

(11) Número de Publicação: **PT 770138 E**

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(51) Classificação Internacional:

C12N 15/74 (2007.10) **C12N 15/31** (2007.10)
C07K 14/35 (2007.10) **C12N 15/62** (2007.10)
A61K 39/04 (2007.10) **A61K 39/116** (2007.10)
C12Q 1/68 (2007.10) **C07K 16/12** (2007.10)
C12N 1/21 (2007.10)

(22) Data de pedido: **1995.08.30**

(30) Prioridade(s): **1994.09.02 FR 9410585**

(43) Data de publicação do pedido: **1997.05.02**

(45) Data e BPI da concessão: **2008.02.13**
084/2008

(73) Titular(es):

INSTITUT PASTEUR
25-28, RUE DU DOCTEUR ROUX 75724 PARIS
CÉDEX 15 **FR**

(72) Inventor(es):

BRIGITTE GICQUEL **FR**
ENG MONG LIM **FR**
DENIS PORTNOI **FR**
FRANÇOIS-XAVIER BERTHET **FR**
JULIANO TIMM **FR**

(74) Mandatário:

MARIA SILVINA VIEIRA PEREIRA FERREIRA
RUA CASTILHO, N.º 50, 5º - ANDAR 1269-163 LISBOA **PT**

(54) Epígrafe: **VECTORES DE PESQUISA E/OU DE EXPRESSÃO FUNCIONAIS EM MICOBACTÉRIAS**

(57) Resumo:

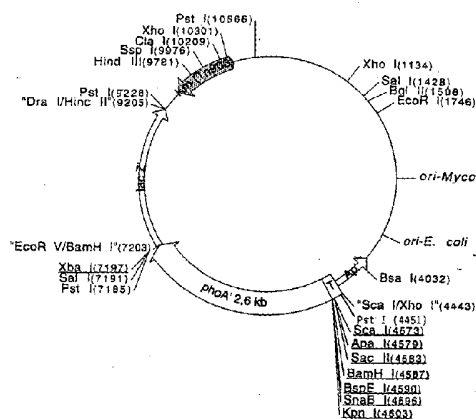
RESUMO

"VECTORES DE PESQUISA E/OU DE EXPRESSÃO FUNCIONAIS EM MICOBACTÉRIAS"

A invenção refere-se a um vector recombinante de pesquisa e/ou de clonagem e/ou de expressão, caracterizado por se replicar em micobactérias, por conter: 1) um replicação funcional nas micobactérias; 2) um marcador de selecção; 3) um bloco integrado relator que compreende a) um sítio de clonagem múltipla (*polylinker*), b) um terminador de transcrição activo nas micobactérias, situado a montante do *polylinker*, e c) uma sequência nucleotídica codificadora proveniente de um gene que codifica para um marcador de expressão e/ou de exportação e/ou de secreção de proteína, sequência nucleotídica esta que está desprovida do seu codão de início e das suas sequências de regulação. Este vector é utilizado para a identificação e para a expressão de polipeptídeos exportados, tais como o antigénio P28 de *Mycobacterium tuberculosis*.

pJEM11

(10 857 pb)



DESCRIÇÃO
"VECTORES DE PESQUISA E/OU DE EXPRESSÃO FUNCIONAIS EM
MICOBACTÉRIAS"

O género *Mycobacterium* inclui agentes patogénicos humanos importantes, tais como *M. leprae* e *M. tuberculosis*, os agentes responsáveis pela lepra e pela tuberculose, que continuam a ser problemas graves de saúde pública em todo o mundo.

M. bovis e *M. tuberculosis*, os agentes causais da tuberculose, são bactérias facultativas intracelulares. Apesar dos problemas importantes de saúde ligados a estes organismos patogénicos, sabe-se pouco sobre as suas proteínas exportadas e/ou segregadas. Análises em SDS-PAGE de um filtrado de cultura de *M. tuberculosis* mostram pelo menos 30 proteínas segregadas (1, 19, 38). Algumas destas foram caracterizadas, os seus genes clonados e sequenciados (7, 35, 37). Outras, se bem que se trate de antigénios imunodominantes de grande importância para induzir uma imunidade protectora (2, 21), não estão totalmente identificados. Além disso, é provável que numerosas proteínas exportadas fiquem fixadas sobre a membrana celular e, por consequência, não estejam presentes nos sobrenadantes de cultura. Foi demonstrado que as proteínas localizadas à superfície externa de diversas bactérias patogénicas, tais como a invasina de 103 kDa de *Yersinia pseudotuberculosis* (14) ou a internalina de 80 kDa de *Listeria monocytogenes* (10), desempenham um papel importante nas interacções com as células hospedeiras e, por consequência, na patogenicidade como na indução de respostas protectoras. Deste modo, uma proteína ligada à

membrana poderia ser importante para a infecção por M. tuberculosis como para a indução de uma resposta protectora contra esta infecção. Estas proteínas poderão revestir um interesse certo para a preparação de vacinas.

O BCG (Bacilo de Calmette e Guérin), uma estirpe avirulenta derivada de M. bovis, tem sido muito utilizada como vacina contra a tuberculose. É igualmente um vector muito interessante para a construção de vacinas recombinantes vivas, em particular por causa do seu forte poder imunogénico. Por consequência, o estudo da biologia molecular das micobactérias suscita actualmente grande interesse.

O desenvolvimento de novas vacinas contra as micobactérias patogénicas ou o melhoramento das vacinas disponíveis necessitava da disponibilidade de ferramentas específicas que permitissem isolar ou obter sequências de polipeptídeos imunogénicos.

Os inventores definiram e realizaram para este objectivo novos vectores que permitem a pesquisa de sequências de ADN de micobactérias a fim de identificar de entre estas sequências ácidos nucleicos que codificam para proteínas de interesse.

Foram definidos vectores para avaliar a eficácia de sequências de regulação da expressão no seio de micobactérias.

A invenção tem também como objectivo novos polipeptídeos de micobactérias que foi possível isolar por meio dos vectores precedentes e susceptíveis de entrar na realização de

composições para a detecção de uma infecção por micobactérias ou para a protecção contra uma infecção devida a micobactérias.

O pedido de patente descreve um vector recombinante de pesquisa e/ou de clonagem e/ou de expressão caracterizado por se replicar em micobactérias, por conter

- 1) um replicação funcional nas micobactérias;
- 2) um marcador de selecção;
- 3) um bloco integrado relator que compreende
 - a) um sítio de clonagem múltipla (*polylinker*),
 - b) um terminador de transcrição activo nas micobactérias a montante do *polylinker* e
 - c) uma sequência nucleotídica codificadora proveniente de um gene que codifica para um marcador de expressão e/ou de exportação e/ou de secreção de proteína, sequência nucleotídica esta que está desprovida do seu códon de início e das suas sequências de regulação.

O marcador de exportação e/ou de secreção é uma sequência de nucleótidos cuja expressão seguida da exportação e/ou da secreção depende dos elementos de regulação que controlam a sua expressão.

Por "sequências ou elementos de regulação da expressão" entende-se uma sequência promotora da transcrição, uma sequência que compreende o sítio de ligação ao ribossoma (RBS), as sequências responsáveis pela exportação e/ou pela secreção tais como a sequência sítio sequência de sinal.

Um primeiro marcador interessante de exportação e/ou de expressão é uma sequência codificadora proveniente do gene *phoA*. Eventualmente, esta sequência está truncada de tal

modo que a actividade de fosfatase é no entanto susceptível de ser restaurada quando a sequência codificadora truncada é colocada sob o controlo de um promotor e de elementos de regulação apropriados.

Podem ser utilizados outros marcadores de exposição e/ou de exportação e/ou de secreção. Serão mencionadas a título exemplificativo uma sequência do gene da β -agarase ou da nuclease de um estafilococo ou da β -lactamase de uma micobactéria.

O terminador de transcrição deve ser funcional nas micobactérias. Um terminador vantajoso é, a este respeito, o terminador do colifago T4 (tT4). Podem ser isolados outros terminadores apropriados para a realização da invenção utilizando a técnica apresentada nos exemplos, por exemplo por meio do vector pJN3.

Um vector particular descrito é o plasmídeo pJEM11 depositado na CNCM (*Collection Nationale de Cultures de Microorganismes* de Paris - França) sob o n.º I-1375, em 3 de Novembro de 1993.

Para a selecção ou a identificação de sequências de ácidos nucleicos de micobactérias que codificam para produtos susceptíveis de serem incorporados nas composições imunogénicas ou antigénicas para a detecção de uma infecção por micobactérias, o vector descrito compreenderá, num dos sítios do *polylinker*, uma sequência de nucleótidos duma micobactéria na qual se pesquisa a presença de sequências reguladoras associadas à totalidade ou a parte de um gene de interesse que permite, quando o vector que possui estas sequências (vector recombinante) está integrado ou se

replica num hospedeiro celular de tipo micobactéria, obter a exposição ao nível da membrana ou da parede celular do hospedeiro e/ou a exportação e/ou a secreção do produto de expressão da referida sequência de nucleótidos.

A sequência em questão de micobactérias pode ser qualquer sequência de que se pretende detectar se ela contém elementos de regulação da expressão associados à totalidade ou a parte de um gene de interesse e susceptíveis de permitir ou de favorecer a exposição ao nível da membrana celular de um hospedeiro no qual ela será expressa e/ou a exportação e/ou a secreção de um produto de expressão duma sequência codificadora dada e a título experimental do marcador de exportação e/ou de secreção.

De preferência, esta sequência é obtida por digestão enzimática do ADN genómico ou do ADN complementar de um ARN de uma micobactéria e de preferência de uma micobactéria patogénica.

De acordo com uma primeira forma de realização do vector descrito, a digestão enzimática do ADN genómico ou do ADN complementar é efectuada a partir de M. tuberculosis.

De preferência, esta ADN é digerido com uma enzima como, por exemplo, sau3A.

Podem naturalmente ser utilizadas outras enzimas de digestão, tais como ScaI, ApaI, ScaII, KpnI ou ainda exonucleases ou polimerases, uma vez que estas permitem a obtenção de fragmentos cujas extremidades podem ser inseridas num dos sítios de clonagem do *polylinker* do vector da invenção.

Eventualmente, serão efectuadas simultaneamente digestões com diferentes enzimas.

São escolhidos vectores recombinantes descritos de entre os vectores recombinantes seguintes depositados na CNCM em 8 de Agosto de 1994:

pExp53 depositado na CNCM sob o n.º I-1464

pExp59 depositado na CNCM sob o n.º I-1465

pExp410 depositado na CNCM sob o n.º I-1466

pExp421 depositado na CNCM sob o n.º I-1467

Os vectores da invenção podem igualmente ser utilizados para determinar a presença de sequências de interesse, de acordo com o que foi exposto anteriormente, em micobactérias tais como M. africanum, M. bovis, M. avium ou M. leprae cujo ADN ou ADNc tenha sido tratado com enzimas determinadas.

O pedido de patente descreve também um processo de pesquisa de sequências de nucleótidos provenientes de micobactérias para determinar a presença no seio destas sequências de elementos - de regulação que controlem a expressão num hospedeiro celular de sequências de ácido nucleico que as contém e/ou a exposição na superfície do hospedeiro celular e/ou a exportação e/ou a secreção de sequências polipeptídicas resultantes da exportação das ditas sequências de nucleótidos, caracterizado por compreender as fases seguintes:

a) a digestão das sequências de ADN de micobactérias por pelo menos uma enzima determinada e a recuperação dos fragmentos de digestão obtidos.

b) a inserção dos fragmentos de digestão num sítio de clonagem compatível com a enzima da fase a) do *polylinker* de um vector referido,

c) se necessário, a amplificação do fragmento de

digestão contido no vector, por exemplo, por replicação deste último após inserção do vector assim modificado numa célula determinada, por exemplo E. coli,

d) a transformação de hospedeiros celulares pelo vector amplificado na fase c) ou, na ausência de ampliação, pelo vector da fase b),

e) a cultura dos hospedeiros celulares transformados num meio que permite evidenciar o marcador de exportação e/ou de secreção contido no vector,

f) a detecção dos hospedeiros celulares positivos (colónias positivas) para a expressão do marcador de exposição e/ou de exportação e/ou de secreção,

g) o isolamento do ADN das colónias positivas deste ADN numa célula idêntica à da fase c),

h) a selecção das inserções contidas no vector, que permitem a obtenção de clones positivos para o marcador de exportação e/ou de secreção,

i) o isolamento e a caracterização dos fragmentos de ADN de micobactérias contidos nestas inserções.

A realização deste processo permite a construção de bancos de ADN que comportam sequências susceptíveis de ser exportadas e/ou segregadas, quando são produzidas no seio de micobactérias recombinantes.

A fase i) do processo pode compreender uma fase de sequenciação de inserções seleccionadas.

De preferência, o vector utilizado é o plasmídeo pJEM11 (CNCM I-1375) e a digestão é efectuada por meio da enzima sau3A.

O processo de pesquisa é, por exemplo, caracterizado por as sequências de micobactérias serem provenientes de uma micobactéria patogénica, por exemplo de M. tuberculosis, M. bovis, M. avium, M. africanum ou M. leprae.

O pedido de patente descreve também as sequências nucleotídicas de micobactérias seleccionadas depois da realização do processo acima descrito.

São sequências seleccionadas, por exemplo, os fragmentos de ADN de micobactérias contidos nos vectores pIPX412 (CNCM I-1463 depositado em 8/08/94), pExp53, pExp59, pExp410 ou pExp421.

Quando a sequência codificadora proveniente do gene marcador de exportação e/ou de secreção é uma sequência proveniente do gene *phoA*, a exportação e/ou a secreção do produto do gene *phoA*, eventualmente truncado, apenas é obtida quando esta sequência está inserida em fase com a sequência localizada a montante, que contém os elementos que controlam a expressão e/ou a exportação e/ou de secreção provenientes de sequência de micobactérias.

O pedido de patente descreve também micobactérias recombinantes que contêm um vector recombinante descrito anteriormente. Uma micobactéria preferida é uma micobactéria do tipo *M. smegmatis*.

A *M. smegmatis* permite de modo vantajoso experimentar a eficácia de sequências de micobactérias para o controlo da expressão e/ou da exportação e/ou da secreção de uma sequência dada, por exemplo de uma sequência que codifica para um marcador tal como a fosfatase alcalina.

Uma outra micobactéria interessante é uma micobactéria do tipo *M. bovis*, por exemplo a estirpe BCG utilizada actualmente para a vacinação contra a tuberculose.

O pedido de patente descreve, além disso, uma micobactéria recombinante, caracterizada por conter um vector recombinante acima definido.

