

República Federativa do Brasil  
Ministério do Desenvolvimento, Indústria  
e do Comércio Exterior  
Instituto Nacional da Propriedade Industrial.

(21) **PI0611347-8 A2**

(22) Data de Depósito: 27/04/2006  
(43) Data da Publicação: 31/08/2010  
(RPI 2069)



(51) *Int.Cl.:*  
C07K 14/35  
A61K 39/04

(54) Título: **COMPOSIÇÃO FARMACÊUTICA, USOS DE UMA PROTEÍNA DE FUSÃO MTB72F OU UM FRAGMENTO IMUNOGÊNICO DA MESMA, E, DE UM ÁCIDO NUCLÉICO**

(30) Prioridade Unionista: 29/04/2005 US 60/676549, 27/02/2006 US 60/777017, 27/02/2006 US 60/777017, 29/04/2005 US 60/676549

(73) Titular(es): GLAXOSMITHLINE BIOLOGICALS S.A. , INFECTIOUS DISEASE RESEARCH INSTITUTE (IDRI)

(72) Inventor(es): RHEA COLER, STEVEN REED, YVES LOBET

(74) Procurador(es): Momsen, Leonardos & CIA.

(86) Pedido Internacional: PCT EP2006004319 de 27/04/2006

(87) Publicação Internacional: WO 2006/117240 de 09/11/2006

(57) Resumo: COMPOSIÇÃO FARMACÊUTICA, USOS DE UMA PROTEÍNA DE FUSÃO Mtb72f OU UM FRAGMENTO IMUNOGÊNICO D MÁ, E, DE UM ÁCIDO NUCLÉICO A presente invenção refere-se a métodos d prevenir reativação de infecções ativas e latentes com M tuberculosis por meio de administração de uma composição farmacêutica compreendendo um ácido nueleico codificando para uma proteína de fusão Mtb72f, ou uma roteina de fusão Mtb72f ou um fragmento imunogênico do mesmo, por exemplo, em conjunto com um adjuvante. O ácido nucleico Mtb72f ou proteína de fusão pode ser administrado com um ou mais agentes quimioterápicos eficazes contra uma infecção com M tuberculosis. Os métodos também pro orcionam o encurtamento do período de um regime quimioterápico contra uma infecção com M tuberculosis.

“COMPOSIÇÃO FARMACÊUTICA, USOS DE UMA PROTEÍNA DE FUSÃO Mtb72f OU UM FRAGMENTO IMUNOGÊNICO DA MESMA, E, DE UM ÁCIDO NUCLÉICO”

#### CAMPO DA INVENÇÃO

5 A presente invenção refere-se a métodos de prevenir ou tratar reativação de uma infecção com *M. tuberculosis* em um mamífero e a métodos de encurtar o decurso da quimioterapia contra uma infecção com *M. tuberculosis*.

#### ANTERIORIDADES DA INVENÇÃO

10 Tuberculose é uma doença infecciosa crônica causado por infecção com *M. tuberculosis* e outras espécies de *Mycobacterium*. Ela é uma doença importante em países em desenvolvimento, e também é um problema crescente em áreas desenvolvidas do mundo, com cerca de 8 milhões novos casos a cada ano. Embora a infecção possa ser assintomática durante um período considerável, a doença é manifestada mais comumente como uma inflamação aguda dos pulmões, resultando em febre e numa tosse não-  
15 produtiva. Se não tratada, resultarão tipicamente sérias complicações e morte.

Embora a tuberculose possa ser geralmente controlada empregando-se terapia estendida com antibióticos, tal tratamento não é  
20 suficiente para prevenir a disseminação da doença. Indivíduos infectados podem ser assintomáticos, mas contagiosos, durante algum tempo. Adicionalmente, embora a adesão com o regime de tratamento seja crítica, o comportamento do paciente é difícil de monitorar. Alguns pacientes não completam o curso de tratamento, o que pode levar a tratamento ineficaz e ao  
25 desenvolvimento de resistência à droga. Mesmo que o curso completo do tratamento seja completado, a infecção com *M. tuberculosis* não é erradicada do indivíduo infectado, mas permanece como uma infecção latente que pode ser reativada.

Para controlar a disseminação da tuberculose, vacinação eficaz e diagnóstico precoce preciso da doença são da maior importância. Correntemente, vacinação com bactérias vivas é o método mais eficiente para induzir imunidade protetora. O *Mycobacterium* mais comum empregado para tal fim é o Bacillus Calmette-Guerin (BCG), uma cepa avirulenta de *M. bovis*. No entanto, a segurança e eficácia do BCG é uma fonte de controvérsia, e alguns países, como os Estados Unidos, não vacinam o público em geral com este agente.

O diagnóstico da tuberculose é obtido comumente usando-se um teste de pele, que envolve exposição intradérmica à tuberculina PPD (derivado purificado de proteína). Respostas de células T específicas para antígeno resultam em endurecimento mensurável no sítio de injeção em torno de 48 a 72 horas após injeção, o que indica exposição a antígenos micobacterianos. Sensibilidade e especificidade tem representado, contudo, um problema com este teste é que indivíduos vacinados com BCG não podem ser diferenciados de indivíduos infectados.

Embora se tenha mostrado que macrófagos atuam como os principais efetoras da imunidade de *Mycobacterium*, células T são os indutores predominantes de referida imunidade. O papel essencial de células T na proteção contra infecção com *Mycobacterium* é ilustrado pela ocorrência freqüente de infecção com *Mycobacterium* em pacientes com AIDS, devido à depleção de células T CD4<sup>+</sup> associadas com infecção com vírus da imunodeficiência humana (HIV). Mostrou-se que células T CD4<sup>+</sup> reativas a *Mycobacterium* são produtores potentes de  $\gamma$ -interferon (IFN- $\gamma$ ), e que, por sua vez, disparam os efeitos anti-micobacterianos de macrófagos em camundongos. Embora o papel do IFN- $\gamma$  em humanos seja menos claro, estudos mostraram que 1,25-di-hidróxi-vitamina D3, seja sozinha ou em combinação com IFN- $\gamma$  ou fator de necrose de tumor alfa, ativa macrófagos humanos para inibir infecção com *M. tuberculosis*. Adicionalmente, é de

conhecimento geral que o IFN- $\gamma$  estimula macrófagos humanos a produzir, 25-  
 di-hidróxi-vitamina D3. De forma análoga, mostrou-se que a interleucina-12  
 (IL-12) desempenha um papel na estimulação da resistência a infecção com  
*M. tuberculosis*. Para uma revisão da imunologia da infecção com *M.*  
 5 *tuberculosis*, ver Chan & Kaufmann, *Tuberculosis: Pathogenesis, Protection*  
*and Control* (Bloom ed., 1994), *Tuberculosis* (2<sup>a</sup> ed., Rom e Garay, eds.,  
 2003), e *Harrison 's Principles of Internal Medicine*, capítulo 150, pp. 953-  
 966 (16<sup>a</sup> ed., Braunwald, *et al*, eds., 2005).

10 Permanece uma necessidade de estratégias de tratamento  
 eficazes para prevenir a reativação de infecções com *Mycobacterium*  
*tuberculosis*, tanto de infecções ativas como de infecções latentes. Esta  
 invenção atende esta e outras necessidades.

#### Descrição das seqüências listadas

- 15 SEQ ID No:1: Mtb72f com marcador 6 His N-terminal (DNA)  
 SEQ ID No:2: Mtb72f com marcador 6 His N-terminal  
 (proteína)  
 SEQ ID No:3: M72 (variante de Mtb72f) com inserção de 2  
 His N-terminal (DNA)  
 20 SEQ ID No:4: M72 (variante de Mtb72f) com inserção de 2-  
 His N-terminal (proteína)  
 SEQ ID No:5: Mtb72f sem inserção de His N-terminal (DNA)  
 SEQ ID No:6: Mtb72f sem inserção de His N-terminal  
 (proteína)

#### Breve resumo da invenção

25 A presente invenção proporciona composições farmacêuticas  
 compreendendo uma proteína de fusão Mtb72f ou um fragmento imunogênico  
 da mesma de uma espécie de um *Mycobacterium* do complexo de tuberculose,  
 por exemplo, em conjunto com um ou mais adjuvantes, incluindo AS01B e  
 AS02A.

A presente invenção baseia-se em parte na descoberta dos inventores de que a administração de uma proteína de fusão Mtb72f ou fragmento imunogênico da mesma, p. ex., em conjunto com um ou mais adjuvantes ou um ácido nucleico que codifica uma proteína de fusão Mtb72f ou fragmento imunogênico da mesma, pode prevenir ou tratar reativação de uma infecção ativa ou inativa com *M. tuberculosis*. Em uma concretização preferida, uma proteína de fusão Mtb72f ou ácido nucleico é administrada com um ou mais agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis*.

10 Em um aspecto, as composições são empregadas em métodos para prevenir ou tratar reativação de tuberculose em um sujeito, sendo que o método compreende a etapa de administrar a um mamífero já infectado com *Mycobacterium tuberculosis* uma quantidade imunologicamente eficaz de uma composição farmacêutica compreendendo uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose e um adjuvante, sendo que a proteína de fusão Mtb72f induz uma resposta imunológica contra *M. tuberculosis*, prevenindo ou tratando com isto a reativação da tuberculose.

20 Em outro aspecto, as composições são empregadas em métodos para prevenir reativação da tuberculose em um sujeito, sendo que o método compreende a etapa de administrar a um mamífero já infectado com *Mycobacterium tuberculosis* uma quantidade imunologicamente eficaz de uma composição farmacêutica compreendendo um ácido nucleico que codifica uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose, sendo que a proteína de fusão Mtb72f expressa induz uma resposta imunológica contra *M. tuberculosis*, prevenindo ou tratando com isto a reativação da tuberculose.

