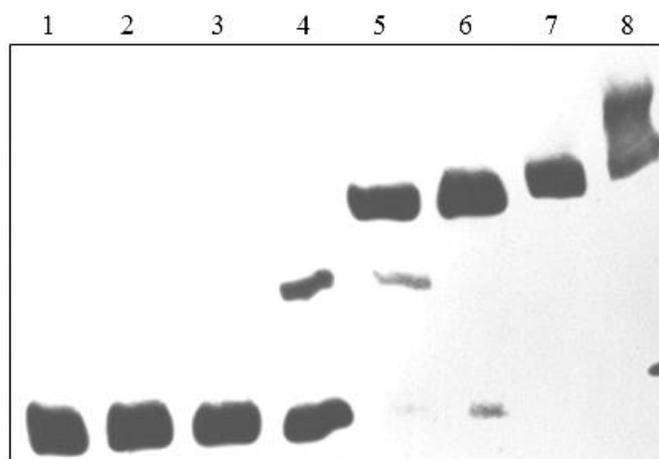


Figura 4.17: A proteína Rv3405c se liga à região intergênica *rv3405-06* (fragmento de PCR de 101pb) formando um complexo com afinidade nanomolar em ensaio de retardo em gel. 1, DNA livre biotilado. 2 a 8, DNA incubado com quantidades crescentes de rRv3405c (4 pM, 40 pM, 400 pM, 4 nM, 40 nM, 400 nM e 4 uM).

4.9. Ensaio de retardo em gel (EMSA) revela que a interação entre Rv3405c e a região intergênica *rv3405-06* pode ser modulada por compostos presentes em bile bovina

Após verificar por EMSA que a proteína recombinante Rv3405c de BCG Pasteur interage com a região intergênica *rv3405-06*, foi verificado se essa interação pode sofrer alguma alteração na presença de determinados compostos que poderiam modular sua atividade de união a DNA. A primeira molécula testada foi taurina **Figura 4.18 (C)**, substrato de uma possível taurina dioxigenase codificada por Rv3406. Conforme resultado da **Figura 4.18 (A)**, em ensaios de retardo em gel a taurina não interfere na formação do complexo Rv3405-região intergênica *rv3405-06*. A adição de concentrações crescentes de taurina (entre 0,19 mM e 10 mM) não alterou a formação do complexo. Averiguamos então se a bile bovina, utilizada por Calmette e Guérin no meio de cultivo que levou à obtenção do BCG original, poderia alterar de alguma maneira a formação do complexo Rv3405 com a região intergênica *rv3405-06*. Como mostrado na **Figura 4.18 (B)**, a adição de concentrações crescentes de bile

inibe a formação do complexo proteína-DNA. Este efeito começa a ser observado com bile na concentração de 0,015% e a inibição total ocorre a 0,035%.

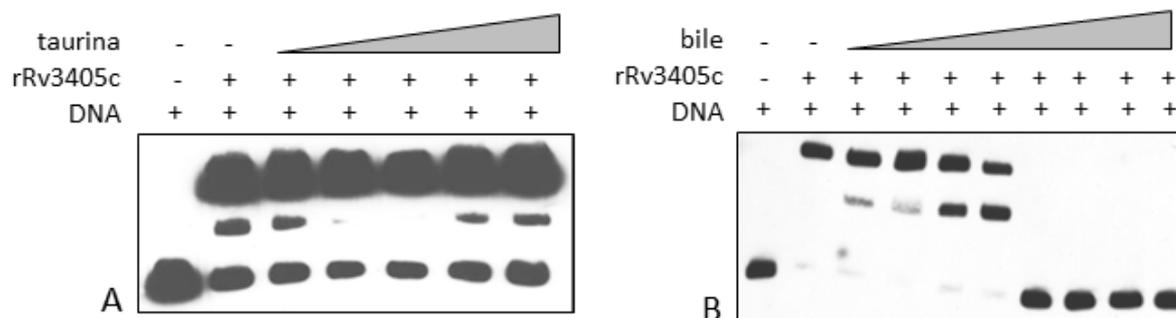


Figura 4.18: Bile bovina mas não taurina é capaz de desfazer o complexo rRv3405c-DNA. Ensaio de retardo em gel (EMSA) para visualização da interação entre o fragmento de DNA relativo à região intergênica 05c-06, biotilado (0,5 ng) e a proteína rRv3405c purificada (40 nM), na presença de concentrações crescentes de taurina (**painel A**; 0,19 mM, 0,56 mM, 1,7 mM, 5 mM, 10 mM) ou bile bovina (**painel B**; 0,015, 0,02, 0,025, 0,03, 0,035, 0,04, 0,045, 0,05%). A presença (+) ou ausência (-) de cada componente está indicada na parte superior da figura.

Uma vez que a bile bovina é composta por várias moléculas, é pertinente determinar qual(is) destes sais biliares poderiam atuar de forma específica, deslocando a proteína do DNA. A **Figura 4.19** mostra os compostos biliares testados, suas estruturas químicas e o resultado dos experimentos de EMSA, com concentrações crescentes dos inibidores. Estes resultados estão resumidos na Tabela 4.1, e indicam que tanto o tauro- quanto o glico-conjugado do ácido quenodeoxicolato se mostraram mais eficientes em inibir a ligação da proteína rRv3405c à região intergênica.

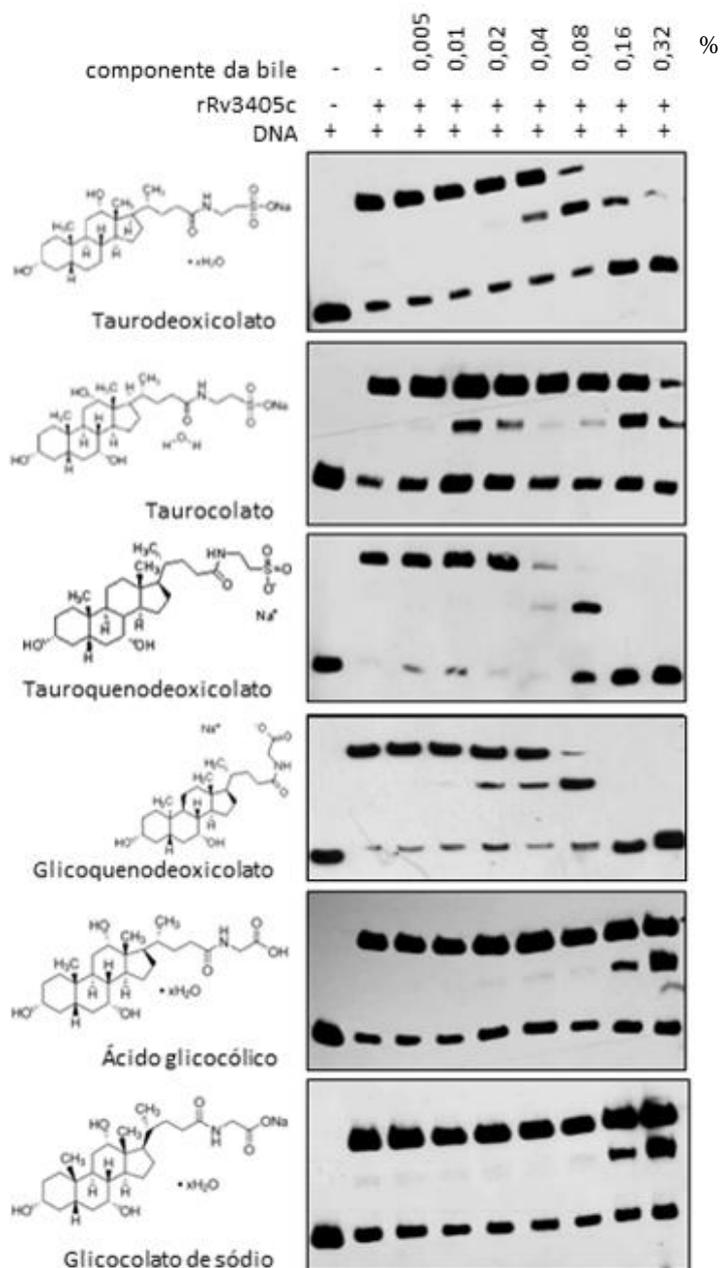


Figura 4.19: Diferentes componentes da bile são capazes de deslocar a proteína rRv3405c do DNA, com eficiência variável. Experimentos de EMSA utilizando DNA biotilado (região intergênica 05c-06; 0,5ng) e a proteína rRv3405c (40nM) para verificar a capacidade de componentes da bile bovina (indicados, com suas respectivas estruturas químicas, à esquerda) para desfazer a ligação de rRv3405c ao DNA. As diferentes concentrações utilizadas estão indicadas.

Tabela 4.1. Concentração inibitória dos diferentes componentes da bile testados no ensaio de ligação de rRv3405c ao DNA da região intergênica 05c-06.

Componente da bile	Concentração
Taurodeoxicolato de sódio	0,32% - inibição parcial
Taurocolato de sódio	0,32% - inibição parcial
Tauroquenodeoxicolato de sódio	0,16% - inibição total
Glicoquenodeoxicolato de sódio	0,16% - inibição total
Ácido glicocólico	0,32% - inibição parcial
Glicocolato de sódio	0,32% - inibição parcial

4.10. RT-PCR para o gene *rv3406* nos cultivos de BCG Pasteur e Moreau mostra que a presença de ligantes de Rv3405c induzem a transcrição de *rv3406*

A fim de testar se as interações entre Rv3405c e bile ou seus componentes observadas em ensaios de retardo em gel ocorrem *in vivo*, foram executados experimentos de RT-PCR com RNA extraído de cultivos de BCG Pasteur e BCG Moreau sob condições específicas. Conforme descrito na metodologia, esses cultivos foram realizados em meio 7H9 e induzidos com concentrações específicas por um período de 19 horas para cada possível modulador selecionado. O RT-PCR para o gene *rv3406* foi utilizado para comprovar se de fato esses possíveis moduladores são capazes de induzir a transcrição, uma vez que há evidência *in vitro* nos ensaios de retardo em gel de que eles inibem a formação do complexo proteína Rv3405c-região intergênica 05-06. Como mostra a **Figura 4.20**, o gene *rv3406* é constitutivamente expresso em BCG Moreau, de forma ligeiramente variável de acordo com a presença de moléculas moduladoras. A adição de taurina em cultura de BCG Pasteur parece ter induzido levemente a expressão de *rv3406*. Já a presença de bile e dos conjugados de ácido quenodeoxicólico, tauro- e glico-, induz de forma clara a expressão de *rv3406*.

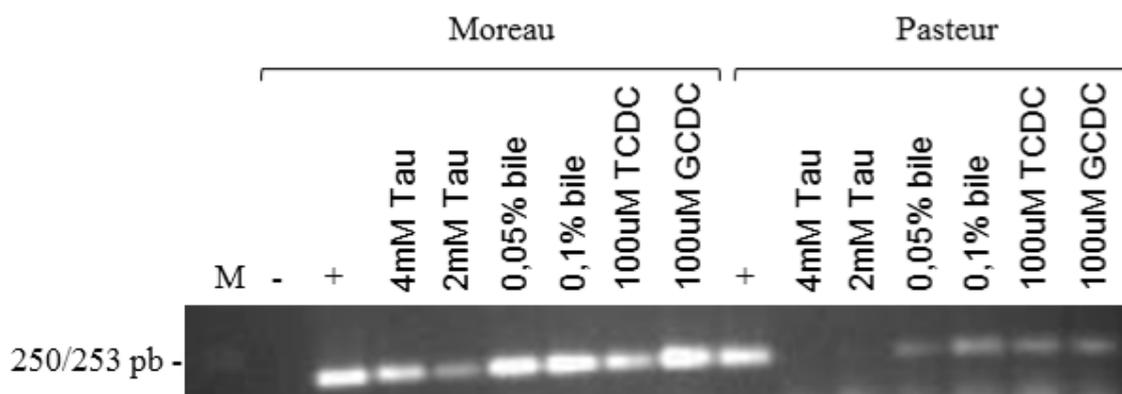


Figura 4.20: Componentes da bile induzem a expressão de *rv3406* em culturas de BCG Pasteur. RT-PCR para detecção de mRNA de *rv3406* em culturas de BCG Moreau e Pasteur cultivadas na presença de 4 mM (0,05%) e 2 mM taurina (0,025%), 0,05 e 0,1% bile e 100 μ M tauroquenodeoxicolato (0,005% TCDC) e glicoquenodeoxicolato (0,005% GCDC), conforme indicado acima da figura. Controles negativo (-), PCR diretamente sobre a preparação de RNA tratada com DNase e, positivo (+), PCR utilizando DNA genômico foram incluídos. M, marcador de tamanho molecular, indicado à esquerda.

4.11. “Western blot” para detecção de Rv3406 nos cultivos de BCG Pasteur e Moreau incubados com os possíveis moduladores confirma a acumulação desta proteína apenas nas culturas de BCG Moreau

O WB com anti-Rv3406 nos cultivos de BCG Pasteur e Moreau incubados com os possíveis moduladores foi utilizado para detectar a proteína Rv3406 nos cultivos onde o RT-PCR revelou presença do transcrito. Contudo, como podemos observar na **Figura 4.21** somente podemos constatar a presença da proteína Rv3406 nos cultivos de BCG Moreau, não tendo sido detectada em nenhum cultivo de BCG Pasteur, inclusive naqueles em que o transcrito de *rv3406* pode ser detectado por RT-PCR.

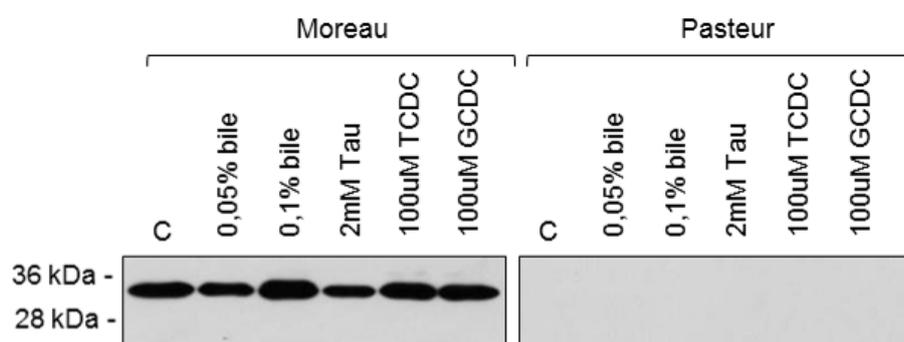


Figura 4.21: WB utilizando o anticorpo primário anti-Rv3406 não detecta Rv3406 em culturas de BCG Pasteur cultivadas na presença dos componentes da bile. 20 μ g de proteínas da fração intracelular, obtidas de culturas de BCG Moreau e Pasteur em meio 7H9/ADC/Tween 80 (C) com adição de 0,05 e 0,1% bile, 2 mM Taurina (0,025 % Tau), e 100 μ M tauroquenodeoxicolato (0,005 % TCDC) e glicoquenodeoxicolato (0,005% GCDC), conforme indicado acima da figura, foram aplicadas por canal em gel SDS-PAGE 12%. Após transferência para membrana de nitrocelulose, incubação com anticorpo policlonal anti-Rv3406 (diluição 1:1000) e revelação, observamos a presença de uma banda positiva apenas nas frações obtidas de BCG Moreau. Marcador de peso molecular, indicado à esquerda.