A invenção tem por objecto uma sequência de nucleótidos proveniente de um gene que codifica para uma proteína

exportada de M. tuberculosis, caracterizada por ser seleccionada de entre as sequências seguintes:

- uma sequência IA que corresponde ao encadeamento de nucleótidos descrito na figura 6A ou uma sequência IB que corresponde ao encadeamento de nucleótidos descrito na figura 6B ou uma sequência que codifica para uma proteína exportada de M. tuberculosis e que hibrida em condições de elevada especificidade com estes encadeamentos,

- uma sequência que compreende o encadeamento de nucleótidos IA ou IB e que codifica para uma proteína P28 de M. tuberculosis que tem um peso molecular teórico de cerca de 28 kDa e um peso molecular observado de 36 kDa, determinado por electroforese em gel de acrilamida desnaturante (PAGE-SDS)

- uma sequência compreendida na sequência IA ou IB e que codifica para um polipeptídeo reconhecido por anticorpos dirigidos contra a proteína P28 de M. tuberculosis, que corresponde ao encadeamento de ácidos aminados VIII A - representado na figura 6A ou ao encadeamento de ácidos aminados VIIIB representado na figura 6B,

- uma sequência IV que compreende as sequências reguladoras do gene que compreende a sequência codificadora IA ou IB, em que estas sequências estão contidas nas sequências nucleotídicas a montante e a jusante representadas na figura 7A e na figura 7B,

- uma sequência V que corresponde ao encadeamento compreendido entre os nucleótidos 1 e 72 da sequência IA ou IB e correspondente à sequência de sinal,

- uma sequência VI que corresponde ao encadeamento compreendido entre os nucleótidos 62 a 687 da sequência IA ou IB,

- uma sequência VII que corresponde ao encadeamento

compreendido entre os nucleótidos 688 a 855 da sequência IA ou IB.

Está abrangido igualmente no âmbito da invenção um polipeptídeo de M. tuberculosis, caracterizado por corresponder ao encadeamento de ácidos aminados VIIIA ou ao encadeamento VIIIB representados na figura 6A e 6B, respectivamente, ou por compreender um destes encadeamentos.

Um polipeptídeo preferido é caracterizado por ter um peso molecular teórico de cerca de 28 kDa, determinado de acordo com a técnica descrita nos exemplos.

A proteína p28 de M. tuberculosis foi caracterizada pela sua capacidade de ser exportada e, por conseguinte, potencialmente localizada através da membrana bacteriana plasmática ou da parede celular. Além disso, como o mostram as sequências apresentadas na figura 6, determinados motivos peptídicos da sequência estão repetidos. Por estas razões, a proteína p28 de M. tuberculosis é agora mais frequentemente designada como proteína ERP e o gene que contém a sequência codificadora desta proteína é chamado seja gene irsa, seja gene erp.

O peso molecular teórico da proteína ERP, avaliado em 28 kDa, corresponde a um peso molecular observado experimentalmente de cerca de 36 kDa (migração electroforética em gel de poliacrilamida desnaturante (PAGE-DOS)).

Um outro polipeptídeo interessante no âmbito da invenção compreende uma parte do encadeamento VIII ou VIIIB de ácidos aminados anteriormente descrito e reage imunologicamente com anticorpos dirigidos contra a proteína p28 de M. tuberculosis.

São sequências de ácidos aminados particularmente interessantes no âmbito da invenção as sequências que

compreendem um dos encadeamentos seguintes numa ou em várias cópias:

PGLTS, PGLTD, PGLTP, PALTN, PALTS, PALGG, PTGAT, PTGLD, PVGLD.

Outras sequências interessantes são, por exemplo, a sequência de sinal compreendida entre as posições de nucleótidos 1 e 72 das sequências da figura 6A ou 6B ou ainda a sequência compreendida entre os nucleótidos 688 e 855 susceptível de se comportar como uma sequência de transmembrana.

Estas sequências polipeptídicas podem ser expressas sob a forma de polipeptídeos recombinantes. Nestes polipeptídeos recombinantes, aquelas podem ser substituídas em parte, em especial no que se refere às sequências de 5 ácidos aminados anteriormente descritas, por sequências de interesse provenientes de micobactérias ou de outros organismos patogénicos, podendo esta substituição conduzir a incluir, no seio dos polipeptídeos recombinantes, epítomos ou determinantes antigénicos de um organismo patogénico ou de uma proteína de interesse contra o qual (a qual) se procuraria obter anticorpos.

Deste modo, os polipeptídeos da invenção, enquanto apresentam eventualmente eles próprios propriedades antigénicas ou mesmo imunogénicas, podem ser utilizados como moléculas portadoras interessantes para, eventualmente, preparar vacinas que têm propriedades variadas.

A invenção tem igualmente como objecto anticorpos monoclonais ou soros policlonais dirigidos contra um polipeptídeo tal como definido anteriormente.

Quando se trata de anticorpos monoclonais, estes são de preferência dirigidos especificamente contra um

polipeptídeo da invenção e não reconhecem, por exemplo, a proteína p28 de M. leprae.

A invenção tem também por objecto uma composição para a detecção in vitro de uma infecção por M. tuberculosis, caracterizada por compreender um polipeptídeo definido acima, capaz de reagir imunologicamente com anticorpos formados num doente infectado por M. tuberculosis.

Uma outra composição para a detecção in vitro de uma infecção por M. tuberculosis é caracterizada por uma sequência de nucleótidos que contém pelo menos 9 nucleótidos, proveniente de uma sequência definida acima, ou uma sequência de nucleótidos que contém pelo menos 9 nucleótidos e que hibrida em condições severas com o ADN de M. tuberculosis e não hibrida nas mesmas condições com o ADN de M. leprae, em que esta sequência é uma sequência de ADN ou de ARN eventualmente marcada.

A invenção tem também por objecto um hospedeiro celular procariótico ou eucariótico caracterizado por ser transformado por uma sequência de nucleótidos tal como descrita nas páginas precedentes, sob condições que permitem a expressão desta sequência e/ou a sua exposição ao nível da membrana do hospedeiro celular e/ou a sua exportação e/ou a sua secreção a partir da referida membrana.

De um modo preferido, os hospedeiros celulares são micobactérias tais como M. smegmatis ou M. bovis BCG.

Outros hospedeiros celulares são, por exemplo, E. coli, células CHO, BHK, Spf9/Baculovirus, leveduras tais como Saccharomyces cerevisiae ou o vírus da vacina.

A invenção tem também por objecto uma composição imunogénica que compreende um polipeptídeo tal como apresentado acima ou um hospedeiro celular tal como definido anteriormente.

O pedido de patente descreve, além disso, um vector para a pesquisa e/ou a clonagem e/ou para a expressão de sequências de nucleótidos funcionais em micobactérias, derivado de um vector precedentemente descrito e caracterizado por a sequência codificadora proveniente de um gene que codifica para um marcador de exportação e/ou de secreção ser substituído por um gene relator ou uma sequência relatora.

O gene ou a sequência relatores podem ser desprovidos das suas sequências reguladoras, em particular das suas sequências de ligação ao ribossoma e/ou das suas sequências que permitem a exportação e/ou a secreção do produto marcador quando o vector é incorporado num hospedeiro celular recombinante.

O gene ou a sequência relatores contém a sequência codificadora do gene *lacZ* ou uma parte desta sequência suficiente para que o polipeptídeo apresente uma actividade de β -galactosidase.

Um vector particular é caracterizado por compreender num dos sítios de clonagem do *polylinker* um encadeamento de nucleótidos que compreende um promotor e eventualmente sequências de regulação, por exemplo para a ancoragem à superfície, a exportação ou, segundo os casos, a secreção de um polipeptídeo que seja produzido sob o controlo do promotor, do qual se pretende avaliar a capacidade de promover ou regular a expressão de uma sequência nucleotídica relatora em micobactérias.

São vectores particulares os plasmídeos seleccionados de entre os plasmídeos pJEM12, pJEM13, pJEM14 ou pJEM15 tais como representados na figura 12.

Um destes vectores pode ser utilizado para avaliar o interesse de sequências de regulação ou de expressão ou

promotores, por exemplo as sequências de promotores pAN, pblaF*, psul3, pgroES/EL1.

O pedido de patente descreve igualmente um processo para determinar a actividade de uma sequência que contém num dos sítios de clonagem do *polylinker* um encadeamento de nucleótidos que compreende um promotor e eventualmente sequências de regulação, por exemplo para a exposição, a exportação ou, segundo os casos, a secreção de um polipeptídeo que seja produzido sob o controlo do promotor, em micobactérias caracterizado por compreender as fases de:

- transformação de uma estirpe de micobactérias, por exemplo M. smegmatis ou M. tuberculosis com um vector anteriormente descrito,
- detecção da actividade normalmente associada à presença do gene relator ou da sequência relatora.

Outras características e vantagens da invenção serão evidentes pela leitura dos exemplos que se seguem assim como nas figuras.

LEGENDA DAS FIGURAS

Fig. 1

Construção de pJEM11.

Ver Material e Métodos. pJEM11 tem origens de replicação (ori) de E. coli e de micobactérias. Trata-se, por conseguinte, de um plasmídeo vaivém. O marcador de selecção é um gene de resistência à canamicina (Km). O gene PhoA truncado de pPH07 (22) está desprovido de promotor, de codão de partida e de sequência de sinal, assim a expressão e a exportação de PhoA dependem da fusão por tradução com as extremidades amino-terminais de outras proteínas. O terminador de transcrição (T) do bloco integrado ómega evita a transcrição por "read-through" a partis de sequências plasmídicas.

Fig. 2**Construção dos plasmídeos pLA71, pLA72 e pLA73.**

A inserção no sítio Bam H1 de pJEM11 de fragmentos BlaF* (34) de 3 comprimentos diferentes conduz à expressão de proteínas de fusão com a actividade phoA. Foram efectuadas dosagens colorimétricas de acordo com a técnica de Brockman e Heppel (8), com o p-nitrofenil-fosfato como substrato. Os teores em proteínas foram medidos com o auxílio da dosagem Bio-Rad. As unidades arbitrárias de fosfatase alcalina (aU) foram calculadas como descrito em Material e Métodos.

Fig. 3**Análises de Western-blot de proteínas de fusão PhoA**

Foram cultivadas em meio Beck que continha canamicina (20 µg/ml) estirpes de M. smegmatis transformadas. Foram solubilizados por SDS extractos totais de bactérias submetidas a ultra-sons, resolvidos por SDS-PAGE e submetidos a imunopressão. A preparação do soro de coelho anti-PhoA foi descrita precedentemente (34). Foram utilizados anticorpos de coelho acoplados a PhoA (promega) e, como substrato, uma mistura de X-P e de azul de nitrotetrazólio (BCIP-NBT, Promega) para revelar as fusões PhoA. Coluna 1: PhoA bacteriana purificada, M. smegmatis transformada por plasmídeos pJEM11; coluna 2, pLA71: coluna 3, pLA72: coluna 4, pLA73: coluna 5, pExp410: coluna 6, pExp53: coluna 7, pExp59: coluna 8, pExp421: coluna 9

Fig. 4

Sequências nucleotídicas e de ácidos aminados deduzidas de segmentos de inserções seleccionados dos plasmídeos pExp410, pExp53, pExp59 et pExp421.

Os clones de M. smegmatis com a actividade da fosfatase alcalina foram seleccionados sobre cápsulas de XP/canamicina. Os seus plasmídeos foram amplificados em E. coli XL-1 B e a sequência nucleotídica das inserções foi determinada como descrito em Material e Métodos. A: pExp410 inclui uma parte da lipoproteína de 19 kDa. O quadro de leitura é mantido na junção com phoA (Bam H1/Sau 3A). B: pExp53 inclui uma parte de um gene que apresenta similitudes com o antigénio de 28 kDa de M. leprae. Os ácidos aminados divergentes estão em caracteres em negrito. O codão de início de tradução é GTG. Os sítios putativos de clivagem por sinal de peptidase estão indicados por flechas. C: pExp59 codifica para uma sequência de sinal característica. Um sítio de ligação putativo aos ribossomas (RBS) está sublinhado. O sítio putativo de clivagem pela peptidase de sinal está indicado por uma flecha. D: pExp421 codifica para motivos de ácidos aminados conservados com proteínas da família das dessaturases de esteroil-proteína-portadora de acilo (ACP). R. comm: R. communis (rícino).

Fig. 5

O gene similar ao gene do antigénio de 28 kDa de M. leprae está presente numa única cópia no genoma de M. tuberculosis.

O ADN genómico de M. tuberculosis foi extraído seguindo os protocolos normais (27), digerido pelas endonucleases Pst I, Sma I, Bst EII, Sph I e Bam HI e submetido a uma migração sobre um gel de agarose a 1%. A hibridação de

Southern-blot foi efectuada de acordo com protocolos normais (27). A sonda marcada com ^{32}P era um fragmento de PCR de 180 pb da inserção de pExp53.

Fig. 6

Sequência de nucleótidos (IA e IB) e de ácidos aminados (VIII A e VIII B) do produto do gene IRSA que codifica para a proteína P28 de M. tuberculosis (são apresentadas duas variantes). Este gene é agora designado pela abreviatura "erp" correspondente à expressão "*exported repetitive protein*".

Fig. 7

Sequências de nucleótidos preliminares de flanco do gene IRSA de M. tuberculosis.

Fig. 8

Genes de bactérias para a regulação do ferro (IRG's).

Fig. 9

Perfil de hidrofília DAS PROT2INAS P28 DE M. leprae e M. tuberculosis

Fig. 10

- A) Alinhamento das sequências de nucleótidos do gene que codifica para as proteínas p28 de M. tuberculosis e de M. leprae.
- B) Alinhamento das sequências de ácidos aminados das proteínas p28 de M. tuberculosis e de M. leprae.

Fig. 11**Construção dos plasmídeos pJN3 e pJN11.**

Apenas se apresentam os sítios de restrição e elementos genéticos pertinentes. Os plasmídeos pRR3 e pJN1 foram descritos na técnica anterior (60) (58). O bloco integrado ω foi obtido por digestão por SmaI de pHP45X (59), seguida da purificação de um fragmento de 2 kb de um gel de agarose utilizando o kit Geneclean (Bio 101 Inc.). As técnicas normais de ADN recombinante foram utilizadas de acordo com a descrição dada no estado da técnica (61). Em pJN3 e pJN11, o gene da β -lactamase (bla) foi interrompido, oriE e oriM designam as origens de replicação de pUC (E. coli) e de pAL5000 (micobactérias), respectivamente.