25 Em outro aspecto, as composições são empregadas em

métodos para reduzir o decurso da quimioterapia contra uma infecção com *M. tuberculosis*, sendo que o método compreende administrar a um mamífero já infectado com *Mycobacterium tuberculosis* um ou mais agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis* e uma quantidade imunologicamente eficaz de uma composição farmacêutica compreendendo uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose e um adjuvante, sendo que referida proteína de fusão Mtb72f ou fragmento imunogênico da mesma induz uma resposta imunológica contra *M. tuberculosis*, permitindo, com isto, reduzir o decurso da quimioterapia contra uma infecção com *M. tuberculosis*. Encurtando o decurso da quimioterapia contra uma infecção com *M. tuberculosis*, os presentes métodos também são eficazes para incrementar a adesão de um indivíduo que está sendo tratado para uma infecção com *M. tuberculosis* para completar um curso inteiro de tratamento.

#### Breve descrição dos desenhos

Figura 1 mostra uma representação gráfica do modelo de reativação da *M. tuberculosis* em camundongos Swiss Webster (SWR/J). A figura mostra momentos determinados para infecção, tratamento com quimioterapia (50 mg de rifampina/85 mg de isoniazida por litro de água potável), imunizações e enumeração de carga bacteriana/unidades formadoras de colônia (CFU).

Figura 2 mostra respostas imunológicas de anticorpos de IgG1 e IgG2a em camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Camundongos foram deixados não-tratados, tratados com quimioterapia (50 mg de rifampina/85 mg de isoniazida por litro de água potável) ou tratados com quimioterapia e imunizados três vezes intra-muscularmente com 8 µg por dose de Mtb72f formulado sem adjuvante. Dez dias após a última imunização, os

camundongos foram sangrados e os soros testados quanto a resposta de anticorpos anti-Mtb72f para ambos os isótopos, IgG1 (vermelho) e IgG2a (preto) por meio de ELISA.

Figura 3 mostra respostas imunológicas de anticorpos de IgG1 e IgG2a em camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Camundongos foram deixados não-tratados, tratados com quimioterapia (50 mg de rifampina/85 mg de isoniazida por litro de água potável) ou tratados com quimioterapia e imunizados três vezes intra-muscularmente com 8 µg por dose de Mtb72f formulado com o adjuvante ASO1B. Dez dias após a última imunização, os camundongos foram sangrados e os soros testados quanto à resposta de anticorpo anti-Mtb72f para ambos os isótopos, IgG1 (vermelho) e IgG2a (preto), por meio de ELISA.

Figura 4 mostra respostas de interferon-gama (IFN-γ) em camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Células de baço foram obtidas de camundongos em diversos momentos determinados e estimuladas *in vitro* durante três dias com 10 µg/ml de rMtb72f ou os componentes (Mtb32c e Mtb39) como indicado. Como controles, culturas de esplenócitos também são estimuladas com PPD (3 µg/ml), lisado de BCG (10 µg/ml), conA (3 µg/ml) ou meio apenas. Infecção com IFN-γ foi medida subsequentemente por meio de ELISA.

Figura 5 mostra respostas de IFN-γ em camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Células de baço foram obtidas de camundongos em diversos momentos determinados e estimuladas *in vitro* durante três dias com 10 µg/ml de rMtb72f ou os componentes (Mtb32c e Mtb39) como indicado. Como controles, culturas de esplenócitos também foram estimuladas com PPD (3 µg/ml), lisado de BCG (10 µg/ml), conA (3 µg/ml) ou meio apenas. A

produção de IFN- $\gamma$  foi medida subsequentemente por meio de ELISA.

Figura 6 mostra respostas de células T CD4<sup>+</sup> e citocina IFN- $\gamma$  em camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Células de baço foram obtidas de camundongos em diversos momentos determinados e estimuladas *in vitro* de um dia para o outro com 10  $\mu$ g/ml de rMtb72f. As células foram então manchadas para CD4 e IFN- $\gamma$ . Como um controle, culturas de esplenócitos também foram estimuladas com meio apenas. A produção de IFN- $\gamma$ <sup>+</sup> específica para células T CD4<sup>+</sup> foi medida subsequentemente por meio de manchamento de citocina intracelular (ICS).

Figura 7 mostra um sumário tabulado dos valores de produção de IFN- $\gamma$ <sup>+</sup> específica para células T CD4<sup>+</sup> e CD8<sup>+</sup> no dia 120 após infecção com Mtb. Células de baço foram obtidas de grupos de camundongos deixados não-tratados, tratados com combinação quimioterapia durante 30, 60 ou 90 dias, ou tratados com quimioterapia de combinação como um adjunto à vacina de Mtb72f. Esplenócitos foram estimulados *in vitro* de um dia para o outro com 10  $\mu$ g/ml de rMtb72f. As células foram então manchadas para CD4, CD8 ou IFN- $\gamma$ . Como um controle, culturas de esplenócitos também foram estimuladas com meio apenas. Produção de IFN- $\gamma$ <sup>+</sup> específica para células T CD4<sup>+</sup> e CD8<sup>+</sup> foi medida subsequentemente por meio de manchamento de citocina intracelular.

Figura 8 mostra a sobrevivência de camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Camundongos foram infectados via aerossol com 50-100 CFU de MtbH37Rv, e iniciou-se quimioterapia (50 mg de rifampina/85 mg de isoniazida por litro de água potável) em um subconjunto de camundongos trinta dias mais tarde. Quimioterapia foi prosseguida durante 60 dias. Metade daqueles camundongos que receberam quimioterapia foram imunizados três vezes intra-muscularmente com 8  $\mu$ g por dose de Mtb72f



formulado com o adjuvante AS10B.

Figura 9 mostra a sobrevivência de camundongos SWR/J infectados com *M. tuberculosis* tratados com quimioterapia e, depois, imunizados com Mtb72f. Camundongos foram infectados via aerossol com 50-100 CFU de MtbH37Rv, e iniciou-se quimioterapia (50 mg de rifampina/85 mg de isoniazida por litro de água potável) em um subconjunto de camundongos trinta dias mais tarde. Quimioterapia foi prosseguida durante 30, 60 ou 90 dias em subconjuntos separados de camundongos. Metade destes camundongos que receberam quimioterapia foram imunizados três vezes intra-muscularmente com 8 µg por dose de Mtb72f formulado com o adjuvante AS10B.

#### Descrição detalhada de concretizações específicas

A presente invenção refere-se a composições compreendendo proteínas de fusão ou ácidos nucleicos de Mtb72f e um adjuvante úteis para tratar, prevenir, ou retardar a reativação de uma infecção ativa ou inativa (i.e., latente) com *Mycobacterium*, e métodos para seu uso. Mais especificamente, as composições da presente invenção compreendem polipeptídeos de fusão Mtb72f ou fragmentos imunogênicos dos mesmos ou ácidos nucleicos que codificam polipeptídeos de fusão Mtb72f ou fragmentos imunogênicos dos mesmos apresentando componentes de uma espécie de *Mycobacterium* do complexo de tuberculose, p. ex., uma espécie, como *M. tuberculosis*, *M. bovis*, ou *M. africanum*, ou uma espécie de *Mycobacterium* que é ambiental ou oportunística e que causa infecções oportunísticas, como infecções do pulmão em hospedeiros imunocomprometidos (p. ex., pacientes com AIDS), p. ex., BCG, *M. avium*, *M. intracellular*, *M. celatum*, *M. genavense*, *M. haemophilum*, *M. kansasii*, *M. simiae*, *M. vaccae*, *M. fortuitum*, e *M. scrofulaceum* (ver, p. ex., Harrison's Principles of Internal Medicine, capítulo 150, pp. 953-966 (16<sup>a</sup> ed., Braunwald, *et al.*, eds., 2005). Os inventores do presente pedido descobriram, com surpresa, que composições compreendendo

polipeptídeos de fusão Mtb72f ou ácidos nucleicos que codificam polipeptídeos de fusão Mtb72f, ou fragmentos imunogênicos dos mesmos, situam-se úteis para tratar, prevenir ou retardar reativação de uma infecção com *M. tuberculosis*. Em uma concretização preferida, um polipeptídeo de fusão Mtb72f ou ácido nucleico é administrado com um ou mais agentes quimioterápicos. Portanto, estas composições, polipeptídeos, e os ácidos nucleicos que codificam os mesmos são úteis para eliciar uma resposta imunológica em mamíferos que é protetora contra reativação de sintomas de doença.

Os polipeptídeos de fusão e ácidos nucleicos de Mtb72f da presente invenção podem compreender adicionalmente outros componentes projetados para incrementar sua antigenicidade ou para aperfeiçoar estes antígenos em outros aspectos. Por exemplo, o isolamento aperfeiçoado dos antígenos de polipeptídeo de fusão pode ser facilitado por meio da adição de uma extensão de radicais histidina no sentido de uma extremidade do antígeno. As composições, polipeptídeos, e ácidos nucleicos da invenção podem compreender cópias adicionais de antígenos, ou polipeptídeos heterólogos adicionais de *Mycobacterium* sp., como antígeno MTB8.4, antígeno MTB9.8, antígeno MTB9.9, antígeno MTB40, antígeno MTB41, antígeno ESAT-6, antígeno de complexo MTB85, antígeno  $\alpha$ -cristalino, ou antígeno NS1.

Alternativamente ou adicionalmente às composições, polipeptídeos, e ácidos nucleicos da invenção podem compreender cópias adicionais de outros antígenos de *Mycobacterium* sp., como Ag85B ou MTCC#2, as composições, polipeptídeos, e ácidos nucleicos da invenção também podem compreender polipeptídeos adicionais de outras fontes. Por exemplo, as composições e proteínas de fusão da invenção podem incluir polipeptídeos ou ácidos nucleicos que codificam polipeptídeos, sendo que o polipeptídeo incrementa a expressão do antígeno, p. ex., NS1, uma proteína

de vírus de influenza (ver, p. ex. WO99/40188 e WO93/04175). Os ácidos nucleicos da invenção podem ser manipulados com base na preferência de códon em uma espécie de escolha, p. ex., humanos.