5. DISCUSSÃO

Este projeto analisou o papel funcional do gene *rv3405c*, especificamente em relação ao impacto da deleção da região RD16 em BCG Moreau, responsável pelo truncamento desse gene e, conseqüentemente, sua possível perda funcional. Resultados prévios de proteômica obtidos em nosso laboratório (Pagani *et al.*, manuscrito em preparação) demonstram a presença da proteína Rv3406 na superfície celular de BCG Moreau mas não de BCG Pasteur. O mecanismo típico de ação de reguladores tipo TetR sugeriria que a proteína Rv3405c de BCG Moreau estaria inativa ao não poder formar dímeros, necessários à sua atividade repressora, e reconhecer putativos moduladores transcricionais. Assim, a correlação entre a diferença qualitativa da expressão de Rv3406 com a presença do gene *rv3405c* em BCG Pasteur e sua ausência em BCG Moreau sugere fortemente que a proteína Rv3405c atua como repressor transcricional de *rv3406*. Esta hipótese é fortalecida pelos nossos resultados de RT-PCR e WB mostrando que o gene *rv3406* é transcrito e que a proteína Rv3406 é acumulada em BCG Moreau e não em BCG Pasteur, e que Rv3405c é uma proteína capaz de se ligar ao DNA. A confirmação do papel de Rv3405c na repressão de *rv3406* foi obtida no laboratório durante a escrita desta tese, por experimento de complementação mostrando que BCG Moreau expressando *rv3405c* intacto (de BCG Pasteur) de forma heteróloga não acumula Rv3406.

Reguladores tipo TetR podem autorregular sua expressão negativa- (Bertram e Hillen, 2008) ou positivamente (Folcher *et al.*, 2001), de forma que o perfil transcricional de *rv3405c* poderia ser similar ao de *rv3406* em BCG Pasteur e Moreau. A detecção da transcrição de *rv3405c* em BCG Moreau e BCG Pasteur mostra que sua inativação não elimina a atividade promotora, sugerindo que este regulador não atua como autoativador como descrito para algumas proteínas desta família, como Pip (Folcher *et al.*, 2001) e MdoR (Park *et al.*, 2011).

Nossos dados apoiam a hipótese de “ganho funcional” (ou *knock-in* funcional) pela perda de uma função gênica, ressaltando a importância do estudo detalhado de mutações que possam afetar os mecanismos de regulação gênica em BCG Moreau. Sabendo que o BCG original, obtido por Calmette e Guérin, resultou da atenuação de *M. bovis* após sucessivas passagens em meio contendo bile bovina (rica em taurina; Calmette *et al.*, 1927), é interessante especular no contexto de evolução adaptativa que a composição deste meio de cultivo levou à seleção da perda da RD16 em BCG Moreau, já que o resultado deste evento é a perda do controle transcricional do gene *rv3406*, codificante para uma putativa enzima responsável pelo catabolismo de taurina. Um exemplo deste tipo de seleção *in vitro* é a duplicação da região DU2, que inclui a glicerol-3-fosfato desidrogenase, *glpD2*, em diferentes cepas de BCG, e que está associada a um aumento de 2,7x na transcrição de *glpD2* em

comparação com *M. bovis* (Brosch *et al.*, 2007). A glicerol-3-fosfato desidrogenase auxilia no metabolismo de glicerol, e seus níveis elevados poderiam acelerar o crescimento de BCG em glicerol, presente no meio usado durante o processo de atenuação (Calmette *et al.*, 1927).

Aspectos técnicos

Neste trabalho, utilizando ferramentas de biologia molecular, conseguimos abordar alguns aspectos da regulação da expressão gênica envolvendo o gene *rv3405c*. Fizemos a clonagem, expressão e purificação das proteínas recombinantes Rv3405c de *M. bovis* BCG Pasteur e Rv3406 de *M. bovis* BCG Moreau em *E. coli*. Para facilitar a clonagem de *rv3405c* e *rv3406* no vetor de expressão pET28a utilizamos a estratégia de primeiro subclonar os produtos de PCR no vetor pBAD TOPO TA. Após verificação por sequenciamento de DNA que as sequências obtidas estavam livres de mutações, prosseguimos na cepa *E. coli* BL21 (DE3), utilizada para expressão das proteínas recombinantes rRv3405c e rRv3406 (construções agora denominadas pET-P3405c e pET-M3406, respectivamente). Essa clonagem no vetor pET28a resultou em construções que geram proteínas com fusão C terminal com cauda de 6-histidinas, sendo que a massa molecular esperada para cada uma é de 32,579 e 33,833 Da, respectivamente.

A fim de melhorar a expressão das proteínas recombinantes, várias condições de temperatura, tempo de indução e concentração do indutor foram testadas. Verificamos que a utilização do meio “overnight express”, independente de ser obtido comercialmente ou fabricado no nosso laboratório, acarretou numa pequena melhora na expressão proteica. Realizamos a purificação das proteínas a partir dos corpos de inclusão utilizando a coluna HisTrap de afinidade ao níquel e as frações coletadas foram analisadas em SDS-PAGE, sendo as mais puras reunidas em *pool* e dialisadas. Após obtenção dessas proteínas, estas foram utilizadas para imunização de camundongos Balb/c para produção de soros policlonais.

A localização de Rv3406 na superfície celular de BCG Moreau, detectada em nossos resultados de WB e proteoma (Pagani *et al.*, manuscrito em preparação), pode estar correlacionado ao maior caráter lipofílico desta proteína, identificada por Malen e colaboradores (2008) como uma proteína hidrofóbica associada à membrana de *M. bovis* BCG Moreau. Esta diferença qualitativa ressalta a importância desta proteína como um possível marcador de superfície da cepa vacinal brasileira assim como um potencial candidato para o controle de qualidade de produção da vacina BCG Moreau.

Reorganização da transcrição na região 00-05c-06

Em BCG Moreau a deleção de RD16 gera uma nova organização cromossômica na qual a sequência de *rv3405c* pode ser transcrita nas duas fitas de DNA, tanto a partir de um promotor a montante a *rv3400* quanto pelo promotor do gene *rv3405c*. Nossos resultados de RT-PCR mostram que a sequência de *rv3405c* é transcrita tanto no sentido 05c-00 (promotor de *rv3405c*, fita inferior) como no sentido 00-05 (promotor de *rv3400*, fita superior) (**Figura 5.2**).

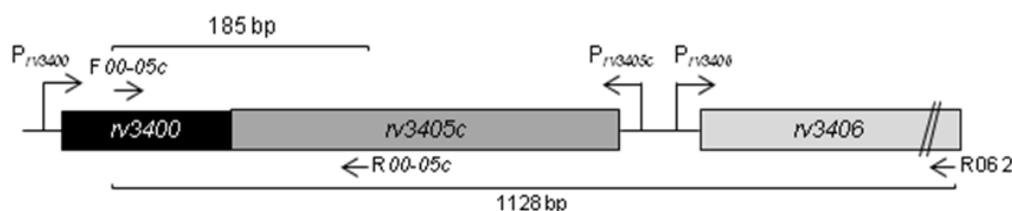


Figura 5.1: Esquema representativo do gene truncado (*rv3400-rv3405*) com seus respectivos promotores funcionais (P_{rv3400} e $P_{rv3405c}$) e *rv3406* (P_{rv3406}). Iniciadores utilizados em ensaios de RT-PCR para detectar transcritos em ambas as fitas do gene truncado (iniciadores F 00-05c e R 00-05c) e para averiguar ausência de acoplamento transcricional entre o gene *rv3400-rv3405* e *rv3406* são indicados na figura.

Uma vez que em bactérias muitos genes são codificados em operons, gerando mRNAs policistrônicos (cotranscritos em um mesmo mRNA), fizemos um RT-PCR para verificar se a deleção de RD16 poderia gerar um acoplamento transcricional a partir do promotor de *rv3400* (genes *rv3400-05c* e *rv3406*), abrindo a possibilidade de que o alto nível de transcrição de *rv3406* possa ser em função da soma da transcrição a partir do seu promotor (P_{rv3406}) e do promotor de *rv3400* (P_{rv3400}) (**Figura 5.1**). Contudo, não foram detectados cotranscritos dos genes *rv3400-05c* e *rv3406* em BCG Moreau (**Figura 4.16**). Este resultado mostra que a transcrição de *rv3406* de BCG Moreau provém exclusivamente do seu promotor, e sugere que o motivo palindrômico presente na região intergênica 05c-06 poderia funcionar como um grampo de parada para a RNA polimerase. Outros experimentos seriam necessários para comprovar a existência de um terminador transcricional nesta região.

Atividade de Rv3405c e sua modulação por pequenas moléculas

Detectamos a interação entre a proteína Rv3405c e a região intergênica *rv3405c-06*, que apresenta um motivo palindrômico, com formação de complexo quando a proteína recombinante está na concentração de 400 pM. Há formação de complexos adicionais, possivelmente dímeros ou tetrâmeros (**Figura 4.17**).

Empregamos EMSAs para identificar possíveis moduladores da atividade de ligação a DNA de Rv3405c. A primeira hipótese testada foi que a taurina, possível substrato para a putativa taurina dioxigenase Rv3406, fosse capaz de atuar como modulador da interação proteína-DNA em EMSA. Contudo, nossos dados não corroboram essa hipótese, uma vez que não ocorreu alteração nenhuma no ensaio. Em contrapartida, averiguamos por EMSA que a bile bovina, utilizada por Calmette e Guérin no meio de cultivo que levou à obtenção do BCG original, atua inibindo a formação complexo rRv3405c-região intergênica *rv3405c-rv3406* (**Figura 4.18**).

Uma vez que a bile bovina é composta por várias moléculas, é pertinente determinar qual(is) destes sais biliares poderiam atuar de forma específica, deslocando a proteína do DNA. Segundo descrito na sessão de resultados, alguns compostos presentes na bile bovina são capazes de inibir a formação do complexo em concentrações variáveis. Vimos que o tauroquenodeoxicolato de sódio e o glicoquenodeoxicolato de sódio são os sais que conseguem a inibição total do complexo em menor concentração (0,16%; 3,1 e 3,4 mM, respectivamente) (**Figura 4.19**). O fato de sais tauro- e glicoconjugados inibirem a formação de complexo Rv3405c-DNA tem precedente na literatura. O regulador CmeR de *Campylobacter jejuni* foi cristalizado na presença de dois ligantes, taurocolato e colato (Lei *et al.*, 2011). De forma surpreendente, apesar da similaridade estrutural entre estas moléculas, cada uma está posicionada de forma diferente no túnel de ligação de CmeR.

Outro exemplo da interação de componentes da bile em eventos de regulação gênica foi observado em *Salmonella typhimurium*. O operon *marRAB* de *S. typhimurium* foi identificado como sendo ativado na presença de bile e esta resposta é dependente da concentração. De todos os sais biliares, o desoxicolato é o único capaz de ativar essa transcrição. Desoxicolato é capaz de interagir com MarR e interferir na sua capacidade de se ligar ao operador *mar*. A incubação de *Salmonella* na presença de concentrações subletais de bile mostra que esta é capaz de aumentar a sua resistência a bile (*mar*-dependente) e cloranfenicol (*mar*-independente). A fim de caracterizar genes putativos regulados por *marRAB*, tendo sua relevância no fenótipo de resistência, genes *acrAB* que codificam uma bomba de efluxo foram analisados. Em *S. typhimurium*, *acrAB* é requerido para resistência a bile, porém enquanto a transcrição de *acrAB* é ativada por bile, esta ativação é independente de *marRAB* bem como Rob, RpoS ou PhoP-PhoQ. Estes dados sugerem que a bile interage com *Salmonella* para aumentar sua resistência à bile e outros agentes antimicrobianos. Isto pode acontecer por vias *marRAB*- e *acrAB*- dependentes que funcionam independentemente com relação a ativação por bile (Prouty *et al.*, 2004). Estes resultados indicam que o desoxicolato interage com MarR para prevenir a ligação do DNA e que esta interação é

específica para deoxicolato. Em *E. coli*, deoxicolato e quenodeoxicolato são capazes de se ligar ao domínio C-terminal de Rob, que afeta a transcrição de *acrAB* (Rosenberg *et al.*, 2003).

A fim de testar se as interações entre Rv3405c e bile ou seus componentes observadas em ensaios de retardo em gel ocorrem *in vivo*, foram executados experimentos de RT-PCR com RNA extraído de cultivos de BCG Pasteur e BCG Moreau crescidos na presença destes compostos. Estes resultados demonstram que o gene *rv3406* é constitutivamente expresso em BCG Moreau, de forma ligeiramente variável de acordo com a presença das moléculas moduladoras testadas. A adição de taurina em cultura de BCG Pasteur parece ter induzido levemente a expressão de *rv3406*. Já a presença de bile, tauroquenodeoxicolato de sódio e glicoquenodeoxicolato de sódio induzem de forma clara a expressão de *rv3406* em cultura de BCG Pasteur. Porém, estes resultados são de apenas um único experimento e a imagem não foi analisada de forma quantitativa, de modo que será necessário confirmar esses dados. A presença da proteína Rv3406 não foi detectada nestas culturas de BCG Pasteur, sugerindo possível regulação pós-transcricional dos níveis de Rv3406.

Ensaio preliminares executados no laboratório durante a redação desta tese mostram que a atividade enzimática de rRv3406 é detectada quando as reações contém tauroquenodeoxicolato de sódio, mas não taurina (dados não mostrados). Contudo, as reações enzimáticas devem ser otimizadas para confirmar estes resultados. É interessante ressaltar que enzimas da família TauD podem usar outros substratos de forma eficiente (Eichhorn *et al.*, 1997), e que tratamento de *M. tuberculosis* com baixas concentrações de SDS (que contém grupamento sulfato) induz a transcrição de *rv3406* (Manganelli *et al.*, 2001). Assim, a aparente baixa indução de *rv3406* por taurina e a incapacidade de Rv3406 de usá-la como substrato (pelo menos nas condições testadas até o momento) podem indicar uma limitação reconhecida na anotação automática de proteínas da superfamília das dioxigenases alfacetogluturato-dependentes (Kundu, 2012). Neste contexto, será interessante definir os substratos usados por Rv3406, incluindo os sais biliares glicoconjugados que deslocam Rv3405c do DNA e induzem transcrição de *rv3406* em BCG Pasteur.

Genes putativos do regulon de Rv3405c

A fim de analisar a conservação da região intergênica entre *rv3405c* e *rv3406* foi feita uma busca por homólogos sintênicos, que apresentam tanto o gene *rv3405c* quanto o gene *rv3406*, através do programa BLAST. O alinhamento das regiões reguladoras encontradas mostra que o motivo palindrômico, alvo de Rv3405c, é conservado entre diversas espécies de micobactérias.

Uma conexão entre o regulon putativo de Rv3405c e uma possível função desse repressor transcricional é encontrada em um experimento descrevendo o perfil transcricional de *M. tuberculosis* em resposta ao surfactante de pulmão bovino (Schwab *et al.*, 2009). O *rv3405c* apresenta um perfil transcricional alterado em *M. tuberculosis* (indução) após tratamento com surfactante.

Rv3406, a tuberculose, e propriedades vacinais de BCG Moreau

Um fator importante na progressão da tuberculose é o surfactante pulmonar, uma mistura complexa de proteínas e lipídios sintetizados, empacotados, estocados e secretados por células epiteliais no ambiente alveolar. As interações entre o *M. tuberculosis* e componentes do surfactante influenciam na interação entre a micobactéria e macrófagos (revisado em Ferguson e Schlesinger, 2000; Li *et al.*, 2012). De fato, a interação do surfactante de pulmão com patógenos (bacterianos, eucarióticos e virais) é importante na oscilação entre resposta imune eficaz e infecção (revisado em Glasser e Mallampalli, 2012).