Fig. 12.**estrutura dos plasmídeos da série pJEM.**

(A) Na representação esquemática dos plasmídeos, apenas são indicados os elementos genéticos pertinentes. pJEM15 resultou da clonagem, no sítio ScaI de pRP3, i) de um fragmento obtido por amplificação por PCR (utilizando OJN1 : 5¹-AAGCTTCGATTTCGTAGAGCC-3' e OJN2: 5'-GGGCTCGAGCTGCAG TGGATGACCTTTTGA-3' como iniciadores; e pJN11 como matriz) e que contém tT4 e a extremidade N-terminal de cII; ii) oligonucleótidos de síntese correspondentes a MCS1; e iii) o fragmento lacZ' de HindIII-DraI de pNM480. pJEM12-13-14 foram obtidos por clonagem do fragmento amplificado por PCR descrito acima no sítio ScaI de pRR3. Os oligonucleótidos de síntese correspondentes a MCS2 foram em seguida inseridos. Por fim, cada uma das três formas da série pNM480 foi introduzida no sítio HindIII em MCS2. (B) Sequências nucleotídicas das regiões entre o iniciador OJN1 e o 8.º codão de lacZ'

(marcado ****). Estas sequências foram verificadas experimentalmente. A região tT4 está sublinhada e o RBS de síntese está em caracteres negrito. A sequência de ácidos aminados da extremidade N-terminal de cII é dada por debaixo da sequência de ADN. Os sítios HindIII estão marcados por um asterisco, uma vez que não são únicos. Para descrições complementares, ver a legenda da fig. 11.

EXEMPLOS

I) Identificação de genes que codificam para proteínas exportadas de *M. tuberculosis*.

Os resultados aqui relatados descrevem a definição para micobactérias de um método genético de identificação de proteínas exportadas. Esta metodologia está baseada na fusão de tradução com a fosfatase alcalina bacteriana (PhoA). Estas proteínas de fusão devem ser exportadas para ter a actividade PhoA (6, 13, 16). Um gene PhoA foi utilizado após supressão da região do promotor, do sítio de ligação aos ribossomas e do conjunto da região que codifica para a sequência de sinal cujo codão de início de tradução foi utilizado. Deste modo, a actividade de fosfatase alcalina depende da fusão de tradução realizada na boa fase de leitura com uma parte de uma proteína exportada. A construção de um vector plasmídico phoA para as micobactérias está descrita em primeiro lugar pois se mostrou que a introdução neste vector do gene da β -lactamase exportada de *M. fortuitum* (blaF*) (34) conduz à produção em *M. smegmatis* de proteínas de fusão que apresentam a actividade enzimática PhoA. Construiu-se em seguida um banco de sequências de fusão entre o ADN genómico de *M. tuberculosis* e o gene phoA. Foram isolados doze clones independentes que exportavam proteínas de fusão. Entre estes, foi possível identificar a lipoproteína exportada de 19 kDa já descrita em *M. tuberculosis*, uma

nova sequência de M. tuberculosis que apresenta semelhanças com a proteína de 28 kDa de M. leprae, uma proteína que comporta resíduos de ácidos aminados conservados com dessaturases de esteroil-proteína portadora-de acilo (ACP) e outras sequências novas.

MATERIAL E MÉTODOS

Estirpes bacterianas, plasmídeos e condições de cultura

As estirpes bacterianas e os plasmídeos utilizados neste estudo estão apresentados no quadro 1. O crescimento de estirpes de E. coli e de M. smegmatis, a electroporação, a pesquisa sobre gelose contendo 20 µg/ml de canamicina e 20 µg/ml de fosfato de 5-bromo-4-cloro-3-indolilo (X-P) foram realizados como descrito anteriormente (14).

M. tuberculosis, isolado de um doente (estirpe 103) foi cultivado sobre meio sólido de Lowenstein-Jensen.

Manipulação e sequenciação de ADN

A manipulação de ADN e as análises de Southern-blot foram efectuadas com o auxílio de técnicas normais (27). Para as determinações das sequências, os oligonucleótidos (5'-GGCCGACGAGTCCCGC-3' e 5'-TTGGGGACCCTAGAGGT-3') foram realizados com vista à sequenciação através das junções de fusão das inserções de M. tuberculosis em pJEM11 (ver acima). As sequências de ADN plasmídico de cadeia dupla foram determinadas pelo método de terminação de cadeias por didesoxi (28) utilizando o kit de sequenciação T7 (Pharmacia) de acordo com as instruções do fabricante, ou com o kit de sequenciação Ciclo Terminador Taq Dyc Desoxi (Applied Biosystems), sobre um sistema de PCR GeneAmp 9600 (Perkin Elmer), e foram passadas sobre um sistema de análise de ADN - Modelo 373 (Applied Biosystems).

Análises dos bancos de dados

As sequências nucleotídicas foram comparadas com as dos bancos de dados de EMBL e de GenBank utilizando o algoritmo FASTA (23) e as sequências de proteínas derivadas foram analisadas para determinar uma semelhança eventual com as sequências contidas nos bancos de dados de proteínas PIR e SwissProt utilizando o algoritmo BLAST (1).

construções dos plasmídeos

pJEM11: A construção de pJEM11 está resumida na fig. 1. Em resumo, pJEM2 foi construído a partir do plasmídeo vaivém pRP3 de *E. coli*-micobactérias (26), por inserção do fragmento lacZ truncado de pNM480 (18), de um multissítio de clonagem ou *poly-linker* (MCS) e do terminador de transcrição do bloco integrado ómega (24). O fragmento N-terminal EcoRV-KpnI de lacZ é substituído pelo fragmento phoA truncado de pPH07 (11) sem codão de início ou de sequência de sinal para dar pJEM10. Por fim, um codão de início potencial no MCS foi eliminado para dar pJEM11.

pLA71, pLA72 e pLA73: Inseriram-se fragmentos de comprimento diferente de blaF* (34) obtidos por amplificação por PCR no sítio Bam H1 de pJEM11 para dar pLA71, pLA72 e pLA73 (fig. 2). Os oligonucleótidos (Genset, Paris) utilizados para a amplificação por PCR eram para a Frente 5'-CGGGATCCTGCTCGGCGACTCCGGG-3' e para Trás 5'-CGGGATCCGGTCATCGATCGGTGCCGCGAA-3', 5'-CGGGATCCC GCCGTGCTCGG CCATCTGCAG-3' e 5'-CGGGATCCAGAGTAAG GACGGCAGCACCAG-3', para pLA71, pLA72 e pLA73, respectivamente. As amplificações por PCR foram efectuadas num Cycler Térmico de ADN (Perkin Elmer), utilizando a polimerase Taq (Cetus), de acordo com as recomendações do fabricante.

Construção dos bancos genómicos de M. tuberculosis

Extraíu-se ADN genómico de M. tuberculosis de acordo com os protocolos normais (27) Este ADN foi parcialmente digerido por Sau3A (com 1 U por 2 µg) a 37 °C durante 2 min e 30 s. A digestão foi feita parar pela adição de fenol. Fez-se então migrar este ADN sobre agarose a baixo ponto de fusão (Gibco, BRL). A fracção que continha os fragmentos de 400 a 2000 pb foi extraída por agarase (GELase, Epicentre Technologies) e foi inserida no sítio Bam HI compatível de pJEM11 com a ligase T4 de ADN Boehringer Mannheim), a 16 °C durante uma noite.

Dosagem da fosfatase alcalina

Para as dosagens da fosfatase alcalina, cultivou-se M. smegmatis no caldo L adicionado de tilaxopol a 0,05% (Sigma) a 37 °C durante 48 h. A actividade da fosfatase alcalina foi doseada pelo método de Brockman e Heppel (8) em extractos tratados com ultra-sons como descrito anteriormente (34), utilizando o fosfato de p-nitrofenilo como substrato da reacção. Os teores em proteínas foram medidos com o auxílio da dosagem Bio-Rad (Bio-Rad). A actividade da fosfatase alcalina é expressa em Unidades arbitrárias (aU) = $DO_{420} \times 105 \times 1 \text{ g de proteína}^{-1} \times \text{min}^{-1}$.

Preparações de anticorpos, electroforese em gel de SDS-poliacrilamida e imunopressão

A preparação de um soro de coelho anti-PhoA foi descrita precedentemente (34). Prepararam-se por tratamento por ultra-sons extractos celulares de M. smegmatis, a SDS-PAGE e a imunopressão foram realizadas como descrito precedentemente (36)

RESULTADOS

Construção de um vector plasmídico vaivém (pJEM11) para a produção de proteínas de fusão com PhoA em M. smegmatis

pJEM11 possui um gene phoA truncado de E. coli sem codão de início ou quaisquer elementos de regulação (fig. 1). O sítio múltiplo de clonagem permite a inserção de fragmentos provenientes de genes que codificam para proteínas exportadas putativas ao mesmo tempo que os seus elementos de regulação. Deste modo, foi possível produzir proteínas de fusão que expressavam a actividade da fosfatase alcalina quando a fusão era exportada. pJEM11 é um plasmídeo vaivém de E. coli/micobactérias que inclui o gene de resistência ao antibiótico canamicina de tn903 como marcador de selecção.

A inserção de elementos genéticos responsáveis pela expressão e pela exportação de β -lactamase em pJEM11 conduz à produção de proteínas de fusão PhoA enzimaticamente activas em M. smegmatis

Os três plasmídeos foram construídos por inserção de fragmentos de comprimento diferente provenientes do gene da β -lactamase da estirpe superprodutora M. fortuitum D316 (blaF*) (34) no sítio Bam HI de pJEM11 (fig. 2). Em pLA71, o fragmento de 1384 pb inclui o promotor, o segmento que codifica para os 32 ácidos aminados da sequência de sinal e os 5 primeiros ácidos aminados da proteína madura (não existe sequência Shine-Dalgarno para a fixação ribossómica na sequência de origem de blaF*). pLA72 possui um fragmento de 1550 pb que inclui os elementos que codificam para a sequência de sinal e os 61 primeiros ácidos aminados da proteína madura. Em pLA73, o fragmento de 2155 pb contém a blaF* inteira. Estes plasmídeos foram utilizados para transformar M. smegmatis e os transformantes foram

pesquisados para as fusões PhoA enzimaticamente activas por revelação sobre meios de gelose que continham canamicina e X-P. O X-P é solúvel e incolor, mas após clivagem do fosfato pela fosfatase alcalina, produz-se um precipitado azul. Deste modo, era possível identificar facilmente os clones que produziam a fosfatase alcalina pela sua coloração azul. A expressão de pLA71, 72 e 73 em M. smegmatis conduz a colónias azuis, enquanto que as colónias com pJEM11 continuaram brancas. Por análises de Western-blot verificou-se a produção de proteínas de fusão phoA com um peso molecular aparente de cerca de 47,5 kDa, 54 kDa e 76 kDa, para pLA71, pLA72 e pLA73 respectivamente (fig. 3, coluna 3, 4, 5). Estes pesos moleculares estão de acordo com o comprimento da proteína madura fundida com a fosfatase alcalina (p. M. aparente de 46 kDa, fig. 3, coluna 1). Em pJEM11, não há expressão de PhoA, como previsto (fig. 3, coluna 2). A dosagem da actividade da fosfatase alcalina (ver a fig. 2) destas bactérias confirma a expressão de uma actividade enzimática com as 3 construções pLA. No entanto, M. smegmatis com pLA73 expressa aproximadamente duas vezes mais actividade em comparação com pLA72 e 71. Em experiências distintas, confirmámos que a produção intracelular de phoA sob o controlo de um promotor micobacteriano, sem fusão com uma proteína exportada, não estava associada à expressão da actividade da fosfatase alcalina. O conjunto destes resultados indica que neste sistema a actividade da fosfatase alcalina depende da fusão de tradução e da exportação propriamente dita do produto. Por consequência, pJEM11 convém para a identificação genética das proteínas exportadas pelas micobactérias.

Construção em M. smegmatis de um banco de fusões phoA com fragmentos de ADN genómico de M. tuberculosis

O ADN genómico de um produto isolado clínico de M. tuberculosis foi purificado e parcialmente digerido por Sau3A. A fracção de 400/2000 pb foi inserida no sítio Bam HI compatível de pJEM11. Os produtos de ligação foram transferidos para E. coli XL-1 azul por electroporação para obter uma fase de amplificação. Cerca de 2500 clones que continham plasmídeos com inserções cresceram sobre um meio de gelose que continha canamicina. Os plasmídeos purificados a partir de transformantes foram reunidos e transferidos por electroporação para M. smegmatis MC²155. As bactérias transformadas foram expostas sobre gelose L-canamicina-X-P. Foram obtidos cerca de 14000 clones. Após 4 dias de incubação, observaram-se as primeiras colónias azuis, portanto PhoA+. As cápsulas foram verificadas todos os dias e foram isoladas novas colónias PhoA+. As colónias clonadas foram lisadas e o seu ADN foi introduzido por electroporação em E. coli XL-1 azul, para as preparações de plasmídeos. Ao todo, foram isoladas e sequenciadas 12 inserções diferentes que permitiam a expressão de phoA. Três sequências apresentavam semelhanças com sequências conhecidas.

Fusão de PhoA com o gene da lipoproteína de 19 kDa de M. tuberculosis

Um dos plasmídeos (pExp410) possui uma inserção correspondente a uma parte do gene da proteína de 19 kDa já conhecida. Este gene codifica para uma lipoproteína exportada (5, 31). A fig. 4 mostra a sequência de ADN correspondente à fusão entre este gene e phoA. Como previsto, a mesma fase de leitura é mantida entre as duas proteínas. O peso molecular esperado da proteína de fusão,

segundo a sequência, deverá ser de aproximadamente 57 kDa. No entanto, o peso molecular verdadeiro observado por análise de Western-blot é idêntico ao da proteína PhoA purificada (fig. 3, coluna 1 e 6), o que sugere que a proteína de fusão está clivada perto da junção PhoA.

Fusão com uma sequência similar ao gene da proteína de 28 kDa de M. leprae.

A proteína de 28 kDa de M. leprae é um antigénio principal que é frequentemente reconhecido pelos soros de doentes atingidos pela forma lepromatosa da lepra (9). No banco de inserção de M. tuberculosis preparado, foi identificada uma sequência incluída num vector recombinante (pExp53) que apresenta uma semelhança de 77% com a sequência nucleotídica deste gene e de 68% para a sequência de ácidos aminados deduzida (fig. 4 B). Na análise por Western-blot, o peso molecular da proteína de fusão é de cerca de 52 kDa (fig. 3, coluna 7), o que faz prever cerca de 45 ácidos aminados da proteína micobacteriana na proteína de fusão, após clivagem do peptídeo de sinal. Isto está conforme com o comprimento do fragmento do gene de M. tuberculosis fundido com phoA (fig. 4 B).

Foram realizadas análises de Southern-blot do ADN genómico de M. tuberculosis. Demonstrou-se que um fragmento de 180 pb da inserção de 2 kb do plasmídeo pExp53 não contém qualquer sítio de restrição para as endonucleases PstI, SmaI, BamHI, BstEII e SphI. Este fragmento foi amplificado por PCR. O ADN genómico de M. tuberculosis foi digerido com o auxílio destas enzimas e foi sondado com o fragmento de PCR marcado com 32P. Como se pode ver na fig. 5, foi observada apenas uma banda quando o ADN genómico foi digerido com cada uma das cinco enzimas, o que sugere que o gene está presente numa única cópia no genoma de M. tuberculosis.