As composições de proteína de fusão Mtb72f compreendem usualmente um ou mais adjuvantes, p. ex., AS10B (lipídeo A de monosfosforila (MPL) e QS21 em uma formulação de lipossoma; ver, Publicação de Patente dos E.U.A. nº 2003/0143240); AS02A (3D-MPL e QS21 e uma emulsão óleo em água; ver, Bojang, *et al.*, *Lancet* (2001) 358:1927); ENHANZYN (Detox); 3D-MPL; saponinas incluindo Quil A e seus componentes, p. ex. QS21 e miméticos de saponina; CWS; TDM; AGP; olinucleopeptídeos imunoestimuladores, p. ex. CPG; Leif; e seus derivados. Em uma concretização preferida, um polipeptídeo de fusão Mtb72f é administrado com um ou mais adjuvantes selecionados do grupo que consiste de 3D-MPL e QS21 em uma formulação de lipossoma, p. ex. AS10B e MPL e QS21 e uma emulsão óleo em água (p. ex., AS02A). Adjuvantes AS10B e AS02A são descritos adicionalmente em Pichyangkul, *et al.*, *Vaccine* (2004) 22:3831-40.

Quando se fornece o antígeno de Mtb72f como um ácido nucleico, este pode ser fornecido, por exemplo, em um vetor viral (i.e., um vetor de adenovírus), ou em uma célula hospedeira de bactéria mutante (i.e., uma célula hospedeiro de *Mycobacterium*, *Lactobacillus* ou *Bacillus* mutante, avirulenta incluindo *Bacillus Calmette-Guerin* (BCG) e *Lactococcus lactis*).

Em um aspecto, as composições são empregadas em métodos para prevenir ou tratar reativação da tuberculose em um sujeito, sendo que o método compreende a etapa de administrar a um mamífero já infectado com *Mycobacterium tuberculosis* uma quantidade imunologicamente eficaz de uma composição farmacêutica compreendendo uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose e um adjuvante, sendo que a proteína de fusão

Mtb72f induz uma resposta imunológica contra *M. tuberculosis*, prevenindo com isso reativação de tuberculose. Praticando-se os métodos da presente invenção, é possível retardar a reativação de uma infecção com *M. tuberculosis* (por exemplo, por um período de meses, anos ou indefinidamente).

Em um aspecto, as composições são empregadas em métodos para prevenir ou tratar reativação da tuberculose em um sujeito, sendo que o método compreende a etapa de administrar a um mamífero já infectado com *Mycobacterium tuberculosis* uma quantidade imunologicamente eficaz de uma composição farmacêutica compreendendo um ácido nucleico que codifica uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose, sendo que a proteína de fusão Mtb72f expressa induz uma resposta imunológica contra *M. tuberculosis*, prevenindo, com isso, a reativação de tuberculose.

Em uma concretização, a proteína de fusão ou ácido nucleico de Mtb72f é administrada a um indivíduo com uma infecção ativa de *M. tuberculosis*. Em uma concretização, a proteína de fusão ou ácido nucleico de Mtb72f é administrada a um indivíduo com uma infecção inativa ou latente de *M. tuberculosis*. Em uma concretização, a proteína de fusão ou ácido nucleico de Mtb72f é administrada a um indivíduo infectado com uma cepa de *M. tuberculosis* resistentes a drogas múltiplas. Em uma concretização, a proteína de fusão ou ácido nucleico de Mtb72f é administrada a um indivíduo que foi imunizado previamente com Bacillus Calmette-Guerin (BCG).

Em algumas concretizações, a proteína de fusão ou ácido nucleico de Mtb72f é administrada com um ou mais agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis*. Exemplos de referidos agentes quimioterápicos incluem, embora sem limitação, amicacina, ácido aminossacilíco, capreomicina, ciclosserina, etambutol, etionamida, isoniazida, canamicina, pirazinamida, rifamicinas (i.e., rifampina, rifapentina e

rifabutina), estreptomicina, ofloxacina, ciprofloxacina, claritromicina, azitromicina e fluoroquinolonas. Referida quimioterapia é determinada pelo critério do médico responsável usando combinações de drogas preferidas. Agentes quimioterápicos de "primeira linha" usados para tratar uma infecção com *M. tuberculosis* que não é resistente a droga incluem isoniazida, rifampina, etambutol, estreptomicina e pirazinamida. Agentes quimioterápicos de "segunda linha" usados para tratar uma infecção com *M. tuberculosis* que demonstrou resistência-a-droga a uma ou mais drogas de "primeira linha" incluem ofloxacina, ciprofloxacina, etionamida, ácido aminossacilíco, ciclosserina, amicacina, canamicina e capreomicina.

A proteína de fusão ou ácido nucleico de Mtb72f pode ser administrada antes, concorrentemente com, ou após administração do um ou mais agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis*. Em uma concretização, a proteína de fusão ou ácido nucleico de Mtb72f é administrada cerca de 2 semanas após o início da administração de um ou mais agentes quimioterápicos. O um ou mais agentes quimioterápicos são administrados geralmente durante um determinado período, por exemplo, durante cerca de 1, 2, 3, ou 4 semanas, 2, 3, 4, 5, 6 ou 8 meses, 1 ano ou mais tempo.

Em determinadas concretizações, o efeito de uma proteína de fusão ou ácido nucleico de Mtb72f é incrementado pela administração com *Bacillus Calmette-Guerin* (BCG).

Em algumas concretizações, uma preparação ou primeira administração de um polipeptídeo de fusão ou ácido nucleico de Mtb72f é seguida de uma ou mais administração de "reforço" ou subseqüentes de um polipeptídeo de fusão ou ácido nucleico de Mtb72f (método de "iniciar e reforçar"). Por exemplo, uma primeira administração com um polipeptídeo de fusão ou ácido nucleico de Mtb72f é seguida de uma ou mais administrações subseqüentes de uma proteína de fusão ou ácido nucleico de Mtb72f. Em uma

concretização, uma primeira administração com um polipeptídeo de fusão ou ácido nucleico de Mtb72f é seguida de uma ou mais administrações subseqüentes de um polipeptídeo de fusão Mtb72f. Em uma concretização, uma primeira administração com um polipeptídeo de fusão ou ácido nucleico de Mtb72f é seguida de uma ou mais administrações subseqüentes de um ácido nucleico Mtb72f. Usualmente, a primeira administração ou administração de "iniciação" e a segunda administração ou administração de "reforço" são dadas com intervalo de cerca de 2-12 semanas, ou de até 4 a 6 meses. Administrações de "reforço" subseqüentes são dadas com intervalos de cerca de 6 meses, ou de até 1, 2, 3, 4 ou 5 anos. Tratamento de reforço convencional (p. ex., uma administração de início de proteína, seguida de uma administração de reforço de proteína) também é útil para prevenir ou tratar contra reativação de *M. tuberculosis*.

Em outro aspecto, as composições são empregadas em métodos para reduzir ou encurtar o decurso da quimioterapia contra uma infecção com *M. tuberculosis*, sendo que o método compreende administrar a um mamífero já infectado com *Mycobacterium tuberculosis* um ou mais agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis* e uma quantidade imunologicamente eficaz de uma composição farmacêutica compreendendo um polipeptídeo de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose e um adjuvante, sendo que referido polipeptídeo de fusão Mtb72f induz uma resposta imunológica contra *M. tuberculosis*, permitindo com isto reduzir ou encurtar o decurso da quimioterapia contra uma infecção com *M. tuberculosis*. Usualmente, a administração de um polipeptídeo de fusão ou ácido nucleico de Mtb72f se seguirá a tratamento quimioterápico eficaz contra uma infecção com *M. tuberculosis* dentro de 6 meses, 5 meses, 4 meses, 3 meses, ou menos.

As composições de Mtb72f são administradas usualmente a

humanos, mas são eficazes em outros mamíferos, incluindo mamíferos domésticos (i.e., cães, gatos, coelhos, ratos, camundongos, porquinhos-da-índia, hamsters, chinchilas) e mamíferos agrícolas (i.e., vacas, porcos, ovelha, cabras, cavalos).

5           Em seu aspecto mais geral, uma proteína de fusão Mtb72f de acordo com a invenção é uma proteína compreendendo pelo menos um fragmento imunogênico de cada um dos 3 antígenos Ra12-TbH9-Ra35.

10           Na nomenclatura da aplicação, Ra35 refere-se à extremidade N de Mtb32A (Ra35FL), compreendendo pelo menos cerca dos primeiros 205 aminoácidos de Mtb32A de *M. tuberculosis*, cuja seqüência de nucleotídeos e aminoácidos é revelada na Figura 4 do Pedido de Patente dos E.U.A. 09/597.796, ou a região correspondente de outra espécie de *Mycobacterium*. Da forma mais típica, Ra35 refere-se à porção da SEQ ID No: 2 revelada no presente pedido correspondente aos radicais de 535 a 729. Alternativamente  
15 isto refere-se a uma variante em Ra35 em que o aminoácido Ser correspondente a 710 na SEQ ID No: 2 é substituído por Ala.

20           Ra12 refere-se à extremidade C de Mtb32A (Ra35FL), compreendendo pelo menos cerca dos últimos 132 aminoácidos de MTB32A de *M. tuberculosis*, cuja seqüência é revelada como SEQ ID NO:4 (DNA) e SEQ ID NO:66 (seqüência de aminoácidos predita) no Pedido de Patente dos E.U.A. nº 09/072.967, ou a região correspondente de outra espécie de *Mycobacterium*. Da forma mais típica, Ra12 refere-se à porção da SEQ ID No: 2 revelada no presente pedido correspondente aos radicais de 8 a 139.