No trabalho de Schwab *et al.*, 2009 foi observada a resposta transcricional de *M. tuberculosis* ao surfactante e subfrações de seus componentes. O foco particular foi na indução/repressão de genes relevantes nas interações macrófago-*M. tuberculosis* e os genes codificantes para lipases bacterianas que devem utilizar os fosfolípídeos surfactantes como substrato. Em suma, os dados obtidos por eles indicam que o surfactante pulmonar afeta diretamente a transcrição de genes do *M. tuberculosis*, modulando sua interação com os macrófagos e proporcionando também o aproveitamento de lipídeos para seu crescimento. São necessários estudos mais específicos para confirmar se há uma associação entre Rv3405c e a resposta ao surfactante pulmonar, em especial no contexto funcional de BCG Moreau. A presença de mRNA de *rv3406* foi detectada em escarro de pacientes com TB. A transcrição desse gene foi significativamente maior ($p < 0,05$) em escarro do que em crescimento aeróbico (Garton *et al.*, 2008), implicando este gene na fisiologia da tuberculose.

Supomos um paralelo entre a interação entre *M. tuberculosis* e surfactante e o crescimento em bile bovina durante atenuação de *M. bovis*, usada por seu efeito detergente (Watt *et al.*, 1984), minimizando o crescimento em grumos da bactéria. Verificamos a similaridade estrutural entre os sais de bile, que são os componentes principais da bile sintetizados a partir do colesterol sendo glico- ou tauro-conjugados (revisto por Begley *et al.*, 2005) e colesterol e seus ésteres presentes no surfactante pulmonar (Hallman & Gluck, 1976) e nos granulomas caseosos (Kim *et al.*, 2010). O uso do colesterol como fonte de carbono é necessário para virulência do *M. tuberculosis* (Pandey & Sasseti, 2008; Yam *et al.*, 2009; Chang *et al.*, 2009; Nesbitt *et al.*, 2010), sendo também necessário para sobrevivência em macrófagos (Rengarajan *et al.*, 2005). Os níveis de colesterol e seus ésteres são elevados nos granulomas caseosos em comparação com tecido pulmonar normal (Kim *et al.*, 2010), e o colesterol sulfato está presente no pulmão de coelhos (Cui & Iwamori, 1997). Desta forma, sugerimos que o perfil transcricional do *M. tuberculosis* em resposta ao surfactante de pulmão possa ter semelhança com o perfil encontrado em BCG crescido em cultura na presença de bile bovina. A indução de *rv3405c* por surfactante implicaria que os genes de seu regulon (e, portanto, sob seu controle) terão sua transcrição alterada. Estas observações sugerem que enzimas que hidrolisam ésteres de colesterol podem ter um importante papel na liberação do colesterol a ser importado por *M. tuberculosis*. Desta forma, níveis mais altos de *rv3406* em BCG Moreau, assim como alterações nos níveis de outros genes no regulon de *rv3405c*, poderiam resultar em uma alteração no perfil metabólico da bactéria e em sua propagação, modulando suas características imunoprotetoras.

6. CONCLUSÕES

O presente trabalho contribui para o melhor conhecimento da cepa vacinal brasileira, *M. bovis* BCG Moreau, através da definição do papel funcional de um repressor transcricional (Rv3405c) e seu impacto direto na fisiologia do BCG envolvendo o metabolismo de enxofre via regulação da expressão da alquil sulfatase Rv3406.

- ✓ A região intergênia *rv3405c- rv3406* possui um motivo palindrômico conservado entre diversas espécies de micobactérias.
- ✓ O gene *rv3406* é transcrito de forma constitutiva em BCG Moreau e não é expresso em BCG Pasteur nas condições avaliadas. Esses resultados corroboram dados de proteômica do laboratório que identificam a presença da proteína somente em BCG Moreau.
- ✓ A proteína Rv3406 está presente na fração intracelular e na de superfície de BCG Moreau, porém ausente nas mesmas frações em BCG Pasteur. Já no filtrado de cultura, Rv3406 está ausente tanto em BCG Moreau como em BCG Pasteur.
- ✓ O gene *rv3405c* é transcrito em BCG Pasteur e Moreau, sendo transcrita a sua forma quimérica (*rv3405c-00*) em BCG Moreau.
- ✓ O gene quimérico *rv3400-05c* é transcrito em ambas as direções em BCG Moreau.
- ✓ Os genes *rv3400-05c* e *rv3406* não são co-transcritos em BCG Moreau. Este dado sugere a hipótese de que o motivo palindrômico presente na região intergênica *05c-06* pode funcionar como um grampo de parada para a RNA polimerase.
- ✓ A proteína Rv3405c interage com a região intergênica *rv3405c-06*.
- ✓ A adição de taurina não altera a formação do complexo Rv3405c-DNA, porém esta interação é modulada por componentes da bile bovina.
- ✓ Tanto o tauro- quanto o glico-conjugado do ácido quenodeoxicolato se mostraram mais eficientes em inibir a ligação Rv3405cDNA.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- Alexander DC, Behr MA. Rv1773 is a transcriptional repressor deleted from BCG-Pasteur. 2007. *Tuberculosis (Edinb)* **87**:421-5.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. 1990. *J Mol Biol.* **215**:403-10.
- Barreto AMW, Melo FF, Campos HS, Castello Branco LR, Caldas PCS, Antas PRZ. O adoecimento. Controle da tuberculose: Uma proposta de integração ensino-serviço. Ministério da Saúde, Brasil, Fundação Nacional de Saúde. 2002. 5ª ed. Rio de Janeiro: FUNASA/CRPHF/SBPT.
- Begley M, Gahan CG, Hill C. The interaction between bacteria and bile. 2005. *FEMS Microbiol Rev.* **29**:625-51.
- Behr, MA, Small, PM. A historical and molecular phylogeny of BCG strains. 1999. *Vaccine.* **17**:915-22.
- Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S. Comparative genomics of BCG vaccines by whole-genome DNA microarray. 1999. *Science.* **284**:1520-3.
- Behr MA, Schroeder BG, Brinkman JN, Slayden RA, Barry CE. A point mutation in the *mma3* gene is responsible for impaired methoxymycolic acid production in *Mycobacterium bovis* BCG strains obtained after 1927. 2000. *J Bacteriol.* **182**:3394-9.
- Behr MA. Correlation between BCG genomics and protective efficacy. 2001. *Scand J Infect Dis.* **33**:249-52.
- Behr MA. BCG-different strains, different vaccines? 2002. *Lancet Infect Dis.* **2**:86-92.
- Benévolo-de-Andrade TC, Monteiro-Maia R, Cosgrove C, Castello-Branco LRR. BCG Moreau Rio de Janeiro: an oral vaccine against tuberculosis - review. 2005. *Mem. Inst. Oswaldo Cruz.* **100**:459-65.
- Berrêdo-Pinho M, Kalume DE, Correa PR, Gomes LH, Pereira MP, da Silva RF, Castello-Branco LR, Degraive WM, Mendonça-Lima L. Proteomic profile of culture filtrate from the Brazilian vaccine strain *Mycobacterium bovis* BCG Moreau compared to *M. bovis* BCG Pasteur. 2011. *BMC Microbiol.* **20**:11:80.
- Bertram R, Hillen W. The application of Tet repressor in prokaryotic gene regulation and expression. 2008. *Microb Biotechnol.* **1**:2-16.

- Besra GS, Khoo KH, McNeil MR, Dell A, Morris HR, Brennan PJ. A new interpretation of the structure of the mycolyl-arabinogalactan complex of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and ¹H nuclear magnetic resonance spectroscopy. 1995. *Biochemistry*. **34**:4257-66.
- Bonah C. The 'experimental stable' of the BCG vaccine: safety, efficacy, proof and Standards, 1921-1933. 2005. *Stud Hist Philos Biol Biomed Sci Dec*. **36**:696-721.
- Brandsch M, Ramamoorthy S, Marczin N, Catravas, JD, Leibach, JW, Ganapathy, V, Leibach, FH. Regulation of taurine transport by *Escherichia coli* heat-stable enterotoxin and guanylin in human intestinal cell lines. 1995. *J Clin Invest*. **96**:361-9.
- Brandt L, Feino Cunha J, Weinreich Olsen A, Chilima B, Hirsch P, Appelberg R, Andersen P. Failure of the *Mycobacterium bovis* BCG vaccine: some species of environmental mycobacteria block multiplication of BCG and induction of protective immunity to tuberculosis. 2002. *Infect Immun*. **70**:672-8.
- Brennan PJ, Nikaido H. The envelope of Mycobacteria. 1995. *Annual Review of Biochemistry*. **64**:29-63.
- Brennan PJ. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. 2003. *Tuberculosis*. **83**:91-7.
- Brosch R, Pym AS, Gordon SV, Cole ST. The evolution of mycobacterial pathogenicity: clues from comparative genomics. 2001. *Trends Microbiol*. **9**:452-8.
- Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, Van Soolingen D, Cole ST. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. 2002. *Proc Natl Acad Sci USA*. **99**:3684-9.
- Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C, Inwald JK, Golby P, Garcia JN, Hewinson RG, Behr MA, Quail MA, Churcher C, Barrell BG, Parkhill J, Cole ST. Genome plasticity of BCG and impact on vaccine efficacy. 2007. *Proc Natl Acad Sci USA*. **104**:5596-601.
- Calmette A, Guérin C, Boquet A, Nègre L. La vaccination préventive contre la tuberculose par le "BCG". Paris: Masson et cie; 1927.
- Casanova JL, Abel L. Genetic dissection of immunity to mycobacteria: the human model. 2002. *Annu Rev Immunol*. **20**:581-620.

- Cervino AC, Lakiss S, Sow O, Bellamy R, Beyers N, Hoal-van Helden E, van Helden P, McAdam K P, Hill A V. 7. Fine mapping of a putative tuberculosis-susceptibility locus on chromosome 15q11-13 in African families. 2002. *Hum Mol Genet.* **11**:1599-603.
- Chan J, Kaufmann SHE. Pathogenesis, Protection and Control. In *Tuberculosis*, ed. BR Bloom. 1994. 271-284. Edited by Am Soc Microbiol, Washington DC.
- Chang JC, Miner MD, Pandey AK, Gill WP, Harik NS, Sasseti CM, Sherman DR. *igr* Genes and *Mycobacterium tuberculosis* cholesterol metabolism. 2009. *J Bacteriol.* **191**:5232-9.
- Charlet D, Mostowy S, Alexander D, Sit L, Wiker HG, and Behr MA. Reduced expression of antigenic proteins MPB70 and MPB83 in *Mycobacterium bovis* BCG strains due to a start codon mutation in sigK. 2005. *Mol Microbiol.* **56**:1302-1313.
- Chen JM, Islam ST, Ren H, Liu J. Differential productions of lipid virulence factors among BCG vaccine strains and implications on BCG safety. 2007. *Vaccine.* **25**:8114-22.
- Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry III CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Conner R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. 1998. *Nature.* **393**:537-44.
- Collins FM. Relative susceptibility of acid-fast and non-acid-fast bacteria to ultraviolet light. 1971. *Appl Microbiol.* **21**:411-3.
- Cui Y, Iwamori M. Distribution of cholesterol sulfate and its anabolic and catabolic enzymes in various rabbit tissues. 1997. *Lipids.* **32**:599-604.
- Daffe M, Etienne G. The capsule of *Mycobacterium tuberculosis* and its implications for pathogenicity. 1999. *Tubercle and Lung Disease.* **79**:153-169.
- Dai J, Chen Y, Lauzardo M. Web-accessible database of *hsp65* sequences from *Mycobacterium* reference strains. 2011. *Journal Clin Microb.* **49**:2296-2303.
- Doerks T, van Noort V, Minguéz P, Bork P. Annotation of the *M. tuberculosis* hypothetical orfeome: adding functional information to more than half of the uncharacterized proteins. 2012. *PLoS One* **7**: e34302.

- Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. 1999. WHO Global Surveillance and Monitoring Project. *JAMA*. **282**:677-86.
- Eichhorn E, van der Ploeg JR, Kertesz MA, Leisinger T. Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. 1997. *J Biol Chem*. **272**:23031-6.
- Ferguson JS, Schlesinger LS. Pulmonary surfactant in innate immunity and the pathogenesis of tuberculosis. 2000. *Tuber Lung Dis*. **80**:173-84.
- Folcher M, Morris RP, Dale G, Salah-Bey-Hocini K, Viollier PH, Thompson CJ. A transcriptional regulator of a pristinamycin resistance gene in *Streptomyces coelicolor*. 2001. *J Biol Chem*. **276**:1479-85.
- Garton NJ, Waddell SJ, Sherratt AL, Lee SM, Smith RJ, Senner C, Hinds J, Rajakumar K, Adegbola RA, Besra GS, Butcher PD, Barer MR. Cytological and transcript analyses reveal fat and lazy persistor-like bacilli in tuberculous sputum. 2008. *PLoS Med*. **5**:e75.
- Glasser JR, Mallampalli RK. Surfactant and its role in the pathobiology of pulmonary infection. 2012. *Microbes Infect*. **14**:17-25.
- Gomes LH, Otto TD, Vasconcellos EA, Ferrão PM, Maia RM, Moreira AS, Ferreira MA, Castello-Branco LR, Degraive WM, Mendonça-Lima L. Genome sequence of *Mycobacterium bovis* BCG Moreau, the Brazilian vaccine strain against tuberculosis. 2011. *J Bacteriol*. **193**:5600-1.
- Gomez JE, McKinney JD. *M. tuberculosis* persistence, latency, and drug tolerance. 2004. *Tuberculosis (Edinb)*. **84**:29-44.
- Hallman M, Gluck L. Phosphatidylglycerol in lung surfactant. III. Possible modifier of surfactant function. 1976. *J Lipid Res*. **17**:257-62.
- Honda I, Seki M, Ikeda N, Yamamoto S, Yano I, Koyama A, Toida I. Identification of two subpopulations of Bacillus Calmette-Guérin (BCG) Tokyo172 substrain with different RD16 regions. 2006. *Vaccine*. **24**:4969-74.
- Kendall SL, Withers M, Soffair CN, Moreland NJ, Gurucha S, Sidders B, Frita R, Ten Bokum A, Besra GS, Lott JS, Stoker NG. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. 2007. *Mol Microbiol*. **65**:684-99.