Outras fusões PhoA que incluem as sequências de sinal putativas

A fig. 4 mostra a sequência de uma inserção incluída num vector recombinante (pExp59) fundido com *phoA*. Esta inserção apresenta uma sequência de sinal típica que permite a exportação de proteínas. A sequência apresentada está conforme com as regras habituais tais como estabelecidas nas bactérias Gram-negativas (25). Esta sequência contém dois ácidos aminados carregados positivamente (Arg, Asn) depois do codão de início, seguidos por um peptídeo hidrófobo, com um Gly, correspondente provavelmente a uma laçada na estrutura tridimensional do peptídeo. Um sítio potencial de clivagem pela peptidase de sinal está indicado por uma flecha, o que dá uma proteína de fusão com um peso molecular próximo do de *phoA*, como mostra a fig. 3, coluna 8, em conformidade.

Proteínas de fusão PhoA com motivos de ácidos aminados conservados com dessaturases estearoil-proteína-portadora de acilo (ACP)

As dessaturases ACP são enzimas implicadas nas vias da biossíntese dos ácidos gordos. Em geral, estas enzimas são proteínas de membrana integrais (29). Por análises do plasmídeo pExp421 do banco preparado foram revelados dois motivos de ácidos aminados conservados com dessaturases ACP, um de 9 ácidos aminados e o segundo de 14 ácidos aminados (fig. 4 D). O resto da sequência não revelou qualquer semelhança significativa com proteínas conhecidas.

DISCUSSÃO

Foram encontradas mais de 30 proteínas segregadas em filtrados de BCG ou de *M. tuberculosis* a curto termo, com

um mínimo de lise da bactéria ((1, 19, 38). Estas proteínas foram classificadas de acordo com o respectivo peso molecular e com as suas reactividades imunológicas. Algumas foram caracterizadas de modo mais completo. Por exemplo, as proteínas segregadas do complexo do antigénio 85 (os antigénios 85 A, B e C) são proteínas de 32 kDa que apresentam reacções serológicas cruzadas (7, 35). Os antigénios 85 A e 85 B apresentam uma afinidade para com a fibronectina e estarão implicados na internalização de M. tuberculosis nos macrófagos. Os genes destas proteínas imunogénicas (7) e de proteínas de 23 kDa (MPB64) (37) e de 19 kDa (5) foram clonados e sequenciados e foram encontradas sequências de peptídeos de sinal características de proteínas exportadas. As proteínas recombinantes produzidas a partir destes genes serão ferramentas interessantes para o diagnóstico serológico da tuberculose. A dismutase de superóxido (SOD) de 23/28 kDa é abundante nos filtrados de cultura a curto termo e estará implicada na sobrevivência das micobactérias no fagolisossoma. O gene que codifica para a SOD na M. tuberculosis foi clonado e sequenciado (39). De modo interessante, não foi encontrada qualquer sequência de peptídeo de sinal característica. Isto sugere uma via específica de secreção desta enzima pelas micobactérias. Proteínas segregadas em duas gamas estreitas de peso molecular (6-10 kDa e 26-34 kDa) são antigénios principais das células T (3) e induzem nos ratinhos respostas imunitárias em células T protectoras frente a uma amostra por micobactérias vivas do complexo de M. tuberculosis (4). Foi sugerido que as diferenças nas respostas imunitárias observadas entre as bactérias vivas e mortas são devidas a estas proteínas exportadas/segregadas (20) Estes diversos resultados preliminares sugerem que uma melhor

caracterização de proteínas exportadas/segregadas das bactérias patogénicas do complexo de M. tuberculosis seria de grande utilidade tanto para a compreensão do seu poder patogénico como para a realização de novas vacinas.

Se as proteínas segregadas foram estudadas por métodos bioquímicos, outras metodologias genéticas mostrar-se-iam necessárias. Utilizando um gene *phoA* truncado, foram estabelecidos sistemas de fusão que permitem a fixação das extremidades amino de outras proteínas sobre PhoA. Esta abordagem baseia-se sobre a fosfatase alcalina bacteriana periplasmática de *E. coli*. Esta enzima deve estar localizada de modo extracitoplasmático para ser activa. Deste modo, a fosfatase alcalina pode ser utilizada como sonda de localização subcelular.

Foi elaborada e aqui descrita uma metodologia PhoA para a identificação de proteínas exportadas pelas micobactérias. A inserção de *blaF** em pJEM11 conduz à produção em M. smegmatis de proteínas de fusão com a actividade de fosfatase alcalina. Além disso, fusões de PhoA com 3 fragmentos diferentes de *BlaF** eram enzimaticamente activas, o que sugere que a maior parte das fusões em fase com proteínas exportadas terão uma actividade de PhoA.

Foi construído e expresso em M. smegmatis um banco de inserções de M. tuberculosis em pJEM11. Neste banco, foi isolada uma parte do gene que codifica para a lipoproteína exportada conhecida de 19 kDa (pExp410). Esta proteína de M. tuberculosis é um dos antigénios serologicamente imunodominantes encontrados neste bacilo. Análises da sequência de ADN do gene que codifica para este antigénio indicam que a região NH₂-terminal hidrófoba é um peptídeo de sinal de lipoproteína (5). Uma parte desta lipoproteína foi fundida à proteína A de superfície externa de Borrelia

burgdorferi para construir uma vacina BCG recombinante capaz de induzir uma forte resposta imunitária (31).

Foram igualmente identificadas duas outras sequências que apresentam semelhanças com as proteínas exportadas ou de membrana.

Verificou-se que pExp53 apresenta semelhanças com o gene do antigénio de 28 kDa de M. leprae. Este antigénio de M. leprae foi encontrado por pesquisa de um banco de λ gt11 com o soro de doentes atingidos com a forma lepromatosa da lepra. É um antigénio principal implicado na resposta imunitária humoral à M. leprae (9). De modo interessante, verificou-se que um peptídeo de 20 ácidos aminados desta proteína apresenta uma semelhança considerável com um peptídeo do antigénio de 19 kDa de M. tuberculosis, e é um epítipo de células T que apresentam reacções cruzadas (12). A sequência de ADN do gene que codifica para o antigénio de 28 kDa de M. leprae sugere que "a sequência de ácidos aminados prevista da proteína contém um peptídeo de sinal potencial na sua extremidade amino-terminal e dois longos domínios hidrófobos, o que sugere que a mesma está dirigida para a localização sobre a membrana plasmática bacteriana ou a parede celular" (9).

Uma proteína de fusão codificada por um plasmídeo do nosso banco (pExp421) partilhava motivos de ácidos aminados com as dessaturases. As dessaturases ACP são enzimas implicadas nas vias da biossíntese dos ácidos gordos. Em geral, estas enzimas são proteínas de membrana integrais (39). Este resultado sugere que é possível ter-se isolado uma parte de um gene importante no metabolismo dos lípidos de M. tuberculosis, talvez implicado na via da biossíntese da parede celular lipídica.

Foi encontrado um outro plasmídeo (pExp59) com uma sequência de sinal putativa característica. -

Em conclusão, os resultados apresentados demonstram que a tecnologia de PhoA para a identificação genética de proteínas exportadas pode ser adaptada com êxito para M. tuberculosis. Pesquisas preliminares de um banco de inserções que dão proteínas de fusão PhoA puseram em evidência sequências que apresentam semelhanças com as proteínas exportadas conhecidas.

II) Expressão da proteína P28 de M. tuberculosis

A BCG é uma vacina viva. É a única vacina utilizada para proteger contra a tuberculose. A sua eficácia tem-se revelado variável segundo as populações vacinadas, indo de desde cerca de 80% na Grã-Bretanha até 0% na Índia. Parece, por conseguinte, indispensável investigar uma vacina mais eficaz. Por outro lado, o facto de utilizar uma vacina viva põe hoje problemas por causa da extensão da epidemia da SIDA.

Vários trabalhos mostraram que antigénios exportados por Mycobacterium tuberculosis, o agente da tuberculose, tinham um efeito protector contra um ensaio pela estirpe virulenta. Os trabalhos aqui relatados consistiram em utilizar um método genético para isolar e estudar os genes de M. tuberculosis que codifica para proteínas exportadas. Nós descrevemos aqui o isolamento e a caracterização de um gene que codifica para uma proteína que possui homologias com a proteína de 28 kDa de Mycobacterium leprae já descrita.

Metodologia para a clonagem de genes que codificam para proteínas exportadas.

A metodologia exposta em pormenor na parte I baseia-se na utilização de fusões de tradução com o gene que codifica

para a fosfatase alcalina de Escherichia coli, PhoA. Estas proteínas de fusão apenas têm uma actividade de fosfatase alcalina detectável quando são exportadas. Construiu-se um vector plasmídico que possuía um gene *phoA* desprovido do seu promotor, do seu sítio de acoplamento ao ARN ribossomal e da sua sequência de sinal. A partir deste vector apenas foi possível observar uma actividade de PhoA após fusão de tradução no quadro correcto de leitura com uma proteína exportada. O vector, designado pJEM11, possui uma origem de replicação para E. coli e uma outra para as micobactérias. Possui igualmente um marcador de selecção, o gene de resistência à canamicina do transposão Tn905. Um sítio múltiplo de clonagem precede o gene *phoA* truncado.

Foi construído em pJEM11 um banco de ADN genómico, proveniente de uma estirpe de M. tuberculosis (Mt103) isolada a partir de um doente tuberculoso, por inserção de fragmentos de ADN provenientes de uma hidrólise parcial pela enzima Sau3a. Os clones seleccionados permitiram identificar um fragmento de nucleótido do gene de 28 kDa de M. tuberculosis homólogo do gene que codifica para a proteína de 28 kDa de M. leprae.

Nos doentes lepromatosos, observam-se anticorpos dirigidos contra esta proteína de 28 kDa, deixando supor que esta proteína é um antigénio imunodominante. Fez-se a hipótese de que em M. tuberculosis, a proteína de 28 kDa que possui homologias com a proteína de 28 kDa de M. leprae poderia ser também um antigénio imunodominante e que poderia servir para a construção de ensaios imunológicos específicos que permitissem a detecção da infecção tuberculosa ou da doença tuberculose. Poderia talvez ser utilizada para a construção de vacinas sub-unitárias em diferentes preparações de vacinas. Além disso, poderia ser útil como vector para a

expressão de antígenos nas micobactérias para a construção de vacinas recombinantes.

Clonagem e sequenciação do gene que codifica para uma proteína de 28 kDa de M. tuberculosis

Utilizando a inserção contida no plasmídeo pExp53 como sonda, clonou-se o gene inteiro que codifica para a proteína de 28 kDa de M. tuberculosis por hibridação sobre colônia de um banco de ADN de M. tuberculosis construído por inserção de fragmentos de ADN de M. tuberculosis de dimensão compreendida entre 2 e 6 kb obtidos por hidrólise total com a enzima PstI no vector pBluescript KS-. O fragmento PstI de M. tuberculosis correspondente ao clone positivo e que compreende uma inserção de 4,1 kb foi sequenciado. A figura 10 mostra a sequência nucleotídica do fragmento e as semelhanças com o gene que codifica para a proteína de 28 kDa de M. leprae. A sequência do gene de 28 kDa de M. tuberculosis é como a de M. leprae precedida de uma sequência que possui semelhanças com as caixas "ferro" encontradas a montante dos genes expressos quando de uma carência em ferro. Verifica-se uma situação de carência de ferro quando do crescimento in vivo. A hipótese feita é de que a expressão deste gene é induzida quando do crescimento nos macrófagos das micobactérias que contêm este gene. Além disso, a proteína de 28 kDa de M. tuberculosis possui na sua parte central duas regiões que contêm motivos de 5 ácidos aminados repetidos em tandem, que estão ausentes da proteína homóloga de M. leprae. Foram identificadas anteriormente estruturas repetidas análogas nos antígenos principais presentes na superfície de outros agentes patogênicos bacterianos ou parasitários, tais como a proteína M (40) dos *Streptococcaceae* e a proteína CS dos *Plasmodiae* (41).

A totalidade ou parte da proteína de 28 kDa de M. tuberculosis cuja sequência do gene é aqui apresentada poderia ser um antigénio protector potencial para a construção de uma vacina contra a tuberculose. Um tal antigénio pode ser obtido por purificação a partir de extractos celulares de M. tuberculosis ou a partir de extractos heterólogos geneticamente recombinados. Além disso, a proteína de 28 kDa de M. tuberculosis ou peptídeos que derivam desta poderiam ser um antigénio utilizável nos ensaios ELISA para a despistagem de doentes tuberculosos.

Utilizando o gene de 28 kDa de M. tuberculosis como sonda, apenas se observou uma hibridação nas condições de elevado rigor com o ADN genómico das estirpes que pertenciam ao complexo M. tuberculosis. Por consequência, a sequência correspondente ao gene de 28 kDa de M. tuberculosis é uma sequência específica que pode ser utilizada para ensaios de detecção dos bacilos da tuberculose, utilizando sondas de ADN ou ARN e métodos de amplificação genética in vitro.

A região reguladora e o gene da 28 kDa de M. tuberculosis podem ser utilizados como moléculas portadoras para expressar antigénios heterólogos na BCG ou qualquer outro vector micobacteriano útil para a construção de vacinas.

III) Expressão de genes de micobactérias; avaliação de diferentes promotores de expressão

Um aspecto importante dos resultados obtidos diz respeito à construção de ferramentas genéticas para o estudo da expressão dos genes nas micobactérias. Foram utilizados vectores sequência de regulação-sonda na técnica anterior para isolar e analisar sequências de regulação num grande número de bactérias (54). A definição pelos inventores destas ferramentas próprias das micobactérias facilita o

estudo dos mecanismos genéticos que regulam a virulência nas espécies patogênicas e o isolamento de novas sequências de regulação que seriam úteis para a realização de vacina BCG recombinante melhorada.

Inicialmente, a expressão de genes micobacterianos foi estudada nos sistemas heterólogos, Escherichia coli e Streptomyces lividans (46) (51) (60). Estas análises sugerem que a maioria dos genes micobacterianos são expressos mais eficazmente em S. lividans que em E. coli. Ulteriormente, foram construídos vectores à base de plasmídeos micobacterianos que poderiam ser utilizados para estudos nos sistemas homólogos. Os vectores pYUB75 e pYUB76 foram concebidos para seleccionar fusões de genes a um gene lacZ truncado de Escherichia coli (42). O plasmídeo pSD7 permite a construção de fusões de operões a um gene da acetiltransferase de cloranfenicol (CAT) sem promotor (47). Utilizando este vector, foi isolado um certo número de sequências de regulação e estas foram avaliadas tanto em E. coli como em Mycobacterium smegmatis.