25           Mtb39 (TbH9) refere-se a uma seqüência essencialmente [] àquela divulgada como SEQ ID NO: 106 (cDNA de comprimento pleno) e SEQ ID NO: 107 (proteína de comprimento pleno) nos Pedidos de Patentes dos E.U.A. nº 08/658.800, nº 08/659.683, nº 08/818.112, e nº 08/818.111 e nos pedidos WO97/09428 e WO97/09429. A seqüência também é divulgada como SEQ ID NO:33 (DNA) e SEQ ID NO:91 (aminoácido) no Pedido de

Patente dos E.U.A. nº 09/056.559. Da forma mais típica, TbH9 refere-se à porção da SEQ ID No: 2 divulgada no presente pedido correspondente aos radicais de 143 a -532.

O texto a seguir proporciona seqüências de alguns antígenos individuais usados nas composições e proteínas de fusão da invenção:

Mtb32A (TbRa35FL ou Ra35FL), cuja seqüência é divulgada como SEQ ID NO: 17 (cDNA) e SEQ ID NO:79 (proteína) nos Pedidos de Patentes dos E.U.A. nums. 08/523.436, 08/523.435, nº 08/658.800, nº 08/659.683, nº 08/818.112, nº 09/056.556, e nº 08/818.111 e nos pedidos WO97/09428 e WO97/09429, ver também Skeiky *et al*, *Infection and Immunity* 67:3998-4007 (1999);

O texto a seguir proporciona seqüências de algumas proteínas de fusão da invenção:

TbH9-Ra35 (Mtb59F), cuja seqüência é divulgada como SEQ ID NO:23 (cDNA) e SEQ ID NO:24 (proteína) no Pedido de Patente dos E.U.A. nº 09/287.849 e no Pedido PCT/US99/07717;

Ra12-TbH9-Ra35 (Mtb72f), cuja seqüência é divulgada como SEQ ID NO:1 ou SEQ ID NO: 5 (DNA) e SEQ ID NO:2 ou SEQ ID NO:6 (proteína) no presente pedido, e também no Pedido de Patente dos E.U.A. nº 09/223.040, e no PCT/US99/07717. As seqüências da SEQ ID NO: 1 e SEQ ID NO:2 incluem um marcador His de radicais 6 His.

M72 que é um mutante de Mtb72f em que o radical serina no aminoácido correspondente à posição 710 na SEQ ID No: 2 foi substituído por Ala, (e também radicais 4 His removidos do marcador His na extremidade N) cuja seqüência é divulgada como SEQ ID No: 3 (DNA) e SEQ ID No: 4 (proteína) no presente pedido. Uma variante destas seqüências em que a proteína apresenta um marcador His de radicais 6 His é divulgada no Pedido de Patente dos E.U.A. nº 09/597.796 e no PCT/US01/19959. Devido à substituição de Ser710 por Ala, acredita-se que M72 é mais resistente a



autólise do que Mtb72f.

O texto a seguir proporciona seqüências de alguns antígenos adicionais usados nas composições e proteínas de fusão da invenção:

Mtb8.4 (DPV), cuja seqüência é divulgada como SEQ ID NO: 101 (cDNA) e SEQ ID NO:102 (proteína) nos Pedidos de Patentes dos E.U.A. num. 08/658.800, nº 08/659.683, nº 08/818.112 e nº 08/818.111 e nos Pedidos WO97/09428 e WO97/09429;

Mtb9.8 (MSL), cuja seqüência é divulgada como SEQ ID NO: 12 (DNA), SEQ ID NO: 109 (seqüência de aminoácidos predita) e SEQ ID NO: de 110 a 124 (peptídeos) nos Pedidos de Patentes dos E.U.A. nº 08/859.381, nº 08/858.998, nº 09/073.009 e nº 09/073.010 e nos Pedidos PCT/US98/10407 e PCT/US98/10514;

Mtb9.9A (MTI, também conhecido como MTI-A), cuja seqüência é divulgada como SEQ ID NO:3 e SEQ ID NO:4 (DNA) e SEQ ID NO:29 e SEQ ID NO: de 51 a 66 (peptídeo de matriz de leitura aberta para MTI) nos Pedidos de Patentes dos E.U.A. nº 08/859.381, nº 08/858.998, nº 09/073.009 e nº 09/073.010 e nos Pedidos PCT/US98/10407 e PCT/US98/10514. Também existem duas outras variantes de MTI, denominadas MTI-B e MTI-C;

Mtb40 (HTCC#1), cuja seqüência é divulgada como SEQ ID NO: 137 (cDNA) e 138 (seqüência de aminoácidos predita) nos Pedidos de Patentes dos E.U.A. nº 09/073.009 e nº 09/073.010 e nos Pedidos PCT/US98/10407 e PCT/US98/10514:

Mtb41 (MTCC#2), cuja seqüência é divulgada como SEQ ID NO: 140 (cDNA) e SEQ ID NO: 142 (seqüência de aminoácidos predita) nos Pedidos de Patentes dos E.U.A. nº 09/073,009 e nº 09/073,010 e nos Pedidos PCT/US98/10407 e PCT/US98/10514;

ESAT-6, cuja seqüência é divulgada como SEQ ID NO: 103 (DNA) e SEQ ID NO: 104 (seqüência de aminoácidos predita) no Pedido de

Patente dos E.U.A. nº 09/072.967. A seqüência de ESAT-6 também é divulgada na Patente dos Estados Unidos nº 5.955.077;

Antígeno  $\alpha$ -cristalino, cuja seqüência é divulgada em Verbon *et al*, *J. Bact.* 174:1352- 1359 (1992);

5 antígeno de complexo 85, cuja seqüência é divulgada em Content *et al*, *Infect. & Immunol.* 59:3205-3212 (1991).

Cada uma das seqüências acima também é divulgada em Cole *et al.* *Nature* 393:537 (1998) e podem ser encontradas, por exemplo, em <http://www.sanger.ac.uk> e <http://www.pasteur.fr/mycdb/>.

10 As seqüências acima são divulgadas nos Pedidos de Patentes dos E.U.A. nº 08/523.435, 08/523.436, 08/658.800, 08/659.683, 08/818.111, 08/818.112, 08/942.341, 08/942.578, 08/858.998, 08/859.381, 09/056.556, 09/072.596, 09/072.967, 09/073.009, 09/073.010, 09/223.040, 09/287.849 e nos Pedidos de Patentes PCT PCT/US98/10407, PCT/US98/10514, 15 PCT/US99/03265, PCT/US99/03268, PCT/US99/07717, WO97/09428 e WO97/09429, WO98/16645, WO98/16646, sendo que cada um é incorporado aqui por referência.

Os antígenos aqui descritos incluem variantes polimórficas e variações modificadas conservativamente, e também homólogos de 20 *Mycobacterium* inter-cepas e inter-espécies. Adicionalmente, os antígenos aqui descritos incluem subseqüências ou seqüências truncadas. As proteínas de fusão também podem conter polipeptídeos adicionais, opcionalmente peptídeos heterólogos de *Mycobacterium* ou outras fontes. Estes antígenos podem ser modificados, por exemplo, por meio de adição de seqüências de 25 peptídeos ligantes como descrito abaixo. Estes peptídeos ligantes podem ser inseridos entre um ou mais componentes que constituem cada uma das proteínas de fusão.

### Definições

O termo "reativação de tuberculose" refere-se à manifestação

tardia de sintomas de doenças em um indivíduo que se testou positivo em um teste de tuberculina, mas que não apresenta sintomas de doença aparentes. O indivíduo é infectado com *M. tuberculosis*, e pode ou não apresentar sintomas de doença ativa previamente manifestada que foi tratada suficientemente para trazer a tuberculose a um estado inativo ou latente. No entanto, é possível iniciar métodos para a prevenção ou tratamento de reativação da tuberculose em um indivíduo que manifesta sintomas ativos de doença.

"Tuberculose primária" refere-se a doença clínica (manifestação de sintomas de doença) diretamente após infecção com *M. tuberculosis*. Ver, *Harrison's Principles of Internal Medicine*, capítulo 150, pp. 953-966 (16<sup>a</sup> ed., Braunwald, *et al*, eds., 2005).

"Tuberculose secundária" ou "tuberculose pós-primária" refere-se à reativação de infecção dormente, inativa ou latente com *M. tuberculosis*. Ver, *Harrison's Principles of Internal Medicine*, supra.

Uma "infecção ativa de *M. tuberculosis*" refere-se a uma infecção com *M. tuberculosis* com sintomas de doença manifestados.

Uma "infecção inativa, dormente ou latente de *M. tuberculosis*" refere-se a uma infecção com *M. tuberculosis* sem sintomas de doença manifestados.

Uma infecção de *M. tuberculosis* "resistente a droga" refere-se a uma infecção com *M. tuberculosis* sendo que a cepa infectante não é mantida estática ou morta (é resistente a) um ou mais assim-chamados agentes quimioterápicos eficazes "de primeira linha" para tratar uma infecção com *M. tuberculosis* (p. ex., isoniazida, rifampina, etambutol, estreptomicina e pirazinamida).

Uma infecção de *M. tuberculosis* "resistente a múltiplas drogas" refere-se a uma infecção com *M. tuberculosis*, sendo que a cepa infectante é resistente a dois ou mais agentes quimioterápicos eficazes de "linha de frente" para tratar uma infecção de *M. tuberculosis*.