- Kendall SL, Burgess P, Balhana R, Withers M, Ten Bokum A, Lott JS, Gao C, Uhia-Castro I, Stoker, NG. Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: *kstR* and *kstR2*. 2010. *Microbiol.* **156**:1362-71.
- Kim MJ, Wainwright HC, Locketz M, Bekker LG, Walther GB, Dittrich C, Visser A, Wang W, Hsu FF, Wiehart U, Tsenova L, Kaplan G, Russell DG. Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism. 2010. *EMBO Mol Med.* **2**:258-74.
- Kramnik I, Dietrich WF, Demant P, Bloom B R. Genetic control of resistance to experimental infection with virulent *Mycobacterium tuberculosis*. 2000. *Proc Natl Acad Sci USA.* **97**:8560-5.
- Kundu S. Distribution and prediction of catalytic domains in 2-oxoglutarate dependent dioxygenases. 2012. *BMC Res Notes.* **5**:410
- Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. 1970. *Nature.* **227**:680-5.
- Lei HT, Shen Z, Surana P, Routh MD, Su CC, Zhang Q, Yu EW Crystal structures of CmeR-bile acid complexes from *Campylobacter jejuni*. 2011. *Protein Sci.* **20**:712-23.
- Lew JM, Kapopoulou A, Jones LM, Cole ST. TubercuList - 10 years after. 2011. *Tuberculosis (Edinb).* **91**:1-7.
- Li Y, Wang Y, Liu X. The role of airway epithelial cells in response to mycobacteria infection. 2012. *Clin Dev Immunol.* **2012**:791392.
- Liu J, Tran V, Leung AS, Alexander DC, Zhu B. BCG vaccines: Their mechanisms of attenuation and impact on safety and protective efficacy. 2009. *Human Vaccines.* **5**:70-8.
- Malen H, Berven FS, Softeland T, Arntzen MO, D'Santos CS, De Souza GA. Membrane and membrane-associated proteins in Triton X-114 extracts of *Mycobacterium bovis* BCG identified using a combination of gel-based and gel-free fractionation strategies. 2008. *Proteomics.* **8**:1859-70.
- Manganelli R, Voskuil MI, Schoolnik GK, Smith I. The *Mycobacterium tuberculosis* ECF sigma factor *sigmaE*: role in global gene expression and survival in macrophages. 2001. *Mol Microbiol* **41**:423-37
- Marquet S, Schurr E. Genetics of susceptibility to infectious diseases: tuberculosis and leprosy as examples. 2001. *Drug Metab Dispos.* **29**:479-83.

- Marri PR, Bannantine JP, Golding GB. Comparative genomics of metabolic pathways in *Mycobacterium* species: gene duplication, gene decay and lateral gene transfer. 2006. FEMS Microbiol Rev. **30**:906-25.
- Mitsos LM, Cardon LR, Fortin A, Ryan L, LaCourse R, North RJ, Gros P. Genetic control of susceptibility to infection with *Mycobacterium tuberculosis* in mice. 2000. Genes Immun. **1**:467-77.
- Mostowy S, Tsolaki AG, Small P M, Behr M A. The *in vitro* evolution of BCG vaccines. 2003. Vaccine. **21**:4270-4.
- Mostowy S, Cleto C, Sherman DR, Behr MA. The *Mycobacterium tuberculosis* complex transcriptome of attenuation. 2004. Tuberculosis (Edinb). **84**:197-204.
- Nesbitt NM, Yang X, Fontán P, Kolesnikova I, Smith I, Sampson NS, Dubnau E. A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. 2010. Infect Immun. **78**:275-82.
- Newton-Foot M, Gey van Pittius NC. The complex architecture of mycobacterial promoters. 2012. Tuberculosis **93**:60-74.
- North RJ, Jung YJ. Immunity to tuberculosis. 2004. Annu Rev Immunol. **22**:599-623.
- Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. 2000. J Mol Biol. **302**:205-17.
- Nunn P, Williams B, Floyd K, Dye C, Elzinga G, Raviglione M. Tuberculosis control in the era of HIV. 2005. Nat Rev Immunol **5**:819-26.
- Oettinger T, Jorgensen M, Ladefoged A, Haslov K, Andersen P. Development of the *Mycobacterium bovis* BCG vaccine: review of the historical and biochemical evidence for a genealogical tree. 1999. Tuber Lung Dis **79**:243-50.
- Ortalo-Magné A, Dupont MA, Lemassu A, Andersen AB, Gounon P, Daffé M. Molecular composition of the outermost capsular material of the tubercle bacillus. 1995. Microbiol. **141**:1609-20.
- Pandey AK, Sasseti CM. Mycobacterial persistence requires the utilization of host cholesterol. 2008. Proc Natl Acad Sci U S A. **105**:4376-80.
- Park H, Ro YT, Kim YM. MdoR is a novel positive transcriptional regulator for the oxidation of methanol in *Mycobacterium* sp. strain JC1. 2011. J Bacteriol. **193**:6288-94.

- Prouty AM, Brodsky IE, Falkow S, Gunn JS. Bile-salt-mediated induction of antimicrobial and bile resistance in *Salmonella typhimurium*. 2004. *Microbiol.* **150**:775-83.
- Pym AS, Brodin P, Brosch R, Huerre M, Cole ST. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. 2002. *Mol Microbiol.* **46**:709-17.
- Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, Zhang, Xiaodong, Gallegos MT, Brennan R, Tobes R. The TetR family of transcriptional repressors. 2005. *Microbiol Mol Biol.* **69**:326-56.
- Raviglione MC, O'Brien RJ. Harrison's principals of internal medicine, 14a ed. EUA: McGraw-Hill; 1998.
- Rengarajan J, Bloom BR, Rubin EJ. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. 2005. *Proc Natl Acad Sci U S A.* **102**:8327-8332
- Riley RL, Knight M, Middlebrook G. Ultraviolet susceptibility of BCG and virulent tubercle bacilli. 1976. *Am Rev Respir Dis.* **113**:413-8.
- Rosenberg EY, Bertenthal D, Nilles ML, Bertrand KP, Nikaido H. Bile salts and fatty acids induce the expression of *Escherichia coli* AcrAB multidrug efflux pump through their interaction with Rob regulatory protein. 2003. *Mol Microbiol.* **48**:1609-1619.
- Russell DG. *Mycobacterium tuberculosis*: here today, and here tomorrow. 2001. *Nat Rev Mol Cell Biol.* **2**:569-77.
- Sachdeva P, Misra R, Tyagi AK, Singh Y. The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. 2010. *FESB J.* **277**:605-26.
- Sakula A. BCG: who were Calmette and Guerin? 1983. *Thorax.* **38**:806-12.
- Saunders BM, Cooper AM. Restraining mycobacteria: role of granulomas in mycobacterial infections. 2000. *Immunol Cell Biol.* **78**:334-41.
- Schwab, U., Rohde K. H., Wang, Z., Chess, P. R., Notter, R. H., Russel, D. G. Transcriptional responses of *Mycobacterium tuberculosis* to lung surfactant. 2009. *Microbial Pathogenesis.* **46**:185-193.
- Spreadbury CL, Pallen MJ, Overton T, Behr MA, Mostowy S, Spiro S, Busby SJ, Cole JA. Point mutations in the DNA- and cNMP-binding domains of the homologue of the cAMP receptor protein (CRP) in *Mycobacterium bovis* BCG: implications for the inactivation of a global regulator and strain attenuation. 2005. *Microbiol.* **151**: 547-56.

- Stewart GR, Robertson BD, Young DB. Tuberculosis: a problem with persistence. 2003. *Nat Rev Microbiol.* **1**:97-105.
- Tufariello JM, Chan J, Flynn JL. Latent tuberculosis: mechanisms of host and bacillus that contribute to persistent infection. 2003. *Lancet Infect Dis.* **3**:578-90.
- Uhía I, Galán B, Medrano FJ, García JL. Characterization of the KstR-dependent promoter of the gene for the first step of the cholesterol degradative pathway in *Mycobacterium smegmatis*. 2011. *Microbiology.* **157**:2670-80
- van der Ploeg JR, Weiss MA, Saller E, Nashimoto H, Saito N, Kertesz MA, Leisinger T. Identification of sulfate starvation-regulated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. 1996. *J Bacteriol.* **178**:5438-46.
- Yam KC, D'Angelo I, Kalscheuer R, Zhu H, Wang JX, Snieckus V, Ly LH, Converse PJ, Jacobs WR Jr, Strynadka N, Eltis LD. Studies of a ring-cleaving dioxygenase illuminate the role of cholesterol metabolism in the pathogenesis of *Mycobacterium tuberculosis*. 2009. *PLoS Pathog.* **5**:e1000344.
- Yu, Z., Reichheld, S.E., Savchenko, A., Parkinson, J., and Davidson, A.R. A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators. 2010. *J. Mol. Biol.* **400**:847-864.
- Watt, SM., Simmonds, WJ. Effects of four taurine-conjugated bile acids on mucosal uptake and lymphatic absorption of cholesterol in the rat. 1984. *J. Lipid Res.* **25**:448-55.
- WHO. TB, a global emergency. 1994. WHO report on the TB epidemic. Geneva.
- WHO. 2012. Global Tuberculosis Report.

8. ANEXOS

ANEXO I - Materiais utilizados neste estudo

Meios de Cultura para *E. coli*:

Meio LB (Luria-Bertani)

Bacto-triptona (Difco)	10,0 g/L
Extrato de levedura (Difco)	5,0 g/L
NaCl	10,0 g/L

pH ajustado para 7,2 com NaOH

Meio LB / ágar: Meio LB acrescido de 1,5% (p/v) bacto-ágar (Difco)

Meios de Cultura para Micobactérias:

Meio Sauton

L-Asparagina	4 g
Ácido Cítrico	2 g
K ₂ HPO ₄	0,5 g
MgSO ₄	0,5 g
Citrato Férrico Amoniacal	0,05 g
Zn ₂ SO ₄	5 mM
Glicerol	60 mL
H ₂ O Milli-RO	qsp 1000 mL

pH ajustado para 7,2 com NaOH

ADC: Enriquecimento para meios Middlebrook 7H9 e 7H10

Albumina, (Sigma, fração V)	50g /L
Dextrose (Glicose)	20g /L
Catalase (beef)	0,03g /L
NaCl	8,5g /L

Cepas bacterianas utilizadas nesse estudo

Cepa	Características	Fonte/referência
<i>E. coli</i> Top 10	<i>F mcrA Δ(mrr-hsdRMS-MCR) φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galK rpsL (Str^R)endA1 nupG</i>	Invitrogen
<i>E. coli</i> BL21(DE3)	F- <i>ompT hsdSB(rB- mB-) gal dcm</i> (DE3)	Novagen
BCG Pasteur	<i>M. bovis</i> BCG cepa Pasteur	Instituto Pasteur
BCG Moreau	<i>M. bovis</i> BCG cepa Moreau	Fundação Ataulpho de Paiva

Plasmídeos utilizados nesse estudo

Plasmídeos	Características	Fonte/Referência
pBadTopo TA (3931 pb)	Amp ^R ; pUC <i>ori</i> ; P _{Lac} ; <i>lacZα</i>	Invitrogen
pET28a (5369 pb)	Kan ^R ; f1 <i>ori</i> ; pBR322 <i>ori</i> ; <i>Lac1</i> ; P _{T7} ; His tag; T7 tag	Novagen
pET-P3405c (5949 pb)	pET28a; gene <i>rv3405c</i> de <i>M. bovis</i> BCG Pasteur	este trabalho
pET-M3406 (6270 pb)	pET28a; <i>rv3406</i> de <i>M. bovis</i> BCG Moreau	este trabalho

Tampões para DNA:**TE 1X**

Tris-HCl (pH8,0)	10 mM
EDTA (sal dissódico; pH 8,0)	1 mM

TBE 10X

Tris-base	108 g/L
Ácido bórico	55 g/L
EDTA 0,5 M (pH8,0)	40 mL /L

Tampão (5X) para aplicação de amostras em gel de agarose

Azul de Bromofenol	0,25% (p/v)
Xileno Cianol FF	0,25% (p/v)
Glicerol	30% (p/v)

Soluções utilizadas para géis SDS-PAGE:**Tampão de Focalização Isoelétrica (IEF)**

Uréia	8 M
Chaps	2 %

Tampão (1X) para aplicação de proteínas em gel SDS-PAGE

Tris-HCl (pH6,8)	62,5 mM
SDS	2%
β -mercaptoetanol	5% (v/v)
Azul de Bromofenol	0,002%
Glicerol	10% (v/v)

Tampão de corrida Laemmli (5X)

Tris-base	15 g/L
Glicina	72 g/L
SDS	5 g/L

Preparo de géis SDS-PAGE:**Gel de separação 15%**

Tris-HCl 1,5 M (pH 8,8)	1,5 mL
Acrilamida 30% / Bis-acrilamida 0,8%	3,0 mL
SDS 10%	60 μ L
APS 10%	60 μ L
TEMED	6 μ L
H ₂ O Milli-Q	qsp 6mL

Gel de empacotamento 4%

Tris-HCl 1,0 M (pH 6,8)	0,5 mL
Acrilamida 30% / Bis-Acrilamida 0,8%	0,67 mL
SDS 10%	40 μ L
APS 10%	40 μ L
TEMED	4 μ L
H ₂ O Milli-Q	qsp 4mL

Soluções de coloração para géis SDS-PAGE:**Coloração com Coomassie Brilliant Blue R-250**

Etanol P.A	30%
Ácido acético glacial	7%
Coomassie Brilliant Blue R-250 (Bio-Rad)	0,5%

Coloração pela prata (Método de Ansorge - Ansorge, 1985)

Reagente	Tempo
10% Etanol/5% HOAc	10'
2% CuCl ₂ /12% TCA/50% Etanol	20'
10% Etanol/5% ácido acético	10'
0,01% KMnO ₄	10'
10% Etanol/5% ácido acético	10'
10% Etanol	10'
H ₂ O	10'
0,1% AgNO ₃	10'
H ₂ O	5''
10% K ₂ CO ₃	1'
2% K ₂ CO ₃ /0,01% formaldeído	4 x 5'
10% EtOH/5% HOAc	5'
H ₂ O ou sol. secagem ou 1% HOAc	

Solução de secagem

Etanol P.A.	20%
Glicerol	2%

Tampões para Western Blot:**Tampão para eletro-transferência de proteínas (Trans-Blot, BioRad)**

Tris-base	25 mM
Glicina	192 mM

Tampão TBS (10X)

1M Tris-HCl, pH 7,5	100 mL
NaCl	87g
H ₂ O	qsp 1000 mL

Tampões para purificação de proteínas recombinantes:**Tampão de lise bacteriana**

1M Tris-HCl, pH 7,5	20 mM
KCl	100 mM
EDTA	5 mM

Solução de lisozima

10 mg/mL em 100 mM Tri-HCl pH 8,0

Tampão Triton X-100 (para lavagem dos corpos de inclusão)

Tris-HCl pH 8,5	50 mM
Triton X-100	0,5%
EDTA	5 mM
NaCl	150 mM

Solução desnaturante para solubilização de corpos de inclusão

Tris-HCl pH 8,5	50 mM
Uréia	8 M
NaCl	300 mM

Tampões para cromatografia de afinidade a metal (Ni^{2+} -IMAC):

Tampão de ligação pH 7,5

NaH_2PO_4	20 mM
Uréia	8 M
NaCl	0,5 M
Imidazol	20 mM

Tampão de eluição pH 7,5

NaH_2PO_4	20 mM
Uréia	8 M
NaCl	0,5 M
Imidazol	500 mM

PBS 10X pH 7,2

NaH_2PO_4	14,4 g/L
NaCl	80 g/L
KCl	2 g/L
KH_2PO_4	2,4 g/L

H₂O DEPC: 0,1 mL DEPC p/ 1 Litro de H₂O Milli-Q autoclavada (0,01%)