Os inventores descreveram outras construções de vectores da série pJEM que têm várias vantagens: estes vectores compreendem um terminador de transcrição e multissítios de clonagem adaptados e permitem fusões tanto de operões como de genes lacZ. Escolheu-se lacZ como gene relator uma vez que a enzima codificada, a β -galactosidase, continua activa quando se fundem sequências heterólogas na sua extremidade amino-terminal (45) (64). A sua actividade pode ser facilmente medida in vitro, mesmo a níveis muito baixos com o auxílio de compostos fluorescentes (48). A β -galactosidase é igualmente muito imunogénica. Esta enzima induz respostas imunitárias tanto humorais como celulares após apresentação ao sistema imunitário de ratinhos por bactérias recombinantes (44) (56). Deste modo, a

β -galactosidase pode igualmente ser utilizada como relator do poder imunogénico de uma vacina recombinante. Utilizando vectores pJEM, é possível isolar novas sequências de regulação activas em BCG e ensaiar as estirpes BCG recombinantes no que se refere à sua capacidade para induzir respostas imunitárias em ratinhos.

Foi igualmente levado a efeito um estudo comparativo das actividades de diversos promotores em M. smegmatis e BCG. Os resultados sugerem que as polimerases de ARN de M. smegmatis e de BCG não partilham a mesma especificidade.

Construção de vectores pJEM. Idealmente, um vector plasmídico promotor-sonda deve conter cinco elementos: i) um replicação, ii) um marcador de selecção e um bloco integrado relator que contém iii) um terminador de transcrição seguido iv) de multissítios de clonagem (MCS) e v) de um gene relator desprovido das suas sequências de regulação.

Para construir um vector de clonagem de promotor as micobactérias, nós utilizámos o replicação derivado do plasmídeo pAL5000 de Mycobacterium fortuitum e o gene da resistência à canamicina (aph) de Tn903 (58). Estes elementos genéticos são os componentes de base da maioria dos plasmídeos actualmente utilizados para a transformação das micobactérias. Parece que estes elementos conferem uma elevada estabilidade aos clones transformados de M. smegmatis e M. bovis BCG tanto *in vitro* como *in vivo* (no ratinho) mesmo na ausência de selecção pelos antibióticos (56). Para facilitar a preparação e a manipulação de ADN episomal, a maior parte destes plasmídeos contém igualmente um replicação de E. coli. Deste modo, nós escolhemos o plasmídeo pRR3, um vector vaivém E. coli-micobactérias que contém estes três elementos genéticos, como vector de base (58).

Não foi ainda caracterizado qualquer terminador de transcrição micobacteriano. Para examinar se o terminador de transcrição do colifago T4 (tT4) estava activo como sítio de terminação para as polimerases de ARN de micobactérias, o interposição ómega (57) foi clonado no plasmídeo pJN3, a montante do elemento sRBS-cII-lacZ, gerando pJN11 (fig. 11). O fragmento ómega é composto por um gene de resistência à estreptomicina/espectinomicina enquadrado por curtas repetições invertidas que contêm tT4. A inserção de ómega num fragmento de ADN conduz à terminação da síntese de ARN em E. coli (57). Construiu-se pJN3 por clonagem, no sítio ScaI de pRR3, de um bloco integrado composto por um lacZ truncado associado a um RBS de síntese (sRBS) e a extremidade 5' do gene de regulação cII do fago lambda e o promotor pL (fig. 11). Transformou-se M. smegmatis mc2155 (61) com pJN3 (pL-sRBS-cII-lacZ) ou pJN11 (pL-XsRBS-cII-lacZ) por electroporação e os clones transformados foram detectados após crescimento sobre cápsulas LB-Xgal. Os clones transformantes que possuem pJN3 deram colónias azuis e os clones transformantes que possuem pJN11 deram colónias brancas. A actividade da β -galactosidase em M. smegmatis (pJN11) era 50 vezes inferior à de M. smegmatis (pJN3) (quadro 2). Deste modo, o tT4 contido na inserção X age como terminador de transcrição eficaz em M. smegmatis.

Sintetizou-se in vitro um fragmento de ADN que contém o segmento tT4 seguido do elemento sRBS-cII-lacZ de pJN11 por amplificação por PCR e adicionou-se um MCS (MCS1), que contém 6 sítios de restrição únicos. O bloco integrado resultante foi subseqüentemente clonado no sítio ScaI de pRR3, dando o vector de fusão de operões pJEM15 (fig. 12). A electroporação de M. smegmatis MC²155 e do BCG com este plasmídeo conduziu a colónias brancas sobre as cápsulas de

LB-Xgal com uma actividade muito fraca da β -galactosidase (quadro 2). Pelo contrário, em E. coli, pJEM15 expressou uma actividade da β -galactosidase mais elevada e, por consequência, uma coloração azul sobre as cápsulas de LB-Xgal. Isto é provavelmente devido ao seu número de cópias elevado. Em E. coli, estão presentes vectores num número de cópias elevado (superior a 500), enquanto que nas micobactérias, os plasmídeos derivados de replicões pAL5000 têm um número de cópias de aproximadamente 3 a 10 (50). O ensaio de fragmentos de ADN para a actividade de promotor, com o auxílio de pJEM15, por pesquisa azul-branco, deve ser, deste modo, efectuado directamente nas micobactérias. Para obter vectores que permitam fusões de genes de lacZ, nós seguimos uma estratégia similar. As três formas de lacZ truncado da série pNM480 (55), que diferem umas das outras por "pôr em fase de tradução" um sítio HindIII localizado na sua extremidade 5', foram clonadas a jusante de tT4 e de um MCS (MCS2) que continha 7 sítios de restrição únicos, no sítio ScaI de pRR3. Os plasmídeos resultantes pJEM12-13-14 (fig. 12) permitem, deste modo, a clonagem de uma ampla gama de fragmentos de restrição em fase com lacZ.

Avaliação de diversos promotores em M. smegmatis e BCG. Construíram-se fusões de operações entre o bloco integrado relator cII-lacZ e os promotores pAN (56), pblaF* (63), psul3 (52) e pgroES/EL1 (49). A actividade destes promotores foi avaliada em M. smegmatis e em M. bovis BCG. Os três primeiros promotores foram isolados a partir de espécies micobacterianas. pblaF* é um mutante de expressão elevada de pblaF, que dirige a expressão do gene da p-lactamase de M. fortuitum; pAN é um promotor de M. paratuberculosis e psul3 um componente de elemento genético móvel de M. fortuitum Tn610. Estes promotores foram localizados sobre a base da cartografia de sítios de início

de transcrição (pblaF* e pAN) ou por análise de eliminação (psul3) (62). pgroES/EL1 é um promotor de Streptomyces albus que regula a expressão do operão groES/EL1 e é activo tanto em M. smegmatis como em BCG (65).

As experiências de clonagem foram efectuadas directamente em M. smegmatis. Isolaram-se fragmentos de ADN que continham cada um dos promotores e inseriram-se em MCS1 de pJEM15 digerido pelas enzimas de restrição apropriadas. As misturas de ligação resultantes foram utilizadas para transformar M. smegmatis mc2155 por electroporação e seleccionaram-se colónias para electroduzir E. coli MC1061 (45) como descrito anteriormente (43). Isolaram-se e analisaram-se plasmídeos destes clones de E. coli. Os que correspondiam às construções pretendidas pJN29 a pJN32 (quadro 2) foram utilizadas para a electroporação de BCG (estirpe Pasteur).

A actividade da β -galactosidase foi doseada sobre extractos submetidos s ultra-sons de M. smegmatis e de BCG (quadro 2). A actividade dos promotores variou consideravelmente tanto entre os promotores num hospedeiro micobacteriano como entre os hospedeiros para cada promotor. A força relativa destes promotores não era a mesma em M. smegmatis e BCG. Se bem que pblaF* fosse o promotor mais potente tanto em M. smegmatis como no BCG, a situação é diferente para os outros promotores. pAN e pgroES/EL1 eram mais activos que psul3 em BCG, mas em M. smegmatis, psul3 era mais activo que pAN ou pgroES/E11.

Das Gupta e colaboradores (47) efectuaram buscas em bancos de ADN de M. smegmatis e de M. tuberculosis para a actividade de promotor em M. smegmatis. Estes autores assinalaram uma frequência de promotores 10 a 20 vezes mais elevada no ADN de M. smegmatis. Além disso, promotores muito activos eram mais raros nos bancos de ADN de M.

tuberculosis do que nos de M. smegmatis. Estes autores sugeriram que os promotores de M. tuberculosis podem ter divergido consideravelmente dos de M. smegmatis. Os resultados aqui apresentados sugerem que a maquinaria de transcrição de M. smegmatis e de M. bovis BCG, espécie muito aparentada com M. tuberculosis, pode ser diferente. Em conclusão, a família de vectores construídos facilita o estudo da expressão de genes nas micobactérias. Fragmentos de uma ampla gama podem ser facilmente clonados em fase em lacZ' (fusão de genes) ou a montante de cII-lacZ (fusão de operões) e podem ser avaliados para a actividade de promotor por pesquisa azul-branco de transformantes micobacterianos sobre cápsulas LB-Xgal. A actividade destes promotores pode também ser medida (por dosagem da actividade da β -galactosidase), as suas sequências determinadas e o seu sítio de início de transcrição cartografado (por análises de extensão de iniciador) utilizando "o iniciador universal" ou sequências aparentadas (53) como iniciador.

IV) Expressão da proteína ERP sob forma recombinante em E. coli

A proteína ERP foi expressa sob forma recombinante em E. coli e foi purificada por cromatografia de afinidade. Foram realizados dois tipos de fusões entre ERP e fragmentos peptídicos que tinham uma forte afinidade para suportes cromatográficos específicos (Amilose, sistema MaleE; Níquel Ni^{2+}) quelado, para o sistema Histidina). Trata-se de

- ERP sem a sua sequência de sinal fundida em C-ter com a proteína que liga a maltose (MaleE) de E. coli (MaleE-ERP);
- ERP sem a sua sequência de sinal (ERP(His)₆ ss) ou na sua totalidade (ERP(His)₆) e que possui 6 ácidos aminados histidina em C-ter.

Após purificação, a análise destas três proteínas de fusão em electroforese PAGE-SDS indica que o polipeptídeo ERP possui um peso molecular (PM) relativo de 36 kDa. Existe uma diferença importante entre o PM calculado a partir da sequência (28 kDa) e o PM observado experimentalmente (36 kDa). Este atraso na migração electroforética poderia resultar do forte teor em resíduos prolina ou de modificações pós-tradução.

REFERÊNCIAS

1. Altschul, B.F. *et al.*, 1990, *J. Mol.. Biol.*, 215: 403-410.
2. Andersen, P. *et al.*, 1991, *Infect. Immun.* 59: 1905-1910.
3. Andersen, P. *et al.*, 1991, *Infect. Immun.* 59: 1558-1563.
4. Andersen, P. *et al.*, 1994, *Immun.* 62: 2536-2544.
5. Ashbridge, K.R. *et al.*, 1989, *Nucl. Acid. Res.* 17: 1249.
6. Boquet, P. *et al.*, 1987, *J. Bacteriol.* 169: 1663-1669.
7. Borremans, M. *et al.*, 1989, *Infect. Immun.* 57: 3123-3130.
8. Brockman, R.W. *et al.*, 1968, *Biochemistry* 7: 2554-2561.
9. Cherayil, B. *et al.*, 1988, *J. Immunol.* 12: 4370-4375.
10. Gaillard, J.-L. *et al.*, 1991, *Cell* 65: 1127-1141.
11. Gutierrez, C. *et al.*, 1989, *Nucl. Acids. Res.* 17: 3999.
12. Harris, D.P. *et al.*, 1991, *J. Immunol.* 147: 2706-2712.
13. Hoffman, C.S. *et al.*, 1985, *Proc. Natl. Acad. Sci. USA* 82: 5107-5111.
14. Isberg, R.R. *et al.*, 1987, *Cell* 50: 769-778.
15. Knapp, s. *et al.*, 1988, *J. Bact.* 170: 5059-5066.
16. Manoil, C. *et al.*, 1990, *J. Bacteriol.* 172: 515-518.

17. Miller, V.L. *et al.*, 1987, *cell*. 48: 271-279.
18. Minton, N.P., 1984, *Gene*. 31: 269-273.
19. Nagal, s. *et al.*, 1991, *Infect. Immun.* 59: 372-382.
20. Orme, I.M., 1988, *Infect. Immun.* 56: 3310-3312.
21. Orme, I.M. *et al.*, 1993, *J. Infect. Disea.* 167: 1481-1497.
22. Pearce, B.J. *et al.*, 1993, *Mot. Microbiol.* 9: 1037-1050.
23. Pearson, W.R. *et al.*, 1988, *Proc. Natl. Acad. Sci. USA.* 85: 2444-2448.
24. Prentki, P. *et al.*, 1984, *Gene*. 29: 303-313.
25. Pugsley, A.P., 1993, *Microbiol. Rev.* 57: 50-108.
26. Ranes, L.G. *et al.*, 1990, *J. Bacteriol.* 172: 2793-2797.
27. Sambrook, J. *et al.*, 1989. *Molecular Cloning: a Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbour, N.Y.
28. Sanger, F. *et al.*, 1977, *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.
29. Shanklin, J. *et al.*, 1991, *Proc. Natl. Acad. Sci. USA* 88: 2510-2514.
30. Snapper, S.B. *et al.*, 1990, *Mot. Microbiol.* 11: 1911-1919.
31. Stover, R.C. *et al.*, 1993, *J. Exp. Med.* 178: 197-209.
32. Taylor, R.K. *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* 84: 2833-2837.
33. Taylor, R.K. *et al.*, 1989, *J. Bact.* 171: 1870-1878.
34. Timm, J. *et al.*, 1994, *Mol. Microbiol.* 12: 491-504.
35. Wiker, H.G. *et al.*, 1992, *Microbiol. Rev.* 56: 648-661.
36. Winter, N. *et al.*, 1991, *Gene*. 109: 47-54.

37. Yamaguchi, R. *et al.*, 1989, *Infect. Immun.* 57: 283-288.
38. Young, D.B. *et al.*, 1992, *Mol. Microbiol.* 6: 133-145.
39. Zhang, Y. *et al.*, 1991, *Mol. Microbiol.* 5: 381-391.
40. Hollingstead S. *et al.*, 1986, *J. Biol. Chem.* 262: 1677-1686.
41. Zavala, F. *et al.*, *J. Exp. Med.* 157: 194-1957.
42. Barletta, R.G. *et al.*, 1992, *J. Gen. Microbiol.* 138: 23-30.
43. Baulard, A. *et al.*, 1992, *Nucleic Acids Res.* 20: 4105.
44. Brown, A. *et al.*, 1987, *J. Infect. Dis.* 155: 86-92.
45. Casabadian, M.J. *et al.*, 1980, *J. Bacteriol.* 143: 971-980.
46. Clark-Curtiss, J.E. *et al.*, 1985, *J. Bacteriol.* 161: 1093-1102.
47. Das Gupta, S.K. *et al.*, 1993, *J. Bacteriol.* 175: 5186-5192.
48. Garcia-del-Portillo, F. *et al.*, 1992, *Mol. Microbiol.* 6: 3289-3297.
49. Guglielmi, G. *et al.*, 1993, *Basic and Applied Genetics*. American Society for Microbiology, Washington, D.C.
50. Hatfull, G.H. *et al.*, 1993. Genetic transformation of mycobacteria. *TIM* 1: 310-314.
51. Kieser, T. *et al.*, 1986, *J. Bacteriol.* 168: 72-80.
52. Martin, C. *et al.*, 1990, *Nature* 345: 739-743.
53. Messing, J., 1983, New M13 vectors for cloning, p. 20-78. In R. Wu, L. Grossman and K. Moldave (eds.), *Methods in Enzymology*. Academic Press, New York.
54. Miller, J.H., 1991, *Bacterial Genetic Systems*, In J.N. Abelson and M.I. Simon (eds.), *Methods in Enzymology*, Academic Press, San Diego.
55. Minton, N.P., 1984, *Gene* 31: 269-273.