Um "agente quimioterápico eficaz para tratar uma infecção com *M. tuberculosis*" refere-se a agentes farmacológicos conhecidos e usados na arte para tratar infecções com *M. tuberculosis*. Agentes farmacológicos exemplificados usados para tratar infecções com *M. tuberculosis* incluem, 5 embora sem limitação, amicacina, ácido aminossacílico, capreomicina, ciclosserina, etambutol, etionamida, isoniazida, canamicina, pirazinamida, rifamicinas (i.e., rifampina, rifapentina e rifabutina), estreptomicina, ofloxacina, ciprofloxacina, claritromicina, azitromicina e fluoroquinolonas. Agentes quimioterápicos de "primeira linha" usados para tratar uma infecção 10 com *M. tuberculosis* que não é resistente a droga incluem isoniazida, rifampina, etambutol, estreptomicina e pirazinamida. Agentes quimioterápicos de "segunda linha" usados para tratar uma infecção com *M. tuberculosis* que demonstrou resistência-a-droga a uma ou mais drogas de "primeira linha" incluem ofloxacina, ciprofloxacina, etionamida, ácido 15 aminossacílico, ciclosserina, amicacina, canamicina e capreomicina. Referidos agentes farmacológicos são revistos no capítulo 48 de *Goodman e Gilman's The Pharmacological Basis of Therapeutics*, Hardman e Limbird eds., 2001.

"FL" refere-se a *full-length* [comprimento total], i.e., um 20 polipeptídeo que tem o mesmo comprimento do polipeptídeo de tipo selvagem.

"Marcador His [*His tag*]" refere-se a um seqüência de radicais His, tipicamente 6 radicais que são inseridos na extremidade N, usualmente e imediatamente após o início do radical Met ou, então, na extremidade C. Eles 25 são usualmente heterólogos com relação à seqüência nativa, mas são incorporados porque facilitam o isolamento por meio de aperfeiçoamento da ligação da proteína sobre resinas de cromatografia de afinidade de metal imobilizadas (IMAC). De uma forma geral, a presença ou ausência de um marcador His não é significativa do ponto de vista de causar uma resposta

imunológica útil contra a proteína antigênica a ser elicitada. No caso de uma reação imune adversa contra o próprio marcador His ser elicitada, é considerado a melhor opção minimizar o comprimento do marcador His, p. ex., a 4 ou menos radicais, em particular dois radicais.

5 O termo "fragmento imunogênico do mesmo" refere-se a um polipeptídeo compreendendo um epitopo que é reconhecido por linfócitos T citotóxicos, linfócitos T auxiliares ou células B. Tipicamente um fragmento imunogênico de Mtb72f será um polipeptídeo contendo 500 ou mais aminoácidos, p. ex. 600 ou mais aminoácidos, p. ex. 700 ou mais  
10 aminoácidos. A invenção também abrange uma pluralidade de fragmentos, p. ex. superpondo fragmentos que, em conjunto, cobrem toda ou substancialmente toda a seqüência (p. ex. 500 ou mais aminoácidos, p. ex. 600 ou mais aminoácidos, p. ex. 700 ou mais aminoácidos) de uma proteína de fusão Mtb72F.

15 O termo "espécie de *Mycobacterium* do complexo de tuberculose" inclui aquelas espécies consideradas tradicionalmente como causando a doença tuberculose, e também espécies ambientais e oportunistas de *Mycobacterium* que causam tuberculose e doença do pulmão em pacientes imunocomprometidos, como pacientes com AIDS, p.  
20 ex., *M. tuberculosis*, *M. bovis*, ou *M. africanum*, BCG, *M. avium*, *M. intracellular*, *M. celatum*, *M. genavense*, *M. haemophilum*, *M. kansasii*, *M. simiae*, *M. vaccae*, *M. fortuitum*, e *M. scrofulaceum* (ver, p. ex., *Harrison's Principles of Internal Medicine*, capítulo 150, pp. 953-966 (16<sup>a</sup> ed., Braunwald, *et al*, eds., 2005).

25 Um adjuvante refere-se aos componentes em uma vacina ou composição terapêutica que incrementa a resposta imunológica específica ao antígeno (ver, p. ex., Edelman, *AIDS Res. Hum Retroviruses* 8:1409-1411 (1992)). Adjuvantes induzem respostas imunológicas da resposta de tipo Th1 e de tipo Th-2. As citocinas de tipo Th1 (p. ex., IFN- $\gamma$ , IL-2, e IL-12) tendem

a favorecer a indução de resposta imunológica mediada com células a um antígeno administrado, enquanto que citocinas de tipo Th-2 (p. ex., IL-4, IL-5, IL-6, IL-10 e TNF- $\beta$ ) tendem a favorecer a indução de respostas imunológicas humorais. Adjuvantes capazes de estimulação preferencial de uma resposta imunológica mediada com célula Th-1 encontram-se descritos no WO 94/00153 e WO 95/17209.

"Ácido nucleico" refere-se a desoxirribonucleotídeos ou ribonucleotídeos e polímeros dos mesmos em forma de filamento simples ou duplo. O termo compreende ácidos nucleicos contendo análogos de nucleotídeos conhecidos ou radicais de estrutura modificada ou ligações, que são sintéticas, naturalmente ocorrentes, e não-naturalmente ocorrentes, que apresentam propriedades de ligação similares ao ácido nucleico de referência, e que são metabolizadas de maneira similar aos nucleotídeos de referência. Exemplos de referidos análogos incluem, sem limitação, fosforotioatos, fosforamidatos, fosfonatos de metila, fosfonatos de metila quirais, ribonucleotídeos de 2-O-metila, ácidos nucleicos de peptídeos (PNAs).

Exceto se indicado de outra forma, uma seqüência de ácido nucleico particular também compreende implicitamente suas variantes modificadas conservativamente (p. ex., substituições de códons degenerados) e seqüências complementares, e também a seqüência indicada explicitamente. Especificamente, é possível obter substituições de códons degenerados por meio de geração de seqüências em que a terceira posição de um ou mais códons selecionados (ou de todos) é substituída por radicais de base mista e/ou desoxinosina (Batzer *et al*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al*, *Mol. Cell. Probes* 8:91-98 (1994)). O termo ácido nucleico é usado de forma intercambiável com gene, cDNA, mRNA, oligonucleotídeo, e polinucleotídeo.

Os termos "polipeptídeo", "peptídeo" e "proteína" são usados intercambiavelmente aqui para referir a um polímero de radicais de

aminoácidos. Os termos aplicam-se a polímeros de aminoácidos em que um ou mais radicais de aminoácidos é um mimético artificial de um aminoácido naturalmente ocorrente correspondente, e também polímeros de aminoácidos naturalmente ocorrentes e polímero de aminoácidos não-naturalmente ocorrentes.

O termo "aminoácido" refere-se a aminoácidos naturalmente ocorrentes e sintéticos, e também análogos de aminoácidos e miméticos de aminoácido que funcionam de maneira similar aos aminoácidos naturalmente ocorrentes. Aminoácidos naturalmente ocorrentes são aqueles codificados pelo código genético, e também aqueles aminoácidos que são modificados posteriormente, p. ex., hidroxiprolina,  $\gamma$ - carboxiglutamato, e O-fosfoserina. Análogos de aminoácidos refere-se a compostos que apresentam a mesma estrutura química básica que um aminoácido naturalmente ocorrente, i.e., um  $\alpha$  carbono que é ligado a um hidrogênio, um grupo carboxila, um grupo amino, e um grupo R, p. ex., homoserina, norleucina, sulfóxido de metionina, metil sulfônio de metionina. Referidos análogos apresentam grupos R modificados (p. ex., norleucina) ou espinhas dorsais de peptídeos modificadas, mas conservam a mesma estrutura química básica que o aminoácido naturalmente ocorrente. Miméticos de aminoácidos referem-se a compostos químicos que apresentam a estrutura que é diferente da estrutura química geral de um aminoácido, mas que funciona de maneira similar a um aminoácido naturalmente ocorrente.

Aminoácidos podem ser referidos aqui por seus símbolos de três letras comumente conhecidos ou pelos símbolos de uma letra só recomendados pela Comissão de Nomenclatura Bioquímica da IUPAC-IUB. Nucleotídeos, da mesma forma, podem ser referidos por seus códigos de uma letra só comumente aceitos.

"Variantes modificadas conservativamente" aplica-se a seqüências de aminoácidos e ácido nucleico. Com relação a seqüências de

ácido nucleico particulares, variantes modificadas conservativamente referem-se àqueles ácidos nucleicos que codificam seqüências de aminoácidos idênticas ou substancialmente idênticas, ou em que o ácido nucleico não codifica uma seqüência de aminoácidos, a seqüências substancialmente idênticas. Devido à degeneração do código genético, um grande número de ácidos nucleicos funcionalmente idênticos codifica qualquer dada proteína. Por exemplo, todos os códons GCA, GCC, GCG e GCU codificar o aminoácido alanina. Assim, em cada posição em que uma alanina é especificada por um códon, o códon pode ser alterado para qualquer um dos códons correspondentes descritos sem alterar o polipeptídeo codificado. Referidas variações de ácido nucleico são "variações silentes", que são uma espécie de variações modificadas conservativamente. Cada seqüência de ácido nucleico aqui que codifica um polipeptídeo também descreve cada possível variação silente do ácido nucleico. Alguém com prática na arte será capaz de reconhecer que cada códon em um ácido nucleico (exceto AUG, que é ordinariamente o único códon para metionina, e TGG, que é ordinariamente o único códon para triptofano) pode ser modificado para dar uma molécula funcionalmente idêntica. Assim, cada variação silente de um ácido nucleico que codifica um polipeptídeo está implícita em cada seqüência descrita.

Quanto às seqüências de aminoácidos, alguém com prática na arte reconhecerá que substituições, deleções ou adições individuais em um ácido nucleico, peptídeo, polipeptídeo, ou seqüência de proteína que alteram, adicionam, ou deletam um único aminoácido ou um pequeno percentual de aminoácidos na seqüência codificada é uma "variante modificada conservativamente" sendo que a alteração resulta na substituição de um aminoácido por um aminoácido quimicamente similar. Tabelas de substituições conservativas que proporcionam aminoácidos funcionalmente similares são bem conhecidos na arte. Referidas variantes modificadas conservativamente são adicionais e não excluem variantes polimórficas,



homólogos inter-espécies, e alelos da invenção.