Anexo II - Oligonucleotídeos utilizados nesse estudo

Nome	Orientação	Sequência	Finalidade
Intergênica 05c-06	Senso	ACGGTAATCAGATCTGTTCAT	EMSA
	Antisenso	TTGCCGACGCGTAGTCATC	
00-05c	Senso	ACTATCCGGAAGTGAGGTCCC	RT-PCR
	Antisenso	CCTGGGCACGAAGCTGACCA	
<i>NcoI/XhoI</i> Rv3406	Senso	GAGTCCATGGGAATGACAGATCTGATTACCGTGAA GAAG	Clonagem em pET28a
	Antisenso	ACTCCTCGAGGCCAGCGATCTCCAT	
<i>NcoI/XhoI</i> Rv3405c	Senso	GAGACCATGGGAATGACTACGCGTCCGG	Clonagem em pET28a
	Antisenso	TCTCCTCGAGGTGCGGTTCCAGGA	
Rv3406	Senso	ACGGCGAATCACCATGCCCCG	RT-PCR
	Antisenso	AGCCAGCGATCTCCATCGGC	
Rv3405c Pasteur/ Moreau	Senso	ACTACGCGTCCGGCAACC	RT-PCR
	Antisenso (Pasteur)	GTGCGGTTCCAGGATCCGGG	
	Antisenso (Moreau)	TCCGGAAGTGAGGTCCC	
3400-05c Rv3406 2	Senso	ACTATCCGGAAGTGAGGTCCC	RT-PCR
	Antisenso	GCAGCACTTCGAATAACACC	

Anexo III - Condições de Ciclagem para PCRs**Região intergênica 05c-06**

Temperatura	Tempo	Número de ciclos
95°C	3`	1x
95°C	30``	
55°C	25``	35x
72°C	25``	
72°C	4`	1x
16°C	∞	∞

Genes *rv3405c* e *rv3406* para clonagem em pET28a

Temperatura	Tempo	Número de ciclos
95°C	5`	1x
95°C	30``	
56°C	30``	35x
72°C	1`45``	
72°C	4`	1x
16°C	∞	∞

RT-PCR para os genes *rv3405c* e *rv3406* em BCG Moreau e Pasteur

Temperatura	Tempo	Número de ciclos
95°C	5`	1x
95°C	30``	
60°C	30``	35x
72°C	30``	
72°C	4`	1x
16°C	∞	∞

RT-PCR para o gene *rv3400-05c* em BCG Moreau

Temperatura	Tempo	Número de ciclos
95°C	5´	1x
95°C	30´´	
56°C	25´´	35x
72°C	25´´	
72°C	4´	1x
16°C	∞	∞

RT-PCR para o gene *rv3400-06* em BCG Moreau

Temperatura	Tempo	Número de ciclos
95°C	5´	1x
95°C	30´´	
56°C	1´	35x
72°C	1´	
72°C	4´	1x
16°C	∞	∞

Anexo IV - Manuscrito submetido para publicação na revista “Tuberculosis”

A *Mycobacterium bovis* BCG Moreau-specific genomic deletion (RD16) inactivates a transcriptional repressor and reshapes transcription of adjacent regions

Teca Calcagno Galvão^{c,&}, Cristiane Rodrigues Lima^{a,&}, Leonardo Henrique Ferreira Gomes^a, Talita Duarte Pagani^a, Marcelo Alves Ferreira^a, Antonio S. Gonçalves^b, Paloma Rezende Correa^a, Wim Maurits Degrave^a, Leila Mendonça-Lima^{a,*}

^aLaboratório de Genômica Funcional e Bioinformática and ^bLaboratório de Biotecnologia e Fisiologia de Infecções Virais, Instituto Oswaldo Cruz, Fiocruz. Av. Brasil 4365, Manguinhos 21040-900 Rio de Janeiro, RJ Brazil

^cLaboratório de Bacteriologia, Centro de Referência Prof. Hélio Fraga, Escola Nacional de Saúde Pública Sergio Arouca, Fiocruz. Estr. de Curicica 2000, Jacarepaguá, 22780-194 Rio de Janeiro, RJ Brazil.

[&]Authors contributed equally to this work.

*Corresponding author. Mailing address: Laboratório de Genômica Funcional e Bioinformática, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz. 21040-900 Rio de Janeiro, RJ Brazil. Phone: (55-21) 3865-8159. Fax: (55-21) 3865-8200. Email: lmlima@ioc.fiocruz.br

Author Email addresses:

Teca Calcagno Galvão: tecagalvao1@gmail.com

Cristiane Rodrigues Lima: crika@ioc.fiocruz.br

Leonardo Henrique Ferreira Gomes: leonardoh@ioc.fiocruz.br

Talita Duarte Pagani: talitadp@ioc.fiocruz.br

Marcelo Alves Ferreira: malves@ioc.fiocruz.br

Antônio S. Gonçalves: ajsg@ioc.fiocruz.br

Paloma Rezende Correa: pacorrea@ioc.fiocruz.br

Wim Maurits Degrave: wdegrave@ioc.fiocruz.br

Leila Mendonça-Lima: lmlima@ioc.fiocruz.br

Abstract

The Brazilian anti-tuberculosis vaccine strain *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) BCG Moreau is unique in having a deletion of 7608 bp (RD16) that results in the truncation of *rv3405c*, encoding a putative TetR transcriptional regulator. We investigated the effect of this truncation on the expression of *rv3406*, lying 81 bp downstream in the opposite orientation. RT-PCR and western blot experiments show that *rv3406* mRNA and Rv3406 accumulate in BCG Moreau but not in BCG Pasteur (strain that bears an intact *rv3405c*), suggesting this to be a result of *rv3405c* truncation. Recombinant Rv3405c forms a complex with the *rv3405c-rv3406* intergenic region, which contains a characteristic transcription factor binding site, showing it to have DNA binding activity. Complementation of *M. bovis* BCG Moreau with an intact copy of *rv3405c* abolishes Rv3406 accumulation. These results show that Rv3405c is a DNA binding protein that acts as a transcriptional repressor of *rv3406*.

Keywords

Mycobacterium bovis BCG

TetR transcriptional regulator

rv3405c

Gain of function mutation

1. Introduction

Mycobacterium bovis BCG is the only available vaccine against tuberculosis (TB), a curable disease that killed 1.4 million people in 2011.¹ The difficulty in preventing and treating TB comes from limited vaccine efficacy in adolescents and adults and from patients abandoning antibiotic use over the six month long course, as well as from public health services' limited reach. There are currently intensive efforts making use of functional genomics to generate and validate novel vaccines and antibiotics.

Bacille Calmette-Guérin (BCG), an attenuated *Mycobacterium bovis* strain, is the only available vaccine against TB. Its efficacy against extra-pulmonary forms of TB in children is greater than 80%,^{2,3} but protection varies from 0 to 80% in studies carried out with adolescents and adults.^{4,5} Even with these limitations, BCG is the most widely used human vaccine, being administered to more than 100 million children every year.⁶ In spite of its importance and an increasing interest in the molecular characterization of BCG, the causes of variable protection are only just starting to emerge. One important factor is the vaccine strain: the passaging of BCG and its distribution to different countries before the establishment of seed lots resulted in a family of daughter strains, each with a set of genomic differences in comparison with *M. bovis*.^{7,8} Genome mapping and sequencing has revealed that the rearrangements (insertions, deletions, duplications) and mutations in each strain yield a different set of genes, thus affecting immunogenicity and protective efficacy [reviewed in reference 9]. The complete genome sequences of BCG strains Pasteur,¹⁰ Tokyo,¹¹ Moreau,¹² Mexico¹³ and Korea¹⁴ have been reported, and draft sequences are available for BCG China, Denmark, Russia, and Tice.¹⁵ That the genomic differences generate strains with characteristic physiological attributes is clear from the striking differences observed in proteome profiles of different BCG strains. For example, the BCG Mexico and BCG Tokyo proteomes share only 48% of their proteins.¹³ The impact of genome particularities on virulence, or propagative capacity, is also of major importance. It has been shown that BCG secreted proteins, with their dominant immunogenicity and role in pathogenesis due to their early availability for interaction with host cells, accumulate differentially in BCG Moreau and BCG Pasteur.¹⁶ Indeed, older strains, such as BCG Moreau, are considered to be more immunogenic¹⁷ and might confer stronger protection against *M. tuberculosis* infection.¹⁸

The BCG strain used in Brazil since 1927 is known as BCG Moreau (reviewed in reference 19). One feature unique to its genome is RD16, a 7608 bp deletion in comparison to the *M. tuberculosis* genome that removes genes *rv3401-rv3404*, truncating and fusing the *rv3400* and *rv3405c* orthologues (Figure 1A). *rv3405c* encodes a putative TetR transcriptional regulator, and the BCG Moreau truncation results in a chimeric gene which maintains the *rv3405c* N-terminal coding sequence (including the putative DNA binding domain) fused to part of *rv3400*, encoded in the opposite strand (Figure 1B). Proper functioning of TetR transcriptional regulators requires dimerization and ligand binding,²⁰ mediated by the C-terminal region lost in the chimeric *rv3405c* gene of BCG Moreau. Comparison of the BCG Moreau and Pasteur cell surface proteomes (Pagani *et al.*, manuscript in preparation) revealed that Rv3406, the protein encoded by the gene adjacent and in the opposite orientation to *rv3405c*, is expressed only in BCG Moreau. The possible effects of *rv3405c* truncation on BCG Moreau physiology led us to examine the hypothesis that this gene encodes a transcriptional

repressor of *rv3406* transcription, and that this regulatory interaction is disrupted in the Brazilian vaccine strain.

2. Materials and methods

2.1. Bacterial strains and culture conditions

The strains and plasmids used in this study are described in Table 1. *E. coli* Top10 was used as a host for cloning, and *E. coli* BL21(DE3) as a host for expression of recombinant Rv3405c and Rv3406. Both *E. coli* strains were grown in Luria-Bertani (LB) medium (Difco) at 37 °C and kanamycin (50 ug ml⁻¹) was used for plasmid selection and maintenance. *M. bovis* BCG Pasteur 1173P2 was obtained from the Institut Pasteur (Paris, France) and maintained at -80 °C in 20% glycerol. *M. bovis* BCG Moreau was supplied in Sauton/potato medium by Fundação Ataulpho de Paiva (FAP, Rio de Janeiro, Brazil). Both strains were cultured as surface pellicles for 2 weeks at 37 °C in 100 ml Sauton medium,²¹ without agitation. To expand the cultures, about 100 µL of surface pellicles from each culture were seeded in fresh Sauton medium and incubated without shaking at 37 °C for 15 days. For cultures grown under agitation, Middlebrook 7H9 Broth (DIFCO) containing 0.05 % Tween 80 and 10 % ADC (5 % albumin, 0.003 % catalase, 2% dextrose) was used. *M. bovis* BCG Pasteur or *M. bovis* BCG Moreau -80°C stocks (1.0 mL at 1.0 OD₆₀₀) were diluted in 4.0 ml of media and grown with shaking at 37 °C. When cultures reached an OD₆₀₀ of 1.0, they were diluted 1:5 in fresh media. Recombinant BCG Moreau cultures were grown in the presence of 25 ug ml⁻¹ kanamycin. Preparation of electrocompetent BCG Moreau cells and electroporation conditions were essentially as described.²²

2.2. Protein Preparation

Culture filtrate proteins (CFPs) were obtained by removing supernatants from 15 day old BCG Moreau or Pasteur standing cultures, and centrifugation at 2,500 × g for 10 min at 4 °C. The supernatant was filtered through a 0.22 µm low protein binding membrane (Millipore Express; Millipore) in order to remove any remaining bacteria. Surface-associated proteins were obtained after vigorous manual shaking of the remaining bacterial pellicles (after removal of culture supernatant) with 2 mm glass beads, essentially as described.²³ The final bacterial sediment was used to obtain the intracellular protein fraction. Bacteria were resuspended in 1.5 mL of Lysis Buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM EDTA) including a cocktail of protease inhibitors (1:100, Sigma), transferred to 2.0 ml screw cap tubes containing 0.5 ml of 0.1 mm diameter zirconia/silica beads and lysed by three 1 min pulses in a mini bead beater apparatus (BioSpec Produces), keeping the tubes in ice for 1 min between pulses. The lysate was cleared by centrifugation at 16,000 × g for 10 min. Proteins in all fractions were precipitated with 17 % (v/v) trichloroacetic acid (TCA), washed with cold acetone and dissolved in IEF buffer (8 M urea, 2% CHAPS) for 1 h at room temperature. Protein concentrations were estimated with the RC DC protein assay kit (Bio-Rad) and preparations were kept at -80°C until use.

2.3. RNA extraction and RT-PCR

For RNA extraction from Sauton standing cultures, a spatula-full of bacterial surface pellicle from 15 day old BCG Moreau or Pasteur cultures was transferred to sterile, RNase-free microcentrifuge

tubes. Bacteria were suspended in 1.0 ml Trizol reagent (Life Technologies) by vortexing, transferred to 2.0 ml screw cap tubes containing 0.5 ml of 0.1 mm diameter zirconia/silica beads and lysed as described above. After centrifugation, the supernatant was transferred to a new tube (containing 0.5 volumes of chloroform). After mixing by inverting rapidly for 15 seconds samples were centrifuged for 15 min at 13,000 x g at 4 °C and the aqueous phase transferred to RNase-free tubes. The material was then extracted with 0.5 volumes of chloroform/isoamyl alcohol (24:1) and centrifuged as above. Nucleic acids were precipitated after addition of 0.7 volumes of isopropanol, incubation at 4 °C (10 min to overnight) and centrifugation for 15 min at 13,000 x g at 4 °C. The RNA pellets were washed with 500 µl 75% ethanol on ice, centrifuged for 15 min, air-dried for 20 min and resuspended in 50 µl of RNase-free water containing 0.5 µl RNasin (Promega). Any remaining contaminating DNA was removed by DNaseI treatment (Turbo DNase, Ambion), and RNA was quantified by measuring the absorbance at 260 and 280 nm. One microgram of RNA was used for cDNA synthesis, using SuperScript III (Life Technologies) following the manufacturer's instructions. The cDNA synthesis reactions included RNasin (Promega) and the specific reverse oligonucleotide for each RT-PCR target (Table 2). Then, PCR reactions were carried out with 2 µl of the specific cDNA, PCR Master Mix 2x (Promega) and forward and reverse oligonucleotides for each target (Table 2). RT-PCR results were analyzed by agarose gel electrophoresis in TBE 0.5 x. The negative controls for DNA contamination consisted of a PCR directly on RNA samples not reverse-transcribed with SuperScript III. The experiment was carried out in RNA obtained from 3 independent cultures of BCG Moreau and Pasteur.

2.4. SDS-PAGE and immunoblotting

Proteins were resolved by SDS-PAGE using 12% polyacrylamide gels²⁴ and transferred to nitrocellulose membrane (Hybond C, GE) using a semi-dry transfer equipment (BioRad), according to the manufacturer's specifications. The polyclonal Rv3406 and AphC antibodies were used at a 1:1000 dilution as primary antibodies, and anti-mouse IgG peroxidase conjugate (Thermo Scientific) as secondary antibody at a 1:10.000 dilution. All incubations with the primary and secondary antibodies were done in TBS-T (10 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.05 % Tween-20) with 5% skimmed milk, for 1.0 h at room temperature in a rocking platform, followed by 3 sequential 5 min washes with TBS-T and 3 more with TBS (10 mM Tris-Cl pH 7.5, 150 mM NaCl). Blots were developed using West Pico SuperSignal Chemiluminescent Substrate (Thermo Scientific). The final image was captured using a GS-800 scanning densitometer (BioRad).