56. Murray, A. *et al.*, 1992, *Mol. Microbiol.* 6: 3331-3342.
57. Prentki, P. *et al.*, 1984, *Gene* 29: 303-313.
58. Ranes, M.G. *et al.*, 1990, *J. Bacteriol.* 172: 2793-2797.
59. Sambrook, J. *et al.*, 1989, *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
60. Sirakova, T.D. *et al.*, 1989, *FEMS Microbiol. Lett.* 59: 153-156.
61. Snapper, S.B. *et al.*, 1990, *Mol. Microbiol.* 4: 1911-1919.
62. Timm, J. *et al.* Unpublished ta.

Quadro 1

Estirpe/plasmídeo	Características relevantes	Referência
<i>E. coli</i> XL1-Azul	supE44 hsdR17 recA1 8yrA46 thi relA1 lac- F'	27
<i>M. smegmatis</i> mc ² 155	Mutante altamente transformado de <i>M. smegmatis</i> ATCC607	30
pRR3	Vector vaivém de <i>E. coli</i> - micobactéria	26
pPHO7	Derivado de pUC com um gene <i>phoA</i> truncado	11
pNM480	Derivado de pUC com um gene <i>lacZ</i> truncado	18
pJEM11	Vector vaivém de <i>E. coli</i> - micobactéria com um gene <i>phoA</i> truncado	este trabalho
pLA71	pJEM11 no qual foi clonado um fragmento de 1.384 pb de <i>biaF*</i>	34, este trabalho
pLA72	pJEM11 no qual foi clonado um fragmento de 1.550 pb de <i>biaF*</i>	34, este trabalho
pLA73	pJEM11 no qual foi clonado o <i>biaF*</i> completo	34, este trabalho
pExp410	pJEM11 no qual foi clonada parte do gene do antígeno de <i>M. tuberculosis</i> de 19este trabalho kDa	este trabalho
pExp53	pJEM11 no qual foi clonada parte de um gene de <i>M. tuberculosis</i> semelhante ao gene do antígeno de <i>M. leprae</i> de 28 kDa	este trabalho
pExp59	pJEM11 no qual foi clonada a sequência de sinal de um gene não identificado de <i>M. tuberculosis</i>	este trabalho
pExp421	pJEM11 no qual foi clonado um gene de <i>M. tuberculosis</i> que codifica para uma proteína com motivos de ácidos aminados semelhantes a dessaturases	este trabalho

REIVINDICAÇÕES

1. Sequência de nucleótidos proveniente de um gene que codifica para uma proteína exportada de M. tuberculosis, caracterizada por ser selecionada de entre as sequências seguintes:

- uma sequência IA que corresponde ao encadeamento de nucleótidos representado na figura 6A ou uma sequência IB que corresponde ao encadeamento de nucleótidos representado na figura 6B ou uma sequência que codifica para uma proteína exportada de M. tuberculosis e que hibrida em condições de elevada especificidade com uma sequência IA ou IB,
- uma sequência que compreende o encadeamento de nucleótidos IA ou compreende o encadeamento de nucleótidos IB e que codifica para uma proteína P28 de M. tuberculosis que tem um peso molecular teórico de cerca de 28 kDa,
- uma sequência compreendida na sequência IA ou na sequência IB e que codifica para um polipeptídeo reconhecido por anticorpos dirigidos contra a proteína P28 de M. tuberculosis, que corresponde ao encadeamento de ácidos aminados VIIIA representado na figura 6A ou ao encadeamento de ácidos aminados VIIIB representado na figura 6B.
- uma sequência IV que compreende as sequências reguladoras do gene que compreende a sequência codificadora IA ou a sequência codificadora IB, em que estas sequências estão contidas nas sequências nucleotídicas a montante e a jusante representadas na

figura 7A e na figura 7B,

- uma sequência V que corresponde ao encadeamento compreendido entre os nucleótidos 1 a 72 da sequência IA ou da sequência IB e correspondente à sequência de sinal,
- uma sequência VI que corresponde ao encadeamento compreendido entre os nucleótidos 62 a 687 da sequência IA ou da sequência IB,
- uma sequência VII que corresponde ao encadeamento compreendido entre os nucleótidos 688 e 855 da sequência IA ou da sequência IB.

2. Polipeptídeo de M. tuberculosis, caracterizado por corresponder ao encadeamento de ácidos aminados VIIIA ou ao encadeamento de ácidos aminados VIIIB representados respectivamente na figura 6A e na figura 6B ou por compreender este encadeamento.

3. Polipeptídeo de acordo com a reivindicação 2, caracterizado por ter um peso molecular teórico de cerca de 28 kDa.

4. Polipeptídeo de acordo com a reivindicação 2 ou 3, caracterizado por ter um peso molecular observado de cerca de 36 kDa, determinado por migração electrofónica em gel de poliacrilamida desnaturante (PAGE-DOS).

5. Polipeptídeo caracterizado por compreender uma parte do encadeamento VIIIA ou do encadeamento VIIIB de ácidos aminados de acordo com a reivindicação 2 e por reagir

imunologicamente com anticorpos dirigidos contra a proteína P28 de M. tuberculosis que corresponde ao encadeamento VIIIA ou ao encadeamento VIIIIB de ácidos aminados.

6. Polipeptídeo de acordo com a reivindicação 2, 3 ou 4, caracterizado por compreender um ou vários dos encadeamentos seguintes: PGLTS, PGLTD, PGLTP, PALTN, PALTS, PALGG, PTGAT, PTGLD, PVGLD.

7. Polipeptídeo de acordo com uma das reivindicações 2 a 4, caracterizado por um ou vários dos encadeamentos seguintes:

PGLTS, PGLTD, PGLTP, PALTN, PALTS, PALGG, PTGAT, PTGLD, PVGLD

estar substituído por uma sequência de ácidos aminados de um epítopo ou de um determinante antigénico de um organismo patogénico determinado.

8. Anticorpos monoclonais ou soro policlonal dirigidos contra um polipeptídeo de acordo com uma qualquer das reivindicações 2 a 7.

9. Composição para a detecção in vitro ou in vivo de uma infecção por M. tuberculosis, caracterizada por compreender um polipeptídeo de acordo com uma qualquer das reivindicações 2 a 7 capaz de reagir imunologicamente com anticorpos formados num doente infectado por M. tuberculosis.

10. Composição para a detecção in vitro de uma infecção por M. tuberculosis, caracterizada por compreender uma sequência de nucleótidos que contém pelo menos 9 nucleótidos, proveniente de uma sequência de acordo com a

reivindicação 1, ou uma sequência de nucleótidos que contém pelo menos 9 nucleótidos e que hibrida em condições severas com o ADN de M. tuberculosis e não hibrida nas mesmas condições com o ADN de M. leprae, em que esta sequência é uma sequência de ADN ou de ARN eventualmente marcada.

11. Utilização de um polipeptídeo ou de uma sequência de ácidos aminados de acordo com uma qualquer das reivindicações 2 a 7 como molécula portadora.

12. Hospedeiro celular procariótico ou eucariótico, caracterizado por ser transformado por uma sequência de nucleótidos de acordo com a reivindicação 1 sob condições que permitem a expressão desta sequência e/ou a sua exposição ao nível da membrana do hospedeiro celular e/ou a sua exportação e/ou a sua secreção a partir da referida membrana.

13. Hospedeiro celular de acordo com a reivindicação 12, caracterizado por se tratar de uma micobactéria, por exemplo uma estirpe de M. smegmatis, M. tuberculosis ou M. bovis.

14. Composição imunogénica, caracterizada por compreender um polipeptídeo de acordo com uma qualquer das reivindicações 2 a 7 ou um hospedeiro celular de acordo com uma qualquer das reivindicações 12 ou 13.

Lisboa, 16 de Abril de 2008

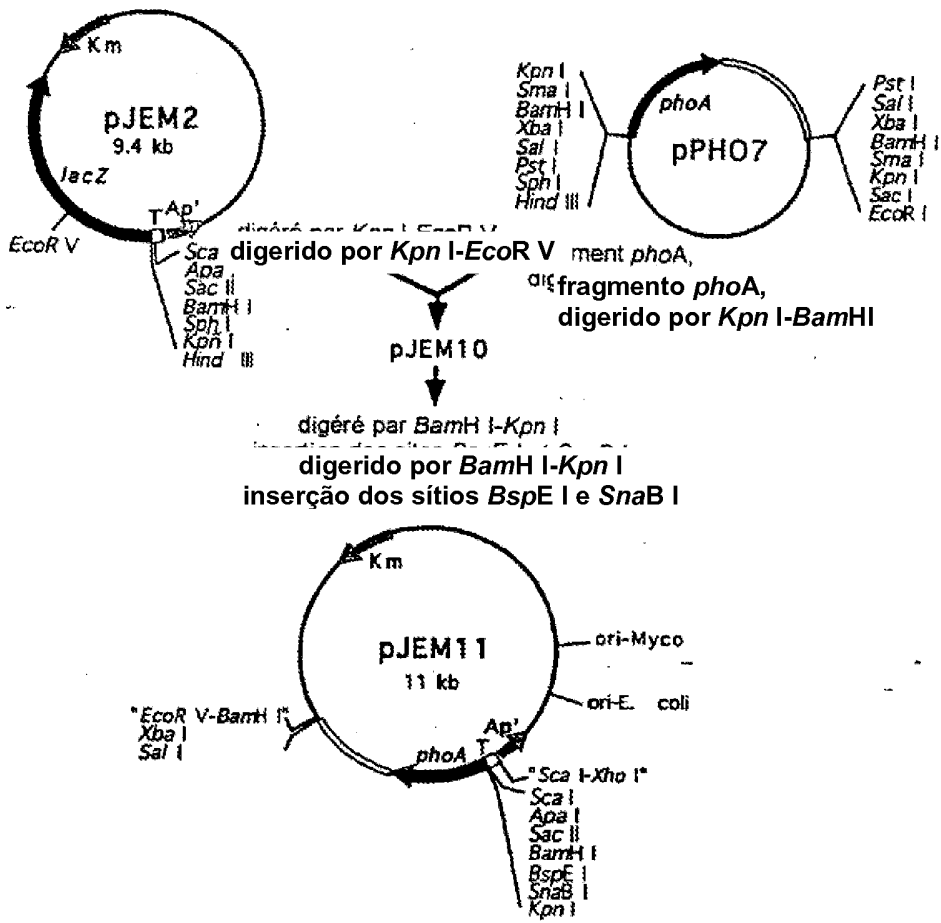


FIGURA 1A

pJEM11
(10 857 pb)