Cada um dos oito grupos a seguir contém aminoácidos que são substituições conservativas mútuas:

- 1) Alanina (A), Glicina (G);
  - 5 2) Ácido aspártico (D), Ácido glutâmico (E);
  - 3) Asparagina (N), Glutamina (Q);
  - 4) Arginina (R), Lisina (K);
  - 5) Isoleucina (I), Leucina (L), Metionina (M),  
Valina (V);
  - 10 6) Fenilalanina (F), Tirosina (Y), Triptofano (W);
  - 7) Serina (S), Treonina (T); e
  - 8) Cisteína (C), Metionina (M)
- (ver, p. ex., Creighton, *Proteins* (1984)).

O termo "heterólogo" quando usado com referência a porções  
 15 de um ácido nucleico, indica que o ácido nucleico compreende duas ou mais  
 subsequências que não são encontradas na mesma relação entre si na natureza.  
 Por exemplo, tipicamente, o ácido nucleico é produzido recombinantemente,  
 apresentando duas ou mais seqüências de genes não-relacionados dispostas  
 para formar um ácido nucleico funcional novo, p. ex., um promotor de uma  
 20 fonte e uma região codificante de outra fonte. De maneira análoga, uma  
 proteína heteróloga indica que a proteína compreende duas ou mais  
 subsequências que não são encontradas na mesma relação entre si na natureza  
 (p. ex., uma proteína de fusão).

"Polipeptídeo de fusão" ou "proteína de fusão" refere-se a uma  
 25 proteína apresentando pelo menos dois polipeptídeos heterólogos de  
*Mycobacterium* sp. ligados covalentemente, seja diretamente ou via um  
 ligante de aminoácido. Os polipeptídeos que formam a proteína de fusão são  
 ligados tipicamente extremidade C à extremidade N, embora também possam  
 ser ligados extremidade C à extremidade C, extremidade N à extremidade N,

ou extremidade N à extremidade C. Os polipeptídeos da proteína de fusão podem ser em qualquer ordem. Este termo também se refere a variantes modificadas conservativamente, variantes polimórficas, alelos, mutantes, subseqüências, e homólogos inter-espécies dos antígenos que formam a

5 proteína de fusão. Antígenos de *Mycobacterium tuberculosis* são descritos em Cole *et al*, *Nature* 393:537 (1998), que divulga todo o genoma de *Mycobacterium tuberculosis*. A seqüência completa de *Mycobacterium tuberculosis* também pode ser encontrada em <http://www.sanger.ac.uk> e at <http://www.pasteur.fr/mycdb/> (MycDB). Antígenos de outras espécies de

10 *Mycobacterium* que correspondem a antígenos de *M. tuberculosis* podem ser modificados, p. ex., usando-se algoritmos de comparação de seqüências, como descrito aqui, ou outros métodos conhecidos por aqueles com prática na arte, p. ex., ensaios de hibridização e ensaios de ligação de anticorpo.

15 Proteínas de fusão Mtb72f exemplares de uso na presente invenção incluem:

Proteínas compreendendo radicais de 8 a 729 da seqüência da SEQ ID No: 2;

20 Proteínas compreendendo ou consistindo da seqüência da SEQ ID No: 2 (=Mtb72f) opcionalmente sem os radicais de 2 a 7 formadores do marcador His de referida seqüência ou com um marcador His de comprimento diferente;

25 Proteínas de fusão compreendendo a seqüência SEQ ID No: 2 opcionalmente sem os radicais de 2 a 7 formadores de marcador His de referidas seqüências ou com um marcador His de comprimento diferente (p. ex., uma proteína compreendendo radicais de 8 a 729 da seqüência da SEQ ID No: 2) em conjunto com um ou mais antígenos de *M. tuberculosis*, por exemplo uma ou mais das proteínas listadas nos parágrafos de [0045] a [0052] acima, ou um fragmento imunogênico das mesmas;

Proteínas compreendendo radicais de 4 a 725 da seqüência da

SEQ ID No: 4 (=M72);

Proteínas compreendendo ou consistindo da seqüência da SEQ ID No: 4 (=M72) opcionalmente sem os radicais de 2 a 3 formadores do marcador His de referida seqüência ou com um marcador His de comprimento  
5 diferente; e

Proteínas de fusão compreendendo a seqüência SEQ ID No: 4 opcionalmente sem os radicais de 2 a 3 formadores de marcador His de referida seqüência ou com um marcador His de comprimento diferente (p. ex. uma proteína compreendendo radicais de 4 a 725 da seqüência da SEQ ID  
10 No: 4) em conjunto com um ou mais antígenos de *M. tuberculosis*, por exemplo um ou mais das proteínas listadas nos parágrafos de [0045] a [0052] acima, ou um fragmento imunogênico de qualquer uma das mesmas;

Fragmentos imunogênicos exemplares de uma proteínas de fusão Mtb72f de uso na presente invenção incluem:

15 Proteínas compreendendo ou consistindo da seqüência de TbH9-Ra35 (Mtb59F); ou TbH9; ou Ra35; ou Ra12; e

Proteínas de fusão compreendendo referidas seqüências em conjunto com um ou mais antígenos de *M. tuberculosis*, por exemplo uma ou mais das proteínas listadas nos parágrafos [0045] a [0052] acima, ou um  
20 fragmento imunogênico de qualquer um dos mesmos.

Fragmentos imunogênicos exemplares adicionais de uma proteínas de fusão Mtb72f de uso na presente invenção incluem:

25 Proteínas compreendendo ou consistindo da seqüência de TbH9-Ra35 (Mtb59F) ou Ra35 em que a posição correspondente a Ser710 na SEQ ID No: 2 foi substituída por Ala; e

Proteínas de fusão compreendendo referidas seqüências em conjunto com um ou mais antígenos de *M. tuberculosis*, por exemplo uma ou mais das proteínas listadas nos parágrafos de [0045] a [0052] acima, ou um fragmento imunogênico de qualquer um dos mesmos.

Mais especificamente a Mtb72f é:

um polipeptídeo compreendendo radicais de 8 a 729 da SEQ ID NO:2; ou

5 um polipeptídeo que consiste de radicais 1 e de 8 a 729 da SEQ ID NO:2 opcionalmente com um marcador His inserido após o radical Met inicial; ou

um polipeptídeo da SEQ ID NO:2; ou

um polipeptídeo compreendendo radicais de 4 a 725 da SEQ ID NO:4; ou

10 um polipeptídeo que consiste de radicais 1 e de 4 a 725 da SEQ ID NO:4 opcionalmente com um marcador His inserido após o radical Met inicial; ou

um polipeptídeo da SEQ ID NO:4; ou

um polipeptídeo da SEQ ID NO:6.

15 Proteínas de fusão Mtb72f exemplares adicionais e fragmentos imunogênicos das mesmas incluem as proteínas indicadas acima em que as extremidades N e/ou C foram encurtadas em, por exemplo, 5 ou 4 ou 3 ou 2 ou 1 radicais de aminoácidos.

20 Proteínas de fusão Mtb72f exemplares adicionais e fragmentos imunogênicos das mesmas incluem as proteínas mencionadas acima em que até 10% dos aminoácidos, p. ex. até 5% dos aminoácidos (p. ex. até 10, p. ex. até 5) aminoácidos foram substituídos por substituições conservativas como definido aqui.

25 Ácidos nucleicos de Mtb72f exemplares de uso na presente invenção incluem ácidos nucleicos (p. ex., moléculas de DNA) codificando as proteínas de fusão Mtb72f exemplares previamente mencionadas e fragmentos imunogênicos das mesmas. Um conjunto de moléculas de DNA específicas que pode ser mencionado compreende nucleotídeos de 63 a 2228 da SEQ ID No: 1. Outro conjunto de moléculas de DNA específicas que pode

ser mencionado compreende nucleotídeos de 10 a 2175 da SEQ ID No: 3. Moléculas de DNA específicas que podem ser mencionadas compreendem ou consistem da SEQ ID No: 1 ou SEQ ID No: 3 ou SEQ ID No: 5.

5 O termo "fundido" refere-se à ligação covalente entre dois polipeptídeos em uma proteína de fusão. Os polipeptídeos são conjugados tipicamente via uma ligação peptídica, seja diretamente entre si ou via uma ligante de aminoácido. Opcionalmente, os peptídeos podem ser conjugados via ligações covalentes não-peptídicas conhecidas por aqueles com prática na arte.

10 A expressão "hibridiza seletivamente (ou especificamente) com" refere-se à ligação, duplexação, ou hibridização de uma molécula apenas com uma seqüência de nucleotídeo particular em condições de hibridização estringentes quando a seqüência está presente em uma mistura complexa (p. ex., RNA ou DNA celular total ou de biblioteca).

15 A expressão "condições de hibridização estringentes" refere-se a condições em que uma sonda hibridará com sua subseqüência-alvo, tipicamente em uma mistura complexa de ácido nucleico, mas não com outras seqüências. Condições estringentes são dependentes da seqüência e serão diferentes em circunstâncias diferentes. Seqüências mais longas hibridam especificamente a temperaturas mais elevadas. Um guia extensivo para a  
20 hibridização de ácidos nucleicos é encontrado em Tijssen, *Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). De uma forma geral, condições estringentes são selecionadas  
25 cerca de 5 a 10°C mais baixas do que o ponto de fusão térmica ( $T_m$ ) para a seqüência específica a um pH definido de intensidade iônica. A  $T_m$  é a temperatura (sob intensidade iônica definida, pH, e concentração nucleica) à qual 50% das sondas complementares ao alvo hibridam com a seqüência-alvo em equilíbrio (como as seqüências-alvos estão presentes em excesso, à  $T_m$ ,

50% das sondas são ocupadas em equilíbrio). Condições estridentes serão aquelas em que a concentração de sal é inferior a cerca de 1,0 M de íon sódio, tipicamente, uma concentração de cerca de 0,01 a 1,0 M de íon sódio (ou outros sais) a pH de 7,0 a 8,3 e a temperatura é de pelo menos cerca de 30°C para sondas curtas (p. ex., de 10 a 50 nucleotídeos) e de pelo menos cerca de 60°C para sondas longas (p. ex., acima de 50 nucleotídeos). Condições estridentes também podem ser obtidas com a adição de agentes desestabilizantes, como formamida.