2.5. Expression and purification of recombinant Rv3405c

For cloning, expression and purification of Rv3405c-6His, the *rv3405c* coding region was amplified from BCG Pasteur genomic DNA with primers For-*rv3405c*-pET and Rev-*rv3405c*-pET (Table 2) and the 586 bp PCR product was subcloned into pBad Topo TA (Life Technologies). DNA inserts were sequenced in an ABI 3730 DNA sequencer (Fiocruz/PDTIS sequencing platform), and the *rv3405c* *NcoI*-*XhoI* restriction fragment was ligated into *NcoI*-*XhoI*-digested pET28a vector (Table 1) using T4 DNA ligase, introducing a histidine tail at the protein's C terminus. Insert containing clones were identified and transformed into *E. coli* BL21(DE3). Cultures of bacteria harboring pET28a-05c were

grown in 500 ml LB at 37 °C to an OD₆₀₀ of 0.6 and recombinant protein expression induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) for 3 h at 37 °C. Cells were harvested, resuspended in 200 ml of Lysis Buffer and lysed with 0.1 mm diameter zirconia/silica beads (BioSpec Products) in a bead beater apparatus (3 pulses of 1 min). The cell lysate was centrifuged at 12,000 x g for 20 min at 4°C. Proteins in the soluble fraction were precipitated by slow addition of ammonium sulfate to 85 % (stirring on ice), and the sample was pelleted by centrifugation at 12,000 x g for 20 min at 4°C. The pellet was resuspended in 10 mM Tris-HCl pH 7.5, 10 mM imidazole, 200 mM NaCl, 0.05 % Tween-20 and loaded onto a 2 ml nickel chelating-sepharose manually packed column pre-equilibrated with the same buffer. After thorough washing bound proteins were eluted with an imidazole step gradient (50, 100, 150 and 200 mM). Rv3405c containing fractions were dialysed at 4°C against 2 liters of buffer containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 50 mM NaCl, 2.5% glycerol, 0.05% Tween-20. After 16 h glycerol was added to a final concentration of 16 % v/v, and aliquots were stored at -70°C.

2.6. *rv3405c* complementation

For complementation studies, *rv3405c* with its upstream region (61 bp) was amplified, from BCG Pasteur genomic DNA, with primers For-*rv3405c*-pUS and Rev-*rv3405c*-pUS using Platinum Supermix High Fidelity (Life Technologies). The amplification product was incubated with Master Mix (Promega) for introduction of A overhangs, and ligated with pGEM-T Easy (Promega) to generate pG05cKH. The ligation was electroporated into *E. coli* top10 cells, and clones containing inserts were sequenced. pG05cKH and pUS977 DNAs were digested with *KpnI* and *HindIII*, and the digestions were used in a ligation reaction with T4 DNA ligase (NEB), to generate pU05c. Clones containing insert were identified by PCR and restriction analysis, and were transformed into electrocompetent BCG Moreau.

2.7. Electrophoretic mobility shift assays (EMSAs)

To generate a probe for EMSA, a 101 bp fragment spanning the 63 bp *rv3405c*-*rv3406* intergenic region was amplified with primers For-*05c-06* (biotinylated at the 5' end) and Rev-*05c-06* (Table 2). The PCR product was purified with the Wizard SV Gel and PCR Clean-Up System kit (Promega), and 2 ng were used in binding reactions with increasing concentration of recombinant, purified Rv3405c. Binding reactions in a final volume of 20 μL contained 2 μL of 10x binding buffer, 1 μL of 50 % glycerol and 1 μL of poly dIdC at 1 $\mu\text{g } \mu\text{L}^{-1}$, from the LightShift™ EMSA Optimization & Control Kit (Pierce). Protein was added last, and reactions were incubated for 30 min at room temperature. Loading buffer containing glycerol was added to the reactions, which were then loaded onto 5 % polyacrylamide gels in 0.5 x TBE buffer. Gels were run at 100 V for 30 min. After electrophoresis, the DNA-protein complexes were transferred to positively charged Hybond-N+ nylon membranes (GE) for 1 hour at 180 milliamps in 0.5 x TBE buffer using a Mini-Transblot equipment (BioRad). DNA was crosslinked by exposure to UV light for 5 min for each side of the membrane. The biotinylated PCR product was detected with the Lightshift chemiluminescent EMSA Kit (Thermo) as described by the manufacturer. Membranes were exposed to X-ray film and developed.

3. Results

3.1. Differential *rv3406* transcription in BCG Moreau and BCG Pasteur

To characterize a possible role for *rv3405c* on *rv3406* transcription, RT-PCR experiments were carried out to assess *rv3406* mRNA levels in BCG Pasteur, bearing an intact *rv3405c*, and BCG Moreau. RNA from standing Sauton media cultures was extracted and a *rv3406* reverse primer (Rev-*rv3406*) was used for cDNA synthesis. PCR with *rv3406* primers (Table 2) on a no reverse transcriptase sample and on BCG Moreau genomic DNA showed the RNA to be DNA-free and the PCR to work, respectively. RT-PCR on the cDNAs shows *rv3406* to be expressed in BCG Moreau but not in detectable levels in BCG Pasteur (Figura 2A). To check whether the mRNA results were reproduced at the protein level, we expressed and purified recombinant Rv3406, and raised polyclonal antibodies in mice. The sera was used in western blots against BCG Moreau and BCG Pasteur cellular fractions (Figure 2B), and showed that Rv3406 is only detected in BCG Moreau. Furthermore, Rv3406 accumulates intracellularly, but most of the protein is found in the surface fraction. Thus, Rv3406 accumulation is differential between these vaccine strains, and this feature is regulated at the transcriptional level, with derepression of *rv3406* transcription in BCG Moreau.

The deletion in RD16 creates a novel genomic arrangement, whereby the *rv3400* and the *rv3405c* promoter regions direct transcription of sequences that were previously distant and presumably isolated by transcriptional terminators. In BCG Moreau the chimeric region resulting from the fusion of *rv3400-rv3405c* could be transcribed from both DNA strands (Figure 3A). To test this, transcription from the two possible promoters was measured by RT-PCR. cDNAs were synthesized either with a *rv3400* or a *rv3405c* specific reverse primer (Rev-*rv3400-05c* or For-*rv3400-05c*, respectively), and separately used in PCR reactions with both these primers. As shown in Figure 3B, the 185 bp fragment comprising the *rv3400-rv3405c* region is amplified from both cDNAs, showing both promoters to be active. This result opens the possibility that *rv3406* is also transcribed as a *rv3400-rv3406* transcript, from the *rv3400* promoter in BCG Moreau. To determine if this is the case, cDNA was synthesized with a reverse primer that anneals to the *rv3406* sequence (Rev-*rv3406*) and used in a PCR with *rv3400* forward and *rv3406* reverse primers (For-*rv3400-05c* and Rev-*rv3400-05c*, respectively). In this manner, any PCR product has to originate from the cDNA resulting from activity of the *rv3400* promoter. However, no PCR product is observed (Figura 3C, lane 5), suggesting that transcription from the *rv3400* promoter terminates before *rv3406*. As a control, *rv3406* transcription from its own promoter was detected by a PCR on the cDNA made with *rv3406* primers (Figura 3C, lane 4).

3.3. *rv3405c* is transcribed in BCG Moreau and Pasteur

The above RT-PCR experiment suggests that Rv3405c represses *rv3406* transcription in BCG Pasteur and its truncation in BCG Moreau results in inactivation and *rv3406* constitutive expression. As TetR type regulators sometimes autoregulate their expression negatively or positively,^{25,26} a similar transcriptional profile might be observed for *rv3405c* in these strains. To check this possibility, RT-PCR experiments were carried out to look at *rv3405c* mRNA levels in BCG Pasteur and BCG Moreau. RNA from standing Sauton media cultures was extracted and the *rv3405c* reverse primers used were specific for the different 3' terminal sequences of the BCG Moreau (Rev-M-*rv3405c*) and Pasteur (Rev-P-*rv3405c*) genes. PCR on the cDNA samples revealed that both BCG Moreau and BCG Pasteur transcribe the regulator gene (Figure 4, lanes 3 and 6), suggesting that *rv3405c* transcription occurs independently of an intact *rv3405c* gene.

3.4. DNA binding by Rv3405c

Our results and genome annotation together suggest that Rv3405c is a TetR-type transcriptional repressor of *rv3406*. As such, DNA binding motifs for this type of regulator should be present in the upstream region and these should be conserved in mycobacteria presenting the *rv3405c-rv3406* region. Blast analysis using *rv3405c* and/or *rv3406*, or the intergenic region alone, reveals the presence of *rv3405c-rv3406* in several mycobacterial sequenced genomes and in *Segniliparus rugosus*, a member of the sub-order Corynebactericiae, as are the Mycobacteriaceae. Analysis of the 61 bp intergenic region reveals the presence of inverted repeats characteristic of TetR binding motifs (Figure 5).

To first analyse how Rv3405c regulates *rv3406* expression, we purified recombinant Rv3405c and used it in gel mobility shift assays with a PCR product encompassing the intergenic *rv3405c-rv3406* region (Figure 6). Rv3405c binds the DNA with nanomolar affinity, shifting the probe in concentrations as low as 0.4 nM (lane 4). At 40 nM all DNA is complexed with the protein (lane 5). The lower (0.4 and 4 nM Rv3405c; lanes 4 and 5), and higher (lanes 5 to 8) DNA-protein complex bands possibly represent DNA bound by one dimer or two dimers, respectively.

3.5. Complementation of *rv3405c* in BCG Moreau abolishes Rv3405c accumulation

To establish a firm link between the presence of intact Rv3405c and *rv3406* transcriptional repression, we cloned the BCG Pasteur gene into vector pUS977, creating pU05c, and introduced it into BCG Moreau. In these experiments, recombinant cultures were grown in 7H9 medium until an OD₆₀₀ of 1.0, when intracellular proteins were extracted and used for western blotting. As shown in Figure 7A, Rv3406 accumulation is observed in BCG Moreau bearing the empty vector, but abolished in the strain bearing pU05c. As a control, the same amount of the lysate proteins were probed with an anti-AhpC antibody, showing bands of similar intensity. The identity of the strains was confirmed by PCR (Figure 7B). Thus, Rv3405c encoded by the pU05c construct represses *rv3406* expression.

4. Discussion

We show by RT-PCR, gel shift assays, and complementation studies that *rv3405c* encodes a transcriptional repressor of *rv3406*, and that the unique deletion in the BCG Moreau genome that truncates *rv3405c* reshapes transcriptional activity in the region, including rendering *rv3406* constitutively transcribed. The genomic rearrangement involving RD16 in BCG Moreau brings *rv3400* closer to *rv3406* and albeit *rv3400* promoter activity was observed, no *rv3400-rv3406* transcript was detected, suggesting the presence of a terminator. In BCG Pasteur there may be a terminator upstream of *rv3405c*, but as the region does not exist in BCG Moreau by virtue of the deletion, transcription is possibly being terminated by sequences encoded in the inverted repeats identified in the intergenic *rv3405c-rv3406* region, forming hairpins in the mRNA. RT-PCR experiments revealed *rv3405c* transcription in BCG Moreau and Pasteur, suggesting that the regulator does not function as an activator of its own transcription as has been described for some TetR type regulators, such as Pip²⁶ or MdoR.²⁷ We postulate that other proteins may regulate *rv3405c* transcription by binding these sites, and this is not affected by Rv3405c inactivation. Although the effects of *rv3406* constitutive expression are unknown, evidence from microarray experiments suggests that Rv3406 may have a role in the physiology of *M. tuberculosis* in the host. *rv3406* mRNA was detected in sputum samples,²⁸ and *rv3405c* transcription is induced by treatment of *M. tuberculosis* with bovine lung surfactant.²⁹ Also, treatment of *M. tuberculosis* cultures with low concentrations of SDS induced *rv3406*.³⁰

BCG Tokyo has a 22 bp deletion in *rv3405c*³¹ that results in a frameshifted protein product that retains the first 74 residues of the 189 amino acid full-length protein, plus a novel sequence of 19 amino acids resulting from this frameshift. Although the TetR-type helix-turn-helix DNA binding domain is maintained, it is presumed that Rv3405c function is inactivated, and the deletion has been correlated with changes in colony morphology and growth advantage in a BCG-mouse infection model.³² Involvement of *rv3405c* in important cellular processes has also been shown by Chen and colleagues, who demonstrated that expression of *rv3405c* in BCG Tokyo partially restores accumulation of cell wall phenolic glycolipids (PGL).³³ However, this result is controversial as there are experiments showing that *rv3405c* is not involved in the regulation PGL synthesis.³⁴

As happened with *rv3405c* in BCG Moreau, the function of the transcriptional repressor encoded by *rv1773* in BCG Pasteur is inactivated by truncation, due to a Pasteur-specific genomic deletion, RD14.³⁵ The functions of *rv1774* and *rv1775* have not been described, but their upregulation in BCG Pasteur was suggested to result in a gain of function that facilitates nutrient utilization.³⁵

Our results gain pertinence in the context of mutations that resulted in the attenuation of *M. bovis* BCG. Between *M. bovis* and *M. bovis* BCG, mutations occurred under selective pressure for the bacteria to propagate in a defined growth media, but it was the resulting strain's inability to cause tuberculosis that led to the vaccine strain selection. Thus, mutations in the genomes of BCG strains may have been selected, at least partially, for conferring improved growth in the culture conditions. Indeed, the duplication of the gene encoding glycerol-3-phosphate dehydrogenase, *glpD2*, in BCG strains,¹⁰ is likely to facilitate the metabolism of glycerol present in the media used during attenuation.³⁶ However, it is only through specific studies that the impact of each mutation on vaccine properties may be elucidated. An interesting example of this is the *rv3676* gene encoding the cAMP-receptor protein (CRP), a global transcriptional regulator polymorphic in all BCG strains in which the gene has been sequenced.³⁷ BCG Pasteur CRP is functional, but its DNA binding properties and regulon differ from that of the *M. tuberculosis* protein.³⁸ Furthermore, substitution of the *M. tuberculosis*

crp gene for the BCG polymorphic copy alters the transcriptome, though it does not result in attenuation.³⁹ However, although the effect of CRP mutation is not critical for attenuation, its role in the immunomodulatory properties of BCG has not been investigated. Other examples of transcriptional regulators whose mutation in BCG may impact vaccine properties are *whiB3* and *trcR*, which suffer partial deletions in BCG strains Sweden and Birkhaug.⁴⁰ The inactivation of *whiB3* in *M. bovis* and of *trcR* in *M. tuberculosis* leads to attenuation and hypervirulence, respectively.^{41,42}

rv3406 encodes a putative taurine deoxygenase, and the protein's accumulation in BCG Moreau due to *rv3405c* truncation could confer selective advantages in culture conditions employed for strain maintenance prior to the establishment of seed lots. Though the precise composition of the media used in the early years of BCG passaging in Brazil is unknown, medium containing ox bile was used by Calmette and Guerin.⁴³ Bile acids are major components of bile, synthesized from cholesterol, and are either glyco- or tauro-conjugated (reviewed in 44); they are similar to cholesterol esters found in human lung surfactant⁴⁵ and caseous granulomas.⁴⁶ Therefore, we speculate that since sub-culturing was done in the presence of bile, constitutive expression of *rv3406* resulting from loss of its transcriptional repressor may have conferred a metabolic advantage, leading to selection of the strain deleted in RD16. The impact of such *in vitro* adaptation on BCG Moreau's immune-modulatory properties remain to be explored.

Acknowledgements

DNA sequencing was carried out by Plataforma Genômica para Sequenciamento de DNA - PDTIS/FIOCRUZ (RPT01A). This study was supported by PDTIS/FIOCRUZ, Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, and the WHO/TDR Special Programme for Research and Training in Tropical Diseases.