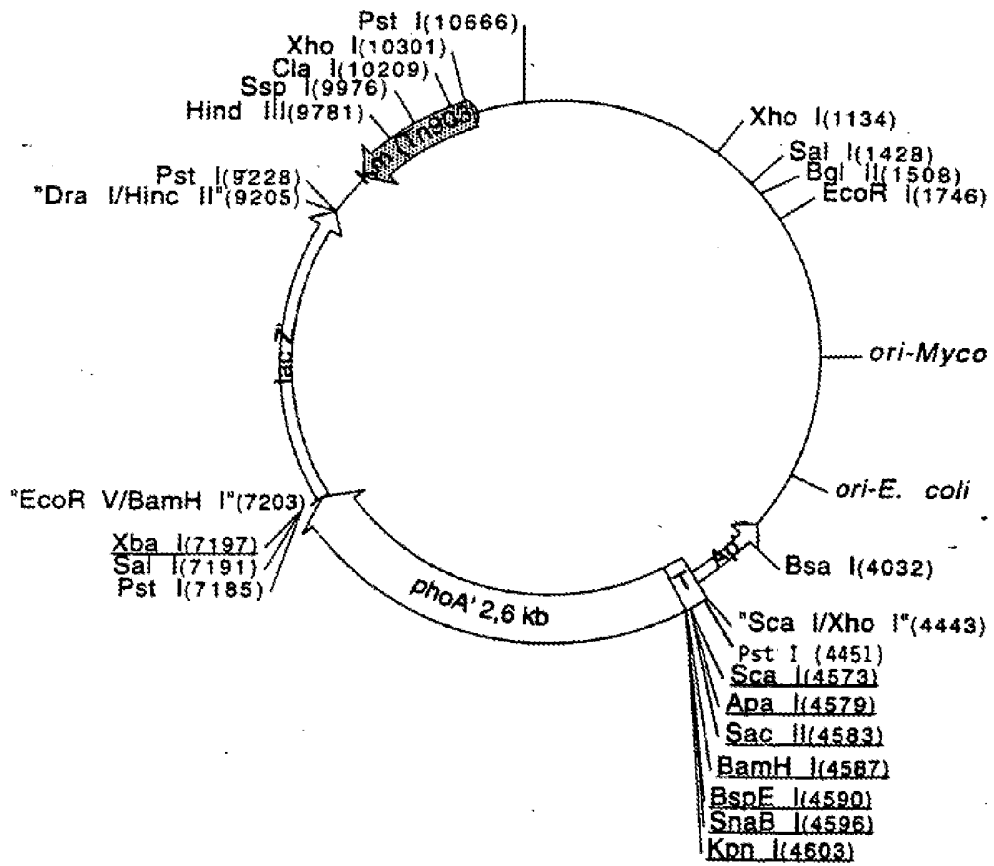


FIGURA 1B

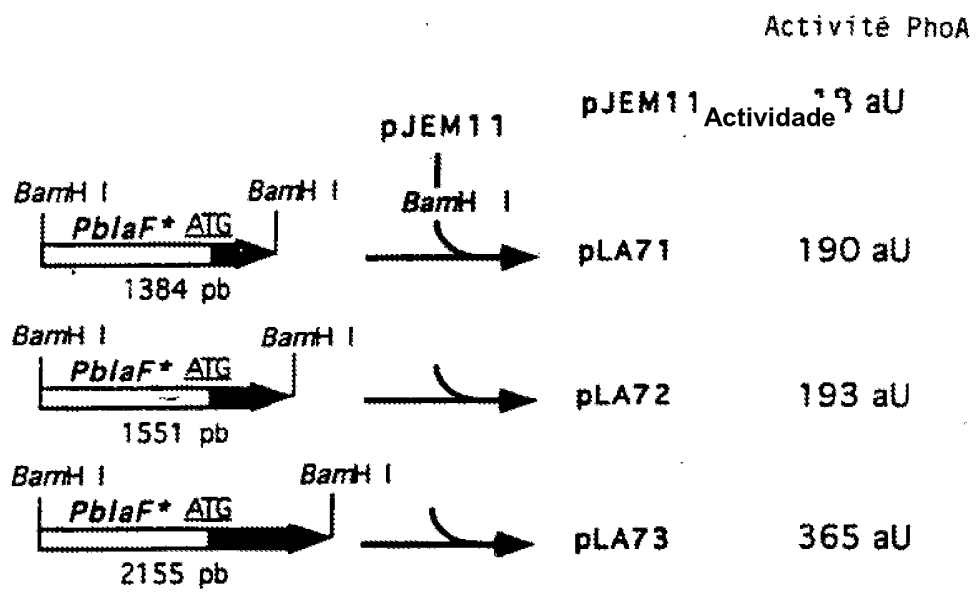


FIGURA 2

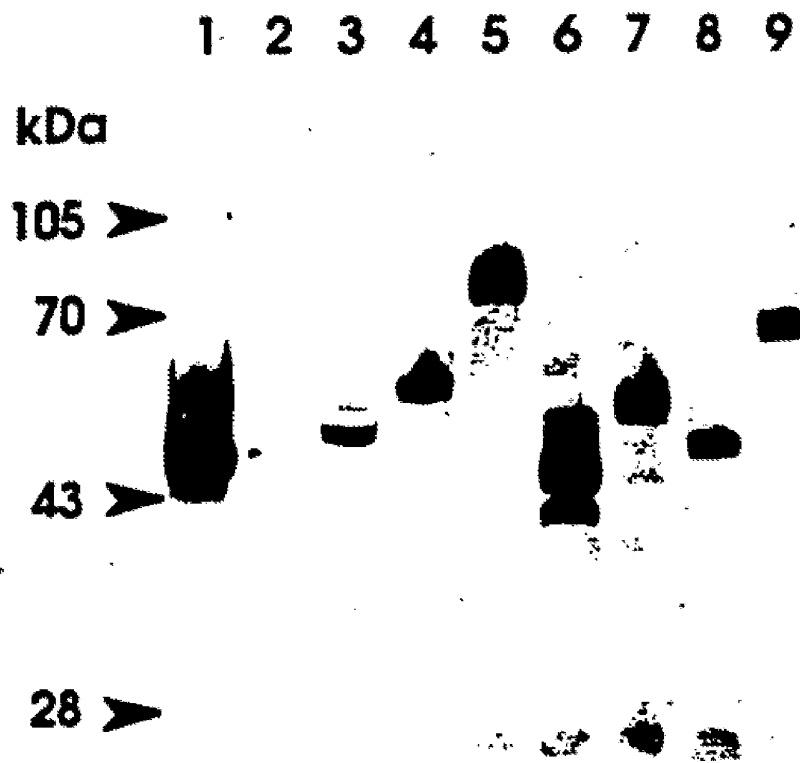


FIGURA 3

A. *M. tuberculosis* 19 kDa (pExp410)

129
 Ser His Tyr Lys Ile
 AGC CAC TAC AAG ATC / C GG ATA CGT ACG
 Bam HI/Sau 3A PhoA reading frame

B. *M. tuberculosis* 28kDa (pExp53)

	I	Start	Pro	Asn	Arg	Ser	Arg	Ser	Lys	Leu	10	
GTTCC	GTG	CCG	AAC	COC	AGC	CGC	AGC	AAG	CTC	TCG	Ser	
	:::	:::	:::	:::	:::A	C::	:::A	T::	:::	:::T		
<i>M. leprae</i>	Start	Pro	Asn	Arg	Arg	Arg	Cys	Lys	Leu	Ser		
	Thr	Ala	Met	Ser	Ala	Val	Ala	Ala	Leu	Ala	Val	
	ACA	GCC	ATG	AGC	GCG	GTC	GCC	GCC	CTG	GCA	GTT	
	:::	:::	:::A	:::	A::	:::	:::	A::	:::A	:::	A:C	
	Thr	Ala	Ile	Ser	Thr	Val	Ala	Thr	Leu	Ala	Ile	
						70						
	Ala	Ser	Pro	-----	Gln	Phe	Gly	Ile				
	GCA	AGT	CCT	-----	CAG	TTC	GGG	ATC / C	GG	ATA	CGT	ACG
	:::C	:::	:::A	-----	:::	:::	:::	Bam HI/Sau 3A	PhoA reading frame			
	Ala	Ser	Pro	-----	Gln	Phe	Gly	Ile				

C. *M. tuberculosis* (pExp59)

	I	Met	Asn	Arg	Ile	Val	Ala			
GTCGAGGAGCCACCG	ATG	AAC	CGG	ATC	GTC	GCC				
	putative RBS									
Pro	Ala	Ala	Ala	Ser	Val	Val	Val	Gly	Leu	
CCC	GCC	GCC	GCA	AGC	GTG	GTG	GTT	GGT	CTG	
Leu	Leu	Ala	Pro	Ala	Ala	Ile				
TTG	CTG	GCG	CCG	GCC	GCG	ATC / C	GG	ATA	CGT	ACG
					Bam HI/Sau 3A	PhoA reading frame				

D. *M. tuberculosis* (pExp421)

	171	Trp	Thr	Ala	Glu	Glu	Asn	Arg	His	Gly	-----	
		TGG	ACC	GCC	GAG	GAG	AAT	CGG	CAC	GGC	-----	
		:::	:::T	:::G	:::A	:::	:::	A:A	:::T	:::T	-----	
<i>R. comm</i>		Trp	Thr	Ala	Glu	Glu	Asn	Arg	His	Gly	-----	
	226	Ser	Phe	Gln	Glu	Leu	Ala	Thr	Arg	Ile	Ser	His
-----		AGT	TTC	CAG	GAA	CTG	GCA	ACC	CGG	ATT	TCG	CAC
-----		TCA	:::	:::	:::	AG:	:::	:::	TTC	:::	:::T	:::T
-----		Ser	Phe	Gln	Glu	Arg	Ala	Thr	Phe	Ile	Ser	His
		Arg	Asn	Thr	-----							
		CGC	AAT	ACC	-----							
		G:G	:::C	:::	-----							
		Gly	Asn	Thr	-----							

FIGURA 4

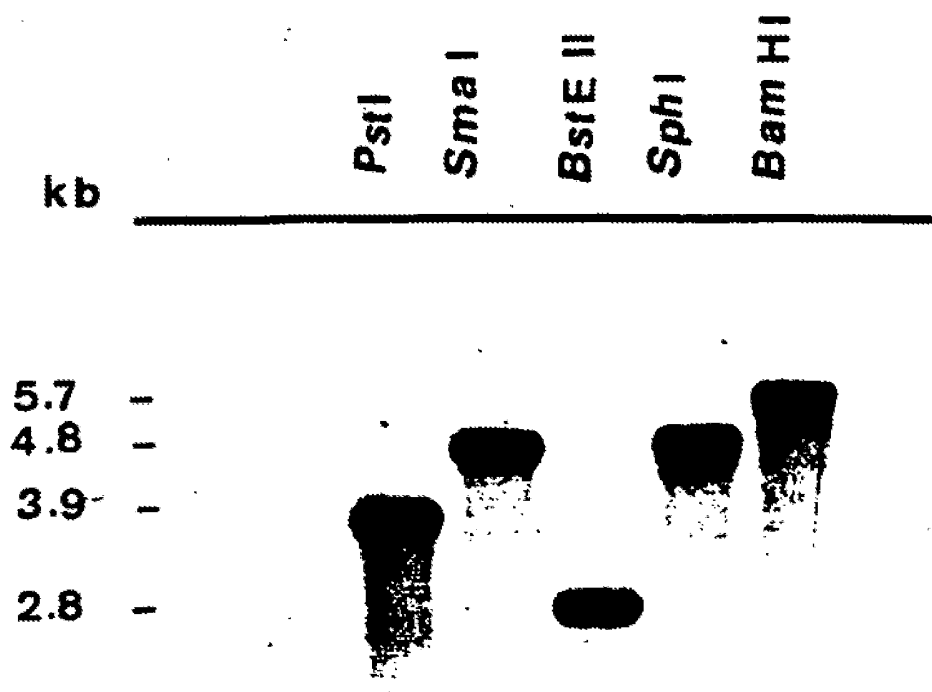


FIGURA 5

1/1 31/11
 GTG CCG AAC CGC AGC CGC AGC AAG CTC TCG ACA GCC ATG AGC GCG GTC GCC GCC CTG GCA
 M P N R S R S K L S T A M S A V A A L A
 61/21 91/31
 GTT GCA AGT CCT TGT GCA TAT TTT CTT GTC TAC GAA TCA ACC GAA ACG ACC GAG CCG CCC
 V A S P C A Y F L V Y E S T E T T E R P
 121/41 151/51
 GAG CAC CAT GAA TTC AAG CAG GCG GCG GTG TTG ACC GAC CTG CCC GGC GAG CTG ATG TCC
 E H H E F K Q A A V L T D L P G E L M S
 181/61 211/71
 GCG CTA TCG CAG GCG TTG TCC CAG TTC GCG ATC AAC ATA CCG CCG GTG CCC AGC CTG ACC
 A L S Q G L S Q F G I N I P P V P S L T
 241/81 271/91
 GGG AGC GGC GAT GCC AGC ACG GGT CTA ACC GGT CCT GGC CTG ACT AGT CCG GGA TTG ACC
 G S G D A S T G L T G P G L T S P G L T
 301/101 331/111
 AGC CCG GGA TTG ACC AGC CCG GGC CTC ACC GAC CCT GCC CTT ACC AGT CCG GGC CTG ACG
 S P G L T S P G L T D P A L T S P G L T
 361/121 391/131
 CCA ACC CTG CCC GGA TCA CTC GCC GCG CCC GGC ACC ACC CTG GCG CCA ACG CCC GGC GTG
 P T L P G S L A A P G T T L A P T P G V
 421/141 451/151
 GGG GCC AAT CCG GCG CTC ACC AAC CCC GCG CTG ACC AGC CCG ACC GGG GCG ACG CCG GGA
 G A N P A L T N P A L T S P T G A T P G
 481/161 511/171
 TTG ACC AGC CCG ACG GGT TTG GAT CCC GCG CTG GGC GGC GCC AAC GAA ATC CCG ATT ACG
 L T S P T G L D P A L G G A N E I P I T
 541/181 571/191
 ACG CCG GTC GGA TTG GAT CCC GGG GCT GAC GGC ACC TAT CCG ATC CTC GGT GAT CCA ACA
 T P V G L D P G A D G T Y P I L G D P T
 601/201 631/211
 CTG GGG ACC ATA CCG AGC AGC CCC GCC ACC ACC TCC ACC GGC GGC GGC GGT CTC GTC AAC
 L G T I P S S P A T T S T G G G G L V N
 661/221 691/231
 GAC GTG ATG CAG GTG GCC AAC GAG TTG GGC GCC AGT CAG GCT ATC GAC CTG CTA AAA GGT
 D V M Q V A N E L G A S Q A I D L L K G
 721/241 751/251
 GTG CTA ATG CCG TCG ATC ATG CAG GCC GTC CAG AAT GGC GGC GTC GCG CCG GCA GCC
 V L M P S I M Q A V Q N G G A V A P A A
 781/261 811/271
 AGC CCG CCG GTC CCG CCC ATC CCC GCG GCC GCG GCG GTG CCA CCG ACG GAC CCA ATC ACC
 S P P V P P I P A A A A V P P T D P I T
 841/281
 GTG CCG GTC GCC TAA
 V P V A *

Sequência de nucleótidos e sequência de ácidos aminados deduzida
 do produto potencial do gene IRSA de *M. Tuberculosis*

FIG. 6A

1/1
 GTG CCG AAC CGA CGC CGA CGC AAG CTC TCG ACA GCC ATG AGC GCG GTC GCC GCC CTG GCA
 M P N R R R R K L S T A M S A V A A L A
 61/21
 GTT GCA AGT CCT TGT GCA TAT TTT CTT GTC TAC GAA TCA ACC GAA ACG ACC GAG CGG CCC
 V A S P C A Y F L V Y E S T E T T E R P
 121/41
 GAG CAC CAT GAA TTC AAG CAG GCG GCG GTG TTG ACC GAC CTG CCC GGC GAG CTG ATG TCC
 E H H E F K Q A A V L T D L P G E L M S
 181/61
 GCG CTA TCG CAG GGG TTG TCC CAG TTC GCG ATC AAC ATA CCG CCG GTG CCC AGC CTG ACC
 A L S Q G L S Q F G I N I P P V P S L T
 241/81
 GGG AGC GGC GAT GCC AGC ACG GGT CTA ACC GGC CTG ACT AGT CCG GGA TTG ACC
 G S G D A S T G L T G P G L T S P G L T
 301/101
 AGC CCG GGA TTG ACC AGC CCG GGC CTC ACC GAC CCT GCC CTT ACC AGT CCG GGC CTG ACC
 S P G L T S P G L T D P A L T S P G L T
 361/121
 CCA ACC CTG CCC GGA TCA CTC GCC GCG CCC GGC ACC ACC CTG GCG CCA ACG CCC GGC GTG
 P T L P G S L A A P G T T L A P T P G V
 421/141
 GGG GCC AAT CCG GCG CTC ACC AAC CCC GCG CTG ACC AGC CCG ACC GGG GCG ACG CCG GGA
 G A N P A L T N P A L T S P T G A T P G
 481/161
 TTG ACC AGC CCG ACG GGT TTG GAT CCC GCG CTG GGC GGC GCC AAC GAA ATC CCG ATT ACG
 L T S P T G L D P A L G G A N E I P I T
 541/181
 ACG CCG GTC GGA TTG GAT CCC GGG GGT GAC GGC ACC TAT CCG ATC CTC GGT GAT CCA ACA
 T P V G L D P G A D G T Y P I L G D P T
 601/201
 CTG GGG ACC ATA CCG AGC AGC CCC GCC ACC ACC TCC ACC GCC GGC GGC GGT CTC GTC AAC
 L G T I P S S P A T T S T G G G L V N
 661/221
 GAC GTG ATG CAG GTG GCC AAC GAG TTG GGC GCC AGT CAG GGT ATC GAC CTG CTA AAA GGT
 D V M Q V A N E L G A S Q A I D L L K L G
 721/241
 GTG CTA ATG CCG TCG ATC ATG CAG GCC GTC CAG AAT GGC GGC GCG GTC GCG CCG GCA GCC
 V L M P S I M Q A V Q N G G A V A P A A
 781/261
 AGC CCG CCG GTC CCG CCC ATC CCC GCG GCC GCG GCG GTG CCA CCG ACG GAC CCA ATC ACC
 S P P V P P I P A A A A V P P T D P I T
 841/281
 GTG CCG GTC GCC TAA
 V P V A *