Para hibridização seletiva ou específica, um sinal positivo é de pelo menos duas vezes o fundo, opcionalmente 10 vezes a hibridização de fundo. Condições de hibridização estridentes exemplares podem ser as seguintes: 50% de formamida, 5x de SSC, e 1% de SDS, incubando a 42°C, ou, 5x SSC, 1% de SDS, incubando a 65°C, com lavagem em 0,2x de SSC, e 0,1% de SDS a 65°C.

Ácidos nucleicos que não hibridam entre si em condições estridentes ainda são substancialmente idênticas se os polipeptídeos que elas codificam forem substancialmente idênticos. Isto ocorre, por exemplo, quando uma cópia de um ácido nucleico é criada usando-se a máxima degeneração de códon permitida pelo código genético. Nesses casos, os ácidos nucleicos hibridam tipicamente em condições de hibridização moderadamente estridentes. "Condições de hibridização moderadamente estridentes" exemplares incluem uma hibridização em um tampão de 40% de formamida, 1 M de NaCl, 1% de SDS a 37°C, e uma lavagem em 1X SSC a 45°C. Uma hibridização positiva é pelo menos duas vezes o fundo. Aqueles com prática ordinária reconhecerão facilmente que condições de lavagem e hibridização alternativas podem ser usadas para proporcionar condições de estridência similar.

"Anticorpo" refere-se a um polipeptídeo compreendendo uma região de matriz de um gene de imunoglobulina ou fragmentos do mesmo que

se ligam especificamente e reconhecem um antígeno. Os genes de imunoglobulina reconhecidos incluem os genes de região constante mu, capa, lambda, alfa, gama, delta e epsilon, e também os muitos genes de região variável de imunoglobulina. Cadeias leves são classificadas como capa ou lambda. Cadeias pesadas são classificadas como gama, mu, alfa, delta, ou epsilon, que, por sua vez, definem as classes de imunoglobulina, IgG, IgM, IgA, IgD e IgE, respectivamente.

Uma unidade estrutural de imunoglobulina exemplar (anticorpo) compreende um tetrâmero. Cada tetrâmero é constituída de dois pares idênticos de cadeias de polipeptídeos, cada par apresentando uma cadeia "leve" (cerca de 25 kDa) e uma "pesada" (cerca de 50 a 70 kDa). A extremidade N de cada cadeia define uma região variável de cerca de 100 a 110 ou mais aminoácidos primariamente responsáveis para reconhecimento de antígeno. Os termos cadeia leve variável ( $V_L$ ) e cadeia pesada variável ( $V_H$ ) referem-se a estas cadeias leve e pesada, respectivamente.

Anticorpos existem, p. ex., como imunoglobulinas intactas ou como uma variedade de fragmentos bem caracterizados produzidos por meio de digestão com várias peptidases. Assim, por exemplo, pepsina digere um anticorpo abaixo das ligações dissulfeto na região de dobradiça para produzir  $F(ab)'_2$ , um dímero de Fab que, por si só, é uma cadeia leve conjugada a  $V_H$ - $C_{H1}$  por uma ligação dissulfeto. O  $F(ab)'_2$  pode ser reduzido em condições brandas para quebrar a ligação dissulfeto na região de dobradiça, convertendo com isso o dímero  $F(ab)'_2$  em um monômero Fab'. O monômero Fab' é essencialmente Fab com parte da região de dobradiça (ver *Fundamental Immunology* (Paul ed., 3<sup>a</sup> ed. 1993)). Embora vários fragmentos de anticorpos sejam definidos em termos da digestão de um anticorpo intacto, alguém com prática perceberá que referidos fragmentos podem ser sintetizados *de novo*, seja quimicamente ou com o uso de metodologia de DNA recombinante. Assim, o termo anticorpo, como usado aqui, também inclui fragmentos de

anticorpo, seja produzidos por meio da modificação de anticorpos inteiros, ou aqueles sintetizados de novo usando metodologias de RNA recombinantes (p. ex., Fv de cadeia simples) ou aqueles identificados usando bibliotecas de apresentação de fago (ver, p. ex., McCafferty *et al*, *Nature* 348:552-554 (1990)).

Para a preparação de anticorpos monoclonais ou policlonais, é possível usar qualquer técnica conhecida na arte (ver, p. ex., Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al*, pp. 77-96 em *Monoclonal Antibodies and Cancer Therapy* (1985)). Técnicas para a produção de anticorpos de cadeia simples (Patente dos Estados Unidos nº 4,946,778) podem ser adaptadas para produzir anticorpos para polipeptídeos desta invenção. Também é possível usar camundongos transgênicos, ou outros organismos, como outros mamíferos, para expressar anticorpos humanizados. Alternativamente, é possível usar tecnologia de apresentação de fago para identificar anticorpos e fragmentos Fab heteroméricos que se ligam especificamente a antígenos selecionados (ver, p. ex., McCafferty *et al*, *Nature* 348:552-554 (1990); Marks *et al*, *Biotechnology* 10:779-783 (1992)).

A expressão "liga-se especificamente (ou seletivamente)" a um anticorpo ou "especificamente (ou seletivamente) imunorreativo com", quando referindo a uma proteína ou peptídeo, refere-se a uma reação de ligação que é determinante da presença da proteína em uma população heterogênea de proteínas e outras biológicas. Assim, em condições determinadas de ensaio imunológico, os anticorpos especificados ligam-se a uma proteína particular pelo menos duas vezes o fundo e não se ligam substancialmente numa quantidade significativa a outras proteínas presentes na amostra. Ligação específica a um anticorpo nessas condições pode exigir que um anticorpo seja selecionado por sua especificidade para uma proteína particular. Por exemplo, anticorpos policlonais desenvolvidos a proteínas de



5 fusão podem ser selecionados de forma a se obter apenas aqueles anticorpos policlonais que são especificamente imunorreativos com proteína de fusão e não com componentes individuais das proteínas de fusão. Esta seleção pode ser obtida subtraindo-se anticorpos que reagem de maneira cruzada com  
10 antígenos individuais. É possível usar uma variedade de formatos de ensaios imunológicos para selecionar anticorpos especificamente imunorreativos com uma proteína particular. Por exemplo, ensaios imunológicos ELISA de fase sólida são usados rotineiramente para selecionar anticorpos especificamente imunorreativos com uma proteína (ver, p. ex., Harlow & Lane, *Antibodies, A*  
15 *Laboratory Manual* (1988) e *Using Antibodies: A Laboratory Manual* (1998), para uma descrição de formatos e condições de ensaio imunológico que podem ser usados para se determinar imunorreatividade específica). Tipicamente, uma reação específica ou seletiva será de pelo menos duas vezes o sinal ou ruído de fundo e, mais tipicamente, de 10 a 100 vezes o fundo.

15 Polinucleotídeos podem compreender uma seqüência nativa (i.e., uma seqüência endógena que codifica um antígeno individual ou porção do mesmo) ou podem compreender uma variante de uma seqüência do tipo referido. Variantes de polinucleotídeos podem conter uma ou mais substituições, adições, deleções e/ou inserções, de tal forma que a atividade  
20 biológica o polipeptídeo de fusão codificado não seja diminuída, relativamente a um polipeptídeo de fusão compreendendo antígenos nativos. Variantes apresentam, de preferência, pelo menos cerca de 70% de identidade, mais preferivelmente, pelo menos cerca de 80% de identidade e, da forma mais preferível, pelo menos cerca de 90% de identidade com  
25 seqüência de polinucleotídeos que codifica um polipeptídeo nativo ou uma porção do mesmo.

Os termos "idêntica" ou percentual de "identidade", no contexto de dois ou mais ácidos nucleicos ou seqüências de polipeptídeos, referem-se a duas ou mais seqüências ou subseqüências que são iguais ou que

apresentam um percentual especificado de radicais de aminoácidos ou nucleotídeos que são iguais (i.e., 70% de identidade, opcionalmente 75%, 80%, 85%, 90%, ou 95% de identidade relativamente a uma região especificada), quando comparado e alinhado para máxima correspondência  
5 relativamente a uma janela de comparação, ou região determinada conforme medido usando-se um dos algoritmos de comparação de seqüências ou por meio de alinhamento manual e inspeção visual. Referidas seqüências são então declaradas como sendo "substancialmente idênticas". Esta definição também se refere ao complemento de uma seqüência de teste. Opcionalmente,  
10 a identidade existe numa região que tem comprimento de pelo menos cerca de 25 a cerca de 50 aminoácidos ou nucleotídeos, ou, opcionalmente, numa região que tem um comprimento de 75 a 100 aminoácidos ou nucleotídeos.

Para comparação de seqüências, tipicamente uma seqüência atua como uma seqüência de referência com que seqüências de teste são  
15 comparadas. Quando se usa um algoritmo de comparação de seqüências, seqüências de teste e de referência são introduzidas em um computador, determina-se coordenadas de subseqüências, se necessário, e determina-se parâmetros de programa de algoritmos de seqüências. É possível usar parâmetros de programa default, ou é possível determinar parâmetros  
20 alternativos. O algoritmo de comparação de seqüências calcula então o percentual de identidade de seqüências para as seqüências de teste relativamente à seqüência de referência, com base nos parâmetros de programa.