References

1. WHO Global Tuberculosis Report 2012
2. Colditz GA, Berkey CS, Mosteller F, Brewer TF, Wilson ME, Burdick E, Fineberg HV. The efficacy of bacillus Calmette-Guérin vaccination of newborns and infants in the prevention of tuberculosis: meta-analyses of the published literature. *Pediatrics* 1995;**96**:29-35.
3. Trunz BB, Fine P, Dye C. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet* 2006;**367**:1173-80.
4. Brewer TF. Preventing tuberculosis with bacillus Calmette-Guérin vaccine: a meta-analysis of the literature. *Clin Infect Dis* 2000;**31**(Suppl. 3):64-7.
5. Colditz GA, Brewer TF, Berkey CS, Wilson ME, Burdick E, Fineberg HV, Mosteller F. Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *JAMA* 1994;**271**:698-702.
6. WHO Position Paper on BCG. *Wkly Epidemiol Rec.* 2004;**79**:27-38
7. Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 1999;**284**:1520-3.
8. Mostowy S, Cleto C, Sherman DR, Behr MA. The *Mycobacterium tuberculosis* complex transcriptome of attenuation. *Tuberculosis* 2004;**84**:197-204.
9. Liu J, Tran V, Leung AS, Alexander DC, Zhu B. BCG vaccines: their mechanisms of attenuation and impact on safety and protective efficacy. *Hum Vaccin* 2009;**5**:70-8.
10. Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, Dos Santos S, Duthoy S, Lacroix C, Garcia-Pelayo C, Inwald JK, Golby P, Garcia JN, Hewinson RG, Behr MA, Quail MA, Churcher C, Barrell BG, Parkhill J, Cole ST. Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A* 2007;**104**:5596-601.
11. Seki M, Honda I, Fujita I, Yano I, Yamamoto S, Koyama A. Whole genome sequence analysis of *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) Tokyo 172: a comparative study of BCG vaccine substrains. *Vaccine* 2009;**27**:1710-6.
12. Gomes LH, Otto TD, Vasconcellos EA, Ferrão PM, Maia RM, Moreira AS, Ferreira MA, Castello-Branco LR, Degraive WM, Mendonça-Lima L. Genome sequence of *Mycobacterium bovis* BCG Moreau, the Brazilian vaccine strain against tuberculosis. *J Bacteriol* 2011;**193**:5600-1.
13. Orduña P, Cevallos MA, de León SP, Arvizu A, Hernández-González IL, Mendoza-Hernández G, López-Vidal Y. Genomic and proteomic analyses of *Mycobacterium bovis* BCG Mexico 1931 reveal a diverse immunogenic repertoire against tuberculosis infection. *BMC Genomics* 2011;**12**:493.

14. Joung SM, Jeon SJ, Lim YJ, Lim J-S, Choi B-S, Choi I-Y, Yu JH, Na K-I, Cho E-H, Shin S-S, Park YK, Kim C-K, Kim H-J, Ryoo SW. Complete genome sequence of *Mycobacterium bovis* BCG Korea, the Korean vaccine strain for substantial production. *Genome Announc* 2013;**1**:e00069-13.
15. Pan Y, Yang X, Duan J, Lu N, Leung AS, Tran V, Hu Y, Wu N, Liu D, Wang Z, Yu X, Chen C, Zhang Y, Wan K, Liu J, Zhu B. Whole-genome sequences of *four* *Mycobacterium bovis* BCG vaccine strains. *J Bacteriol* 2011;**193**:3152-3.
16. Berrêdo-Pinho M, Kalume DE, Correa PR, Gomes LH, Pereira MP, da Silva RF, Castello-Branco LR, Degraive WM, Mendonça-Lima L. Proteomic profile of culture filtrate from the Brazilian vaccine strain *Mycobacterium bovis* BCG Moreau compared to *M. bovis* BCG Pasteur. *BMC Microbiol* 2011;**11**:80.
17. Kozak R, Behr MA. Divergence of immunologic and protective responses of different BCG strains in a murine model. *Vaccine* 2011;**29**:1519-26.
18. Hayashi D, Takii T, Fujiwara N, Fujita Y, Yano I, Yamamoto S, Kondo M, Yasuda E, Inagaki E, Kanai K, Fujiwara A, Kawarazaki A, Chiba T, Onozaki K. Comparable studies of immunostimulating activities in vitro among *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) substrains. *FEMS Immunol Med Microbiol* 2009;**56**:116-28.
19. Benévolo-de-Andrade TC, Monteiro-Maia R, Cosgrove C, Castello-Branco LR. BCG Moreau Rio de Janeiro: an oral vaccine against tuberculosis--review. *Mem Inst Oswaldo Cruz* 2005;**100**:459-65.
20. Yu Z, Reichheld SE, Savchenko A, Parkinson J, Davidson AR. A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators. *J Mol Biol* 2010;**400**:847-64.
21. Sauton, B. Sur la nutrition minérale du bacille tuberculeux. *C R Hebd Seances Acad Sci* 1912;**155**: 860-861.
22. Goude R, Parish T. Electroporation of Mycobacteria. In: Parish T, Brown AC, editors. *Mycobacteria Protocols: Second Edition*, New York: Humana Press; 2008, p. 203-216.
23. Ortalo-Magné A, Dupont MA, Lemassu A, Andersen AB, Gounon P, Daffe M. Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiology* 1995;**141**:1609-20.
24. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;**227**:680-5.
25. Bertram R, Hillen W. The application of Tet repressor in prokaryotic gene regulation and expression. *Microb Biotechnol* 2008;**1**:2-16.
26. Folcher, M., et al. A transcriptional regulator of a pristinamycin resistance gene in *Streptomyces coelicolor*. *J Biol Chem* 2001;**276**:1479-85.
27. Park H, Ro YT, Kim YM. MdoR is a novel positive transcriptional regulator for the oxidation of methanol in *Mycobacterium* sp. strain JC1. *J Bacteriol* 2011;**193**:6288-94.

28. Garton NJ, Waddell SJ, Sherratt AL, Lee SM, Smith RJ, Senner C, Hinds J, Rajakumar K, Adegbola RA, Besra GS, Butcher PD, Barer MR. Cytological and transcript analyses reveal fat and lazy persistor-like bacilli in tuberculous sputum. *PLoS Med* 2008;**5**:e75.
29. Schwab U, Rohde KH, Wang Z, Chess PR, Notter RH, Russell DG. Transcriptional responses of *Mycobacterium tuberculosis* to lung surfactant. *Microb Pathog* 2009;**46**:185-93.
30. Manganelli R, Voskuil MI, Schoolnik GK, Smith I. The *Mycobacterium tuberculosis* ECF sigma factor *sigmaE*: role in global gene expression and survival in macrophages. *Mol Microbiol* 2001;**41**:423-37.
31. Bedwell J, Kairo SK, Behr MA, Bygraves JA. Identification of substrains of BCG vaccine using multiplex PCR. *Vaccine* 2001;**19**:2146-51.
32. Honda I, Seki M, Ikeda N, Yamamoto S, Yano I, Koyama A, Toida I. Identification of two subpopulations of Bacillus Calmette-Guérin (BCG) Tokyo172 substrain with different RD16 regions. *Vaccine* 2006;**24**:4969-74.
33. Chen JM, Islam ST, Ren H, Liu J. Differential productions of lipid virulence factors among BCG vaccine strains and implications on BCG safety. *Vaccine* 2007;**25**:8114-22.
34. Naka T, Maeda S, Niki M, Ohara N, Yamamoto S, Yano I, Maeyama J, Ogura H, Kobayashi K, Fujiwara N. Lipid phenotype of two distinct subpopulations of *Mycobacterium bovis* Bacillus Calmette-Guerin Tokyo 172 substrain. *J Biol Chem* 2011;**286**:44153-61.
35. Alexander DC, Behr MA. Rv1773 is a transcriptional repressor deleted from BCG-Pasteur. *Tuberculosis* 2007;**87**:421-5.
36. Calmette A, Guérin C, Boquet A, Nègre L. La vaccination préventive contre la tuberculose par le "BCG". Paris: Masson et cie; 1927.
37. Spreadbury CL, Pallen MJ, Overton T, Behr MA, Mostowy S, Spiro S, Busby SJ, Cole JA. Point mutations in the DNA- and cNMP-binding domains of the homologue of the cAMP receptor protein (CRP) in *Mycobacterium bovis* BCG: implications for the inactivation of a global regulator and strain attenuation. *Microbiology* 2005;**151**:547-56.
38. Bai G, Gazdik MA, Schaak DD, McDonough KA. The *Mycobacterium bovis* BCG cyclic AMP receptor-like protein is a functional DNA binding protein *in vitro* and *in vivo*, but its activity differs from that of its *M. tuberculosis* ortholog, Rv3676. *Infect Immun* 2007;**75**:5509-17.
39. Hunt DM, Saldanha JW, Brennan JF, Benjamin P, Strom M, Cole JA, Spreadbury CL, Buxton RS. Single nucleotide polymorphisms that cause structural changes in the cyclic AMP receptor protein transcriptional regulator of the tuberculosis vaccine strain *Mycobacterium bovis* BCG alter global gene expression without attenuating growth. *Infect Immun* 2008;**76**:2227-34.

40. Leung AS, Tran V, Wu Z, Yu X, Alexander DC, Gao GF, Zhu B, Liu J. Novel genome polymorphisms in BCG vaccine strains and impact on efficacy. *BMC Genomics* 2008;**9**:413.
41. Steyn AJ, Collins DM, Hondalus MK, Jacobs WR Jr, Kawakami RP, Bloom BR. *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. *Proc Natl Acad Sci U S A* 2002;**99**:3147-52.
42. Parish T, Smith DA, Kendall S, Casali N, Bancroft GJ, Stoker NG. Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect Immun* 2003;**71**:1134-40.
43. Calmette A. Preventive Vaccination Against Tuberculosis with BCG. *Proc R Soc Med* 1931;**24**:1481-90.
44. Begley M, Gahan CG, Hill C. The interaction between bacteria and bile. *FEMS Microbiol Rev* 2005;**29**:625-51.
45. Hallman M, Gluck L. Phosphatidylglycerol in lung surfactant. III. Possible modifier of surfactant function. *J Lipid Res* 1976;**17**:257-62.
46. Kim MJ, Wainwright HC, Locketz M, Bekker LG, Walther GB, Dittrich C, Visser A, Wang W, Hsu FF, Wiehart U, Tsenova L, Kaplan G, Russell DG. Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism. *EMBO Mol Med* 2010;**2**:258-74.
47. Medeiros MA, Dellagostin OA, Armôa GR, Degraive WM, de Mendonça-Lima L, Lopes MQ, Costa JF, McFadden J, McIntosh D. Comparative evaluation of *Mycobacterium vaccae* as a surrogate cloning host for use in the study of mycobacterial genetics. *Microbiology* 2002;**148**:1999-2009.

Table 1. List of strains and plasmids used in this study

	Genotype/Description	Source
<i>Escherichia coli</i> strains		
<i>E. coli</i> Top 10	<i>F mcrA Δ(mrr-hsdRMS-MCR) φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galK rpsL (Str^R)endA1 nupG</i>	Life Technologies
<i>E. coli</i> BL21(DE3)	<i>F- ompT hsdSB(rB- mB-) gal dcm</i> (DE3)	Novagen
<i>M. bovis</i> BCG strains		
BCG Moreau	<i>M. bovis</i> BCG strain Moreau	FAP
BCG Pasteur	<i>M. bovis</i> BCG strain Pasteur	Pasteur Institute
Plasmids		
pBADTopo TA (3931 bp)	Amp ^R ; Kan ^R ; pUC <i>ori</i> ; P _{Lac} ; <i>lacZα</i>	Life Technologies
pGEM-T Easy (3000 bp)	Amp ^R ; P _{Lac} ; <i>lacZ</i>	Promega
pET28a (5369 bp)	Kan ^R ; f1 <i>ori</i> ; pBR322 <i>ori</i> ; <i>Lac1</i> ; P _{T7} ; His tag; T7 tag	Novagen
pG05cKH (3660 bp)	pGEM-T Easy with a 645 bp PCR fragment of <i>rv3405c</i> with its 61 bp upstream region flanked by <i>KpnI</i> and <i>HindIII</i> sites.	This study
pET05 (5803 pb)	Kan ^R ; f1 <i>ori</i> ; pBR322 <i>ori</i> ; <i>Lac1</i> ; P _{T7} ; His tag; T7 tag; <i>rv3405</i> gene from <i>M. bovis</i> BCG Pasteur cloned in the <i>NcoI</i> and <i>XhoI</i> sites, introducing a C-terminal His tail.	This study
pUS977	<i>E. coli</i> -mycobacterial shuttle vector. pUC <i>ori</i> ; <i>oriM</i> ; P _{AN} ; Kan ^R .	Ref. 44
pU05c	pUS977 with a 628 bp fragment (<i>rv3405c</i> with its 61 bp upstream region) cloned in the <i>KpnI</i> and <i>HindIII</i> sites.	This study

Table 2. Primers used in this study

Primer name	Sequence (5' - 3')	Purpose
For- <i>rv3406</i>	ATCATCACGCCGATCAACTC	RT-PCR
Rev- <i>rv3406</i>	GCAGCACTTCGAATAACACC	
For- <i>rv3405c</i>	ACTACGCGTCCGGCAACC	RT-PCR
Rev-P- <i>rv3405c</i>	GTGCGGTTCCAGGATCCGGG	
Rev-M- <i>rv3405c</i>	TCCGGAAGTGAGGTCCCG	
F00-05c	ACTATCCGGAAGTGAGGTCCC	RT-PCR
R00-05c	CCTGGGCACGAAGCTGACCA	
For-05c-06	ACGGTAATCAGATCTGTCAT	EMSA PCR
Rev-05c-06	TTGCCGACGCGTAGTCATC	
For- <i>rv3405c</i> -pET	GAGACCATGGGAATGACTACGCGTCCGG	Cloning (pET28a)
Rev- <i>rv3405c</i> -pET	TTCCTCGAGGTGCGGTTCCAGGA	
For- <i>rv3405c</i> -pUS	GGGTACCTGCTGGCGTCCTCAATGAATG	Cloning (pUS977)
Rev- <i>rv3405c</i> -pUS	CCCAAGCTTTCAGTGCGGTTCCAGGATCCG	
Rev2-M- <i>rv3405c</i>	GCGTGCTGGGTCTG	PCR (recombinant BCG)
pUS-FOR	TACGAGCCACACAGCGCCCG	PCR detection of pUS977
pUS-REV	ACGTTTCCCGTTGAATATGGCT	

Figure legends:

Figure 1. The *rv3405c* region in *M. tuberculosis* and *M. bovis* BCG Moreau. (A) The RD16 region of the *M. tuberculosis* H37Rv genome comprises 7608 bp that are missing in *M. bovis* BCG Moreau. The deletion truncates the homologs of *rv3400* (checked black) and *rv3405c* (checked white), creating ORFs *BCG_M3438* and *BCG_M3439c* (magnified in inset B). (C) Putative TetR domains of *M. tuberculosis* *rv3405c* and its BCG Moreau homolog; the N-terminal 98aa containing the DNA-binding domain are maintained while the C-terminal dimerization domain is lost in BCG Moreau.

Figure 2. *rv3406* is expressed in *M. bovis* BCG Moreau, but not in *M. bovis* BCG Pasteur, at the RNA and protein levels. (A) RT-PCR on RNA from BCG Pasteur (lanes 1 to 3) and BCG Moreau (lanes 4 to 6) cultures. Lanes 1 and 4, *rv3406* PCR on Dnase I treated RNA; lanes 2 and 5, *rv3406* PCR on cDNA synthesised with *rv3406* reverse primer; lanes 3 and 6, *rv3406* PCR on genomic DNA. (B) Western blot with polyclonal anti-Rv3406 on intracellular (IC), surface (SF) and culture filtrate (CF) proteins of BCG Moreau (M) and Pasteur (P).