Sequência de nucleótidos e sequência de ácidos aminados deduzida
 do produto potencial do gene IRSA de *M. Tuberculosis*

FIG. 6B

Fig. 7 Sequências nucleotídicas de flanco do gene IRSA de
M. tuberculosis

A - Sequência nucleotídica a montante:

5'-CGGCTTCGGAATAGGCATTGCCCCGATGTGCGGGCGCCGCTCGAGGACGAGCACGCGCTTGTC
GAGTTGGGTGGACACGGCTCGGCAATCGTCAGGCCGAAGAATCCTGAGCCGACGACGAAAAGGTCA
AAACGAGCGGTCATCGGTTGCATAGGGTAACCGACCTTGCTGGCAAACCCGATTGGCAGCTCGTG
GCGGTCATGGCCCGAACGGGTTTCACCGCAGGTGCGCATGGCCGACCAGTGTGGTTGGCCGGAGGTC
GTTTGGTTCGCGATTGCCTCACGATTCGATATAACCACTCTAGTCACATCAACCACACTCGTACCATC
GAGCGTGTGGGTTTCATGCCATGCACTCGCGACCGCGGGAGCCGGCGAACC CGCGCCACACATAATC
CAGATTGAGGAGACTTCC **GTG COG AAC 3'**

SD Met Pro Asn ... géne IRSA

SD "Shine-Delgarno" sítio putativo de ligação ao
ribossoma

Met Metionina de início do gene IRSA

B - Sequência nucleotídica a jusante

Pro Val Stop

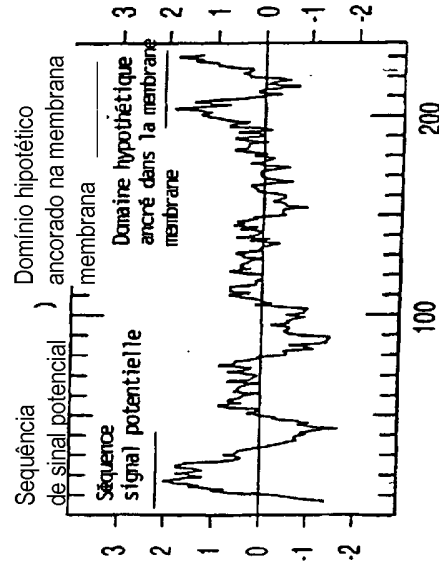
géne IRSA .5'- GTC GCC TAA GCCCGGGTCGGCCGAAAACGCACCCGCGGCCAAGGCG
TCGGTCATTGCTTCGGCCCGTGCAATTAATTCGCCTAAGGGTCGGCTAGGTGTTCTCGAGAGTTTT
ATCGCACCGATTCCGTGTCGTCTCATTAATACCAATAGAAAACACACGTAACATCAGCTGGTGCCGT
CCCGCACCCGCGCGCCGACGACGCTGCTCACCGCGATGGCAGCGACCGTTCGTTCATCGTTCGCGTGGAT
AGCGAATCGTCCACCCGCCAGCTCCCAT 3'

FIGURA 7

Perfil de hidrofília segundo Kyte & Doolittle

Profil d'hydrophilie selon Kyte & Doolittle

Mycobacterium leprae



Mycobacterium tuberculosis

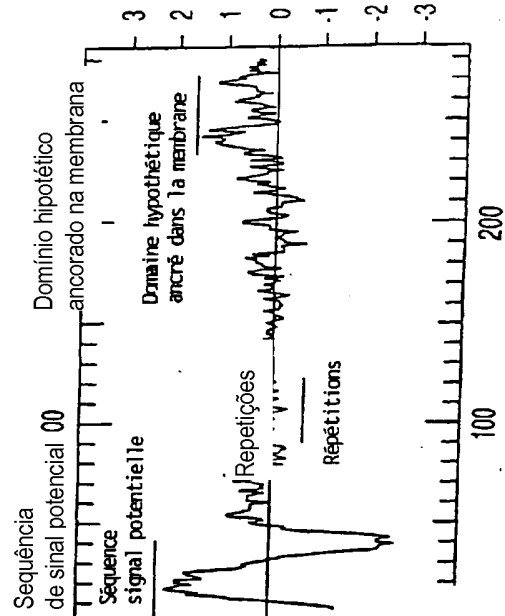


FIGURA 9

<i>M. tuberculosis</i>	1	GTCCGAAACCGCAGCCGCGAGCAAGCTCTCGACAGCCATGAGCCGGTCCG	50
<i>M. Leprae</i>	1	GTCCGAAACCGCAGCCGCGAGCTTCGACAGCCATGAGCCGGTCCG	50
	51	CGCCCTGGCAGTTGCAAGTTCCTTGTGCATATTTCTGTCTACGAATCAA	100
	51	CACCCTAGCAATGCCCAATCCATGCGCATATTTCCCTGTTTACGAACCGA	100
	101	CCGAAACGACCGAGC...GGCCCGAGCACCATGAATTCAAAGCGGGCG	147
	101	CCGGAGCGCCAAACCCGGCGCAACACATATGAATTCAAACAGCAGCA	150
	148	GTGTTGACCGACTGCCCCGGAGCTGATGCCCGGCTATCGCAGGGGTT	197
	151	TCGATAGCCGACTGCCCCGGAGAGTGTGGACGGGATCTCGCAGGGACT	200
	198	GTCCCAGTTCGGGATCAACATACCCCGGTGCCCGAGCTGACCCGGGAGCG	247
	201	GTCCGAGTTCGGCATCAACCTACCCGGGTGCTTCGCTAACCTGGCCACCG	250
	248	GCGATGCCAGCACGGGTCTAACCGGTCTGGCCTGACTAGTCCGGGATTG	297
	251	ATGAT.....CCAGGTAATGGCCTGAGAACTCCCGGTTTG	285
	298	ACCAGCCGGGATGACCAAGCCCGGCTCACCGACCTGCCCTTACCAG	347
	286	ACCAGCCCGATCTGACAAATCAGGAGCTAGGACACCTGTGCTCACCGG	335
	348	TCCGGGCTGACGCCAACCTGCCCGGATCACTCGCCGCGCCCGGCACCA	397
	336	GCCGGCACGGGACTGACACCACTGTGACAGGCAGCCCGATATGTACCG	385
	398	CCCTGGCCCAACCGCCGGGTGGGGCCAAATCCGGGCTCACCAACCCC	447
	386	CAC.....	388

ALINHAMENTO DE SEQUÊNCIAS DE NUCLEÓTIDOS
 FIGURA 10A(1)

<i>M. Tuberculosis</i>	448	GGCTGACCAGCCGACCGGGGGCCAGCCGGGATTGACCAGCCCGACCGGG	497
	389CGGA	392
<i>M. Leprae</i>	498	TTTGGATCCCGCTGGCGGGCCAA CGAAATCCGATACGACGCCGG	547
	393	CCTGAACCTGGTGGCACCTGCCCCAGCGAGGTACCGATCACCACACCAA	442
	548	TCGGATTGGATCCCGGGCTGACGGCACCTATCCGATCCTCGGTATCCA	597
	443	TTTCATTTGGACCCGGGCACCGACGGCACCTATCCGATCCTCGGGATCC.	491
	598	ACACTGGGGACCATACCGAGCAGCCCGCCGCCACCCCTCCACCGGGCGGG	647
	492CTCCACGTTGGGGGTACATCACCGATCAGTACCAGCAGCGG	533
	648	CGGTCTCGTCAACGACGTGATGCAGGTGGCCAA CGAGTTGGCGCCAGTC	697
	534	TGAGCTTGTAATGACCTGCTAAAAGTTGCCAACCGAGTTGGCGCCAGCC	583
	698	AGGCTATCGACCTGTAAAAGGTGCTAATGCCGTGGATCATGCAGGCC	747
	584	AGGTATGGACCTAATCAAGGGTGTGGTATGCCACCGGTTCATGCAGGCC	633
	748	GTCCAGAAATGGCGCGGGTCCGCCCGGCGAGCCAGCCCGCGGTCCCGCC	797
	634	GTCCAGAACGGCAACGTAGCCCGGTGACTTGTGGGCTCAGTAAACGGCGGC	683
	798	CATCCCGGGCGCGGGTGCACCGGACCGAACCCAAATCACCGTCCCGG	847
	684	CGCGATATCACTGATTCCTGTACGGTAG.....	711

ALINHAMENTO DE SEQUÊNCIAS DE NUCLEÓTIDOS

FIGURA 10A (2)

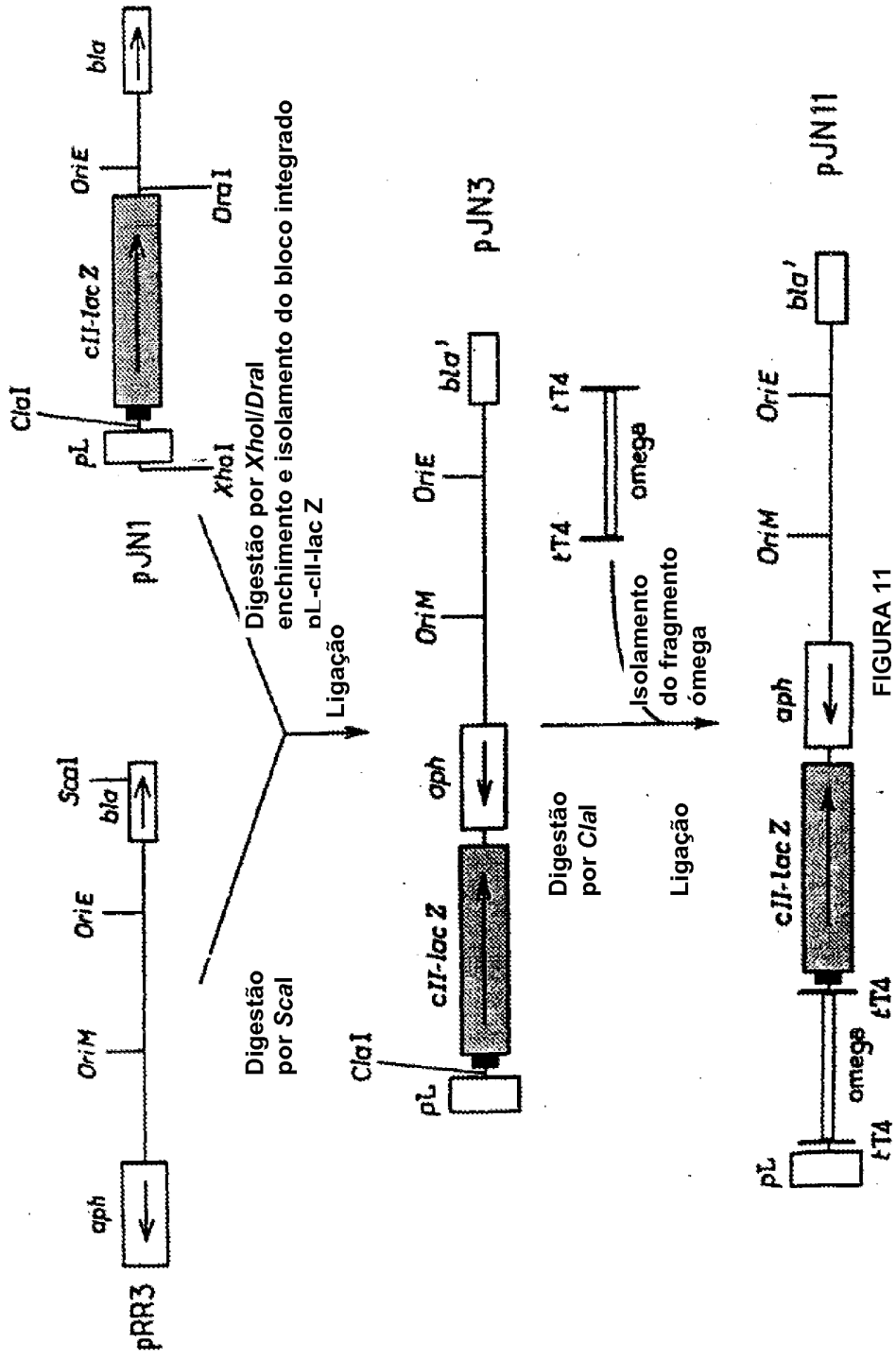
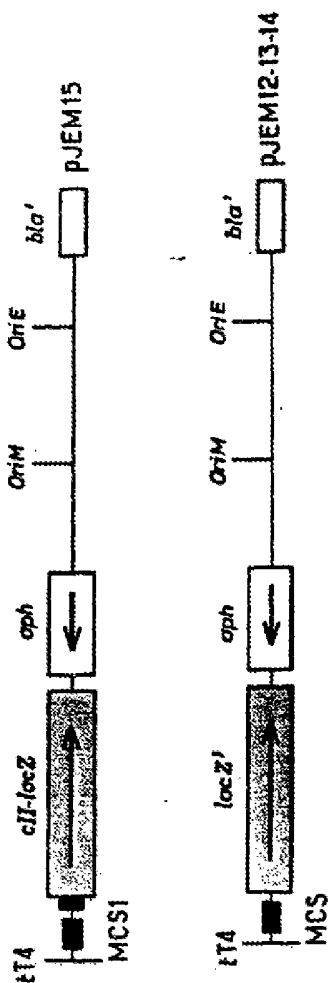


FIGURA 11

A



ATGACCTTAATAGATTATTAATTACTAATTAATTGGGGACCTAGAGGTCCTTTTAAAAAATTTTTCACAAAACGGTTTACAAAGCATAAAGCT
 AGTACTGGGCCC GCGGATCCGATCGGATCGGATCGAAGCTTGATCCGATAACACAGGAACAGATCTATGGTT CGT GCA AAC AAA CGC AAC GAG
 ScaI Apol SacII BamHI SphI KpnI HindIII* Met Val Arg Ala Asn Lys Arg Asn Glu

PJEM15

GCTCTA CGA ATC GGA AGCTTC GAT CCC
 Ala Leu Arg Ile Gly Ser Phe ***

ATGACCTTAATAGATTATTAATTACTAATTAATTGGGGACCTAGAGGTCCTTTTAAAAAATTTTTCACAAAACGGTTTACAAAGCATAAAGCT

AGTACTGGGCCC GCGGATCCGATCGGATCGGATCGAAGCTTCGATCCC
 ScaI Apol SacII BamHI SmaBI KpnI HindIII* GAT CCC
 AspEI GC GAT CCC ***

PJEM12

PJEM13

PJEM14

B

FIGURA 12