Uma "janela de comparação", como usado aqui, inclui  
25 referência a um segmento de qualquer uma de várias posições contíguas selecionadas do grupo que consiste de 25 a 500, usualmente de cerca de 50 a cerca de 200, mais comumente de cerca de 100 a cerca de 150 em que uma seqüência pode ser comparada com uma seqüência de referência com o mesmo número de posições contíguas após as duas seqüências terem sido

alinhadas de maneira ótima. Métodos de alinhamento de seqüências para comparação são bem conhecidos na arte. O alinhamento ótimo de seqüências para comparação pode ser conduzido, p. ex., por meio do algoritmo de homologia local de Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), por meio do algoritmo de alinhamento de homologia de Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), pelo método de busca por similaridade de Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), por meio de implementações computadorizadas destes algoritmos (ESPAÇO, BESTFIT, FASTA, e TFASTA no pacote de programas Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), ou por meio de alinhamento manual e inspeção visual (ver, p. ex., *Current Protocols in Molecular Biology* (Ausubel *et al*, eds., suplemento de 1995)).

Um exemplo de um algoritmo útil é PILEUP. PILEUP cria um alinhamento de seqüências múltiplas de um grupo de seqüências relacionadas usando-se alinhamentos pareados progressivos para mostrar relação e percentual de identidade de seqüências. Ele também plota uma árvore ou dendograma mostrando as relações de aglomeração usadas para criar o alinhamento. PILEUP usa uma simplificação do método de alinhamento progressivo de Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). O método usado é semelhante ao método descrito por Higgins & Sharp, *CABIOS* 5:151-153 (1989). O programa pode alinhar até 300 seqüências, cada uma com um comprimento máximo de 5.000 nucleotídeos ou aminoácidos. O procedimento de alinhamento múltiplo começa com o alinhamento pareado das duas seqüências mais similares, produzindo um aglomerado de duas seqüências alinhadas. Este aglomerado é então alinhado com a seqüência próxima mais relacionada ou com aglomerado de seqüências mais alinhadas. Dois aglomerados de seqüências são alinhados por meio de uma extensão simples do alinhamento pareado de duas seqüências individuais. O alinhamento final é obtido por meio de uma série de alinhamentos pareados progressivos. O

programa é operado determinando-se seqüências específicas e suas coordenadas de aminoácidos ou nucleotídeos para regiões de comparação de seqüência e determinando-se os parâmetros de programa. Usando PILEUP, compara-se uma seqüência de referência com outras seqüências de teste para  
5 determinar a relação de percentual de identidade de seqüências usando-se os parâmetros a seguir: peso espaço default (3,00), peso do comprimento do espaço default (0,10), e espaços de extremidade ponderal. PILEUP pode ser obtido do pacote de programas de análise de seqüências GCG, p. ex., versão 7.0 (Devereaux *et al*, *Nuc. Acids Res.* 12:387-395 (1984).

10                   Outro exemplo de algoritmo que é vantajoso para determinar o percentual de percentual de identidade de seqüências e similaridade de seqüências são os algoritmos BLAST e BLAST 2.0, que são descritos em Altschul *et al*, *Nuc. Acids Res.* 25:3389-3402 (1977) e Altschul *et al*, *J. Mol. Biol.* 215:403-410 (1990), respectivamente. Programa [software] para realizar  
15 análises BLAST encontra-se disponível ao público através do National Center for Biotechnology Information [Centro Nacional de Informação de Biotecnologia] (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo envolve primeiramente identificar pares de seqüências com alta classificação (HSPs, *high scoring sequence pairs*) por meio de identificação de palavras curtas de  
20 comprimento W na seqüência de busca, que ou se equiparam ou satisfazem alguma classificação de limiar T com valor positivo quando alinhada com uma palavra de mesmo comprimento em uma seqüência de banco de dados. T é referido como o limiar de classificação de palavra de vizinhança (Altschul *et al*, *supra*). Estes acertos de palavras de vizinhança iniciais atuam como  
25 sementes para iniciar pesquisas para encontrar HSPs mais longas contendo as mesmas. Os acertos de palavras são estendidos em ambas as direções ao longo de cada seqüência enquanto a classificação de alinhamento cumulativo pode ser incrementada. Classificações cumulativas são calculadas usando, para seqüências de nucleotídeos, os parâmetros M (classificação de

recompensa para um par de radicais de pareamento; sempre  $> 0$ ) e N (classificação de penalidade para radicais malpareados; sempre  $< 0$ ). Para seqüências de aminoácidos, usa-se uma matriz de classificação para calcular a classificação cumulativa. Extensão dos acertos de palavras em cada direção são interrompidas quando: a classificação cumulativa de alinhamento diminui pela quantidade X de seu valor máximo obtido; a classificação cumulativa atinge zero e até menos, devido ao acúmulo de um ou mais alinhamentos de radicais de classificação negativa; ou atinge-se o fim de qualquer das seqüências. Os parâmetros de algoritmo BLAST[:] W, T, e X determinam a sensibilidade e velocidade do alinhamento. O programa BLASTN (para seqüências de nucleotídeos) usa como defaults um comprimento de palavra (W) de 11, uma expectativa (E) ou 10, M=5, N=-4 e uma comparação de ambos os filamentos. Para seqüências de aminoácidos, o programa BLASTP usa como defaults um comprimento de palavra de 3, e expectativa (E) de 10, e a matriz de classificação BLOSUM62 (ver Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alinhamentos (B) de 50, expectativa (E) de 10, M=5, N=-4, e uma comparação de ambos os filamentos.

O algoritmo BLAST também realiza uma análise estatística da similaridade entre duas seqüências (ver, p. ex., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). Uma medida da similaridade proporcionada pelo algoritmo BLAST é a menor probabilidade de soma (P(N)), que proporciona uma indicação da probabilidade com que poderia ocorrer, por mero acaso, uma equiparação entre duas seqüências de nucleotídeos ou aminoácidos. Por exemplo, um ácido nucleico é considerado similar a uma seqüência de referência se a menor probabilidade de soma em uma comparação do ácido nucleico de teste com o ácido nucleico de referência é menor do que cerca de 0,2, mais preferivelmente, menor do que cerca de 0,01, e, da forma mais preferível, menor do que cerca de 0,001.

Composições de polinucleotídeos

Como usado aqui, os termos "segmento de DNA" e "polinucleotídeo" referem-se a uma molécula de DNA que foi isolada livre de DNA genômico total de uma espécie particular. Portanto, uma seqüência de DNA que codifica um polipeptídeo refere-se a um segmento de DNA que contém uma ou mais seqüências codificantes embora seja possível isolar substancialmente, ou livrar por purificação, DNA genômico total das espécies de que se obtém o segmento de DNA. Inclui-se nos termos "segmento de DNA" e "polinucleotídeo" segmentos de DNA e fragmentos menores de referidos segmentos, e também vetores recombinantes, incluindo, por exemplo, plasmídeos, cosmídeos, fagomídeos, fago, vírus, e análogos.

Como o compreenderão aqueles versados na arte, os segmentos de DNA desta invenção podem incluir seqüências genômicas, seqüências extra-genômicas e seqüências codificadas por plasmídeo e segmentos menores de genes manipulados que expressam, ou que podem ser adaptadas para expressar, proteínas, polipeptídeos, peptídeos e análogos. Referidos segmentos podem ser naturalmente isolados, ou modificados sinteticamente pela mão humana.

"Isolado", como usado aqui, significa que um polinucleotídeo encontra-se substancialmente afastado de outras seqüências codificantes, e que o segmento de DNA não contém grandes porções de DNA codificante não-relacionado, como grandes fragmentos cromossômicos ou outros genes funcionais ou regiões codificadoras de polipeptídeo. Evidentemente, isto refere-se ao segmento de DNA como isolado originalmente, e não exclui genes ou regiões codificantes adicionadas mais tarde no segmento pela mão humana.

Como o reconhecerá a pessoa versada na arte, polinucleotídeos podem ser de filamento simples (codificando ou anti-sentido) ou de filamento duplo, e podem ser moléculas de DNA (genômico, cDNA ou sintético) ou RNA. Moléculas de RNA incluem moléculas de HnRNA, que contêm íntrons

e correspondem a uma molécula de DNA de uma maneira um-para-um, e moléculas de mRNA, que não contêm íntrons. Seqüências codificantes ou não-codificantes adicionais podem, embora não necessariamente, estar presentes em um polinucleotídeo da presente invenção, e um polinucleotídeo  
5 pode, embora não necessariamente, estar ligado a outras moléculas e/ou materiais de suporte.

Polinucleotídeos podem compreender uma seqüência nativa (i.e., uma seqüência endógena que codifica um antígeno de *Mycobacterium* ou uma porção do mesmo) ou podem compreender uma variante, ou um  
10 equivalente funcional biológico ou antigênico de uma tal seqüência. Variantes de polinucleotídeos podem conter uma ou mais substituições, adições, deleções e/ou inserções, como descrito adicionalmente abaixo, de preferência, de tal forma que a imunogenicidade do polipeptídeo codificado não seja  
15 diminuída relativamente a uma proteína de tumor nativa. O efeito sobre a imunogenicidade do polipeptídeo codificado pode ser avaliada geralmente como descrito aqui. O termo "variantes" também compreende genes homólogos de origem xenogênica.

Em concretizações adicionais, a presente invenção proporciona polinucleotídeos e polipeptídeos isolados compreendendo vários  
20 comprimentos de extensões contíguas de seqüência idênticas, ou complementares a uma ou mais das seqüências aqui divulgadas. Por exemplo, esta invenção proporciona polinucleotídeos que compreendem pelo menos cerca de 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 ou 1000 ou mais nucleotídeos contíguos de uma ou mais das seqüências aqui divulgadas e  
25 também todos os comprimentos intermediários entre os mesmos. Há de se compreender com facilidade que "comprimentos intermediários", neste contexto, significa qualquer comprimento entre os valores indicados, como 16, 17, 18, 19, etc.; 