Figure 3. The *rv3400* promoter is active in BCG Moreau, but does not transcribe *rv3406*. (A) Schematic diagram of the BCG Moreau *rv3400-rv3405c* genomic region with the presumed promoter of each gene. The *F00-05c* and *R00-05c* primers allow cDNA synthesis of mRNA transcribed from the *rv3405c* and *rv3400* promoters, respectively. (B) Detection of transcripts from the *rv3400* and *rv3405c* promoters by RT-PCR. PCR with *F00-05c* and *R00-05c* on genomic DNA (1), 185 bp; PCR with *F00-05c* and *R00-05c* on DNaseI treated RNA (2); RT-PCR of *rv3400-rv3405c* on cDNA synthesised with *F00-05c* (3) or *R00-05c* (4). (C) Detection of *rv3400-rv3406* transcripts. PCR with genomic DNA for *rv3406* (1), 218 bp and *rv3400-rv3406* (2), 1128 bp; *rv3400-06* PCR on DNaseI treated RNA (3); RT-PCR for *rv3406* (4) and *rv3400-rv3406* (5).

Figure 4. *rv3405c* is transcribed in BCG Moreau and BCG Pasteur. RT-PCR on RNA from BCG Moreau (lanes 1 - 3) and BCG Pasteur (lanes 4 - 6) cultures. Lanes 1 and 4, *rv3405c* PCR on Dnase I treated RNA; lanes 2 and 5, *rv3405c* PCR on genomic DNA; lanes 3 and 6: *rv3405c* PCR on cDNA synthesised with *rv3405c* reverse primer specific for the 3' sequence of each strain.

Figure 5. The *rv3405c-rv3406* region is conserved, and found almost exclusively in mycobacterial genomes. The alignment, constructed with T-Coffee, includes intergenic regions found only in genome sequences with syntenic *rv3405c-rv3406* genes. A palindromic sequence motif is shown in bold and grey; the start codons of *rv3405c*, *rv3406* and their homologs are shown in red and blue, respectively; a putative Shine Delgarno for *rv3406* is underlined. **MOR:** *M. bovis* BCG Moreau (AM412059.2; 3779925-3779991), representing all species in the *Mycobacterium tuberculosis complex*; **SEG:** *Segniliparus rugosus* (ACZ101000111.1; 25640-25703); **MC2:** *M. smegmatis* MC²155 (NC_008596.1; 204171-204227); **ULC:** *M. ulcerans* (NC_008611.1; 947398-947333); **MAR:** *M. marinum* (NC_010612.1; 1370781-1370716); **KAN:** *M. kansasii* (ACBV01000001.1; 171738-171676); **INT:** *M. intracellulare* (NC_016946.1; 4448020-4448085); **COL:** *M. colombiense* (AFVW01000010.1 ; 191991-192056); **AVI:** *M. avium* (NC_008595.1; 4491798-4491857); **ABS:** *M. abscessus* (NC_010397.1; 82011-82088); **MAS:** *M. massiliense* (AHAR01000001.1; 33626-33705); **PAR:** *M. parascrofulaceum* (ADNV01000278.1; 1019-954); **IND:** *M. indicus pranii* (CP002275.1; 4607732-4607845); **LIF:** *M. liflandii* (CP003899.1; 1372556-1372434); **XEN:** *M. xenopi* (AJFI01000021.1; 109596-109693).

Figure 6. Rv3405c binds the *rv3405c-rv3406* intergenic region. Increasing amounts of recombinant Rv3405c was incubated with a 101 bp biotinylated PCR fragment spanning the 61 bp *rv3405c-rv3406* intergenic region, reactions were ran on native acrylamide:bisacrylamide gels, the DNA transferred to nitrocellulose membrane, and visualized with na HRP-streptavidin conjugate. 1: free probe; lanes 2 to 8: probe with Rv3405c at 4pM, 40 pM, 400 pM, 4 nM, 40 nM, 400 nM, 4 uM.

Figure 7. Complementation of BCG Moreau with *rv3405c* from BCG Pasteur abolishes Rv3406 accummulation. (A) Western blot with anti-AhpC (lanes 1 and 2) and anti-Rv3406 (lanes 3 and 4) sera on BCG Moreau carrying pUS977 (lanes 1 and 3) or pU05c (lanes 2 and 4). Equal amounts of intracellular proteins were loaded on each lane. (B) PCR on corresponding recombinant BCG Moreau culture lysates confirming the presence of pUS977 (lanes 1-3) or pU05c (lanes 4-6). Lanes 1 and 4, amplification of 448 bp fragment spanning the *rv3405c* promoter and part of its coding sequence found both in BCG Moreau and Pasteur; lanes 2 and 5, amplification of 640 bp fragment of the *rv3405c* promoter and coding sequence cloned into pU05c; lanes 3 and 6, amplification with pUS977 primers flanking the MCS yielding a 382 bp (empty vector) or 818 bp (pU05c) fragments.

FIG.1

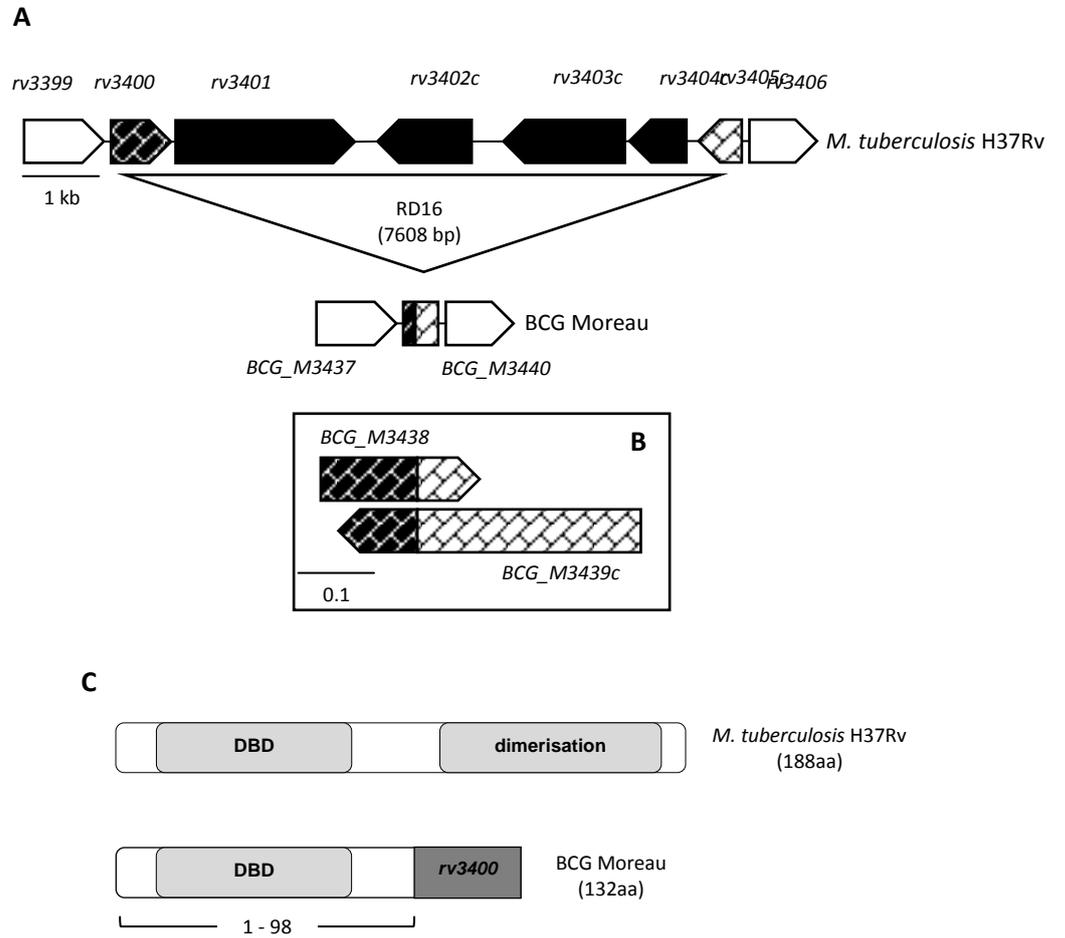


FIG.2

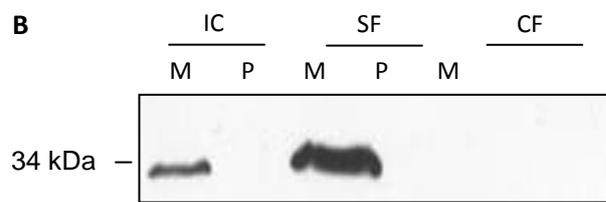
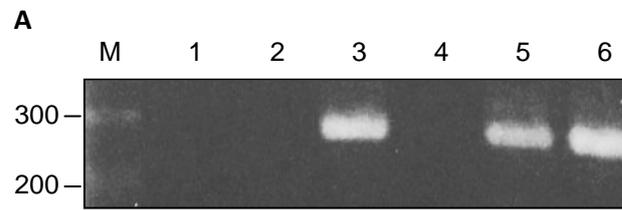


FIG.3

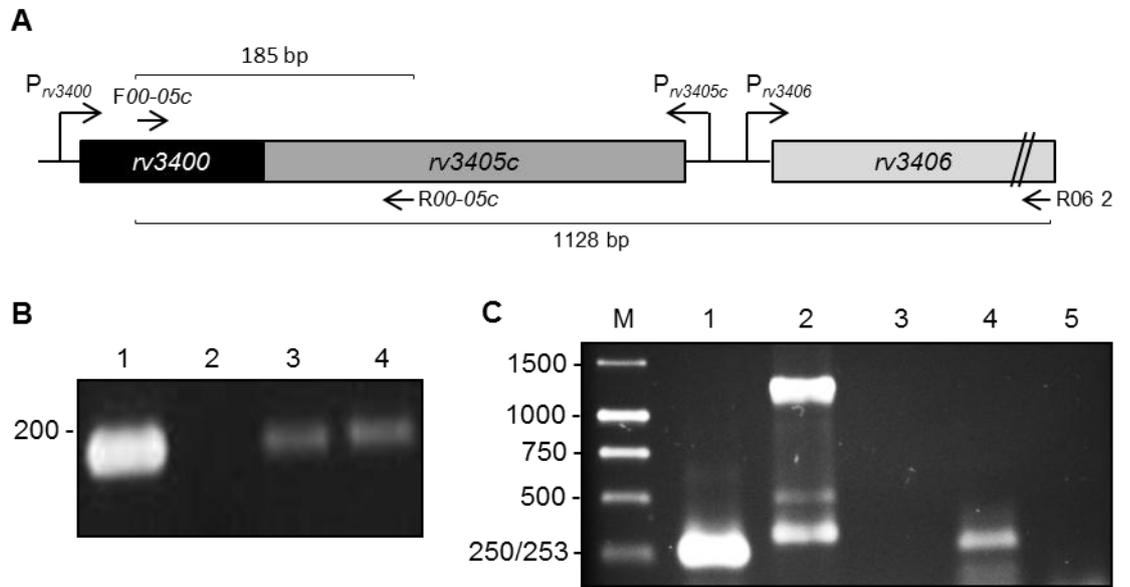


FIG.4

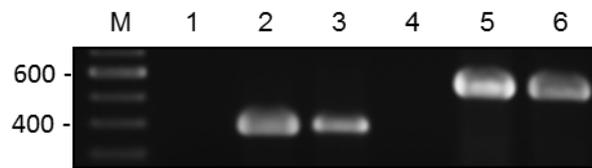


FIG.5

```

MOR -----CAICCGGCTACAGTAACCGA-----TGTAGTCACTGACTACACTAA-----CCATTCAITGAGGACGCCA--GCANTG 67
SEG -----CATGGCTACAATAGCGGCAAC--ATTTGTAGTCACTGACTACAAITG-----AAAGGAACG-GACC-ATG 64
MC2 -----CACTCGACTACATTAGCG-----TGTAGTCACTGACTACAAACG-----TGAGGAGCG-CGCAGATG 57
ULC -----CACAGACTACACTAGCCGA-----TGTAGTCACTGACTACACCCT-----ACGTTTGCTTAAGGAACA--GCANTG 66
MAR -----CACAGACTACACTAGCCGA-----TGTAGTCACTGACTACACCCT-----ACGTTTGCTTAAGGAACA--GCANTG 66
KAN -----CACTCGACTACGCTACCCGA-----TGTAGTCACTGACTACACC-----GTCT-CCGAAAGGACCAT--GCANTG 63
INT -----CAICTGACTACACTACAAA-----CGTAGTCACTGACTACACCG-----TTTTGACTCGAGGAAC-ACGAC-ATG 66
COL -----CAICTGACTACACTACAAA-----TGTAGTCACTGACTACACCG-----TTTTGAGCCAAGGA--ACGACANTG 66
AVI -----CAICTGACTACAGTAGCATT-----TGTAGTCACTGACTACACCGTTTCGGTCCCTCTTCAAGCCAAGGACGACGACANTG 80
PAR -----CAITACGACTACAGTACTCGA-----TGTAGTCACTGACTACACCG-----CTTGACACTAGGATTGATGACANTG 66
ABS -----CAITGTGTACAGTAGCGGATGACCAAGTGTAGTCACTGACTACACCGATTAT-----CGGACGGAAAGGAACG-GGTC-ATG 78
MAS -----CAITGTGTACAGTAGCGGATGACCAAGTGTAGTCACTGACTACACCGATTAT-----TCGGACGGAAAGGAACG-GGTC-ATG 80
IND -----CACTTCGCGCCGCGGTCGGGGGTTCCGCGCCCTCTTTCGGACGTGTAGTCACTGACTACACCAAA-----TTTTGACTCGAGGAAC-ACGAC-ATG 120
LIN CGCCGCGGCTACCTCTTCCCGTCCGCTGGGGTCTTTCGAGAACGGAITGCCGAACTGTAGTCACTGACTACACTAGCCGA-----TGTAGTCACTGACTACACCCT-----ACGTTTGCTTAAGGAACA--GCANTG 129
XEN -----CACCCGG-----CCTCGTTTACCTGTTGACACCGGTAGTCATATGACTACAGTAGCGTA-----TGTAGTCACTGACTACATACG-----CTACGC-CAAGGAAT-----CCGATG 98

```

** * ,***, ** , .*****. *****.. ***

FIG.6

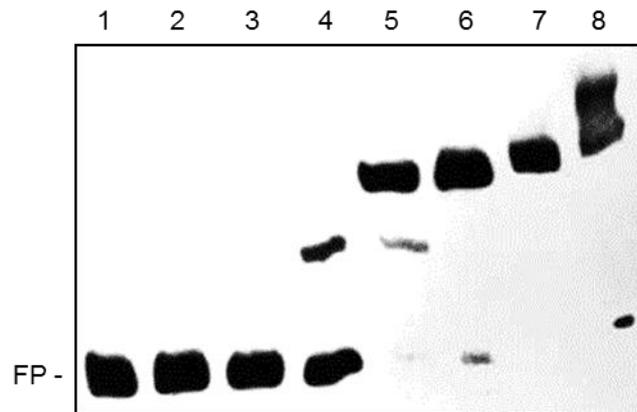


FIG.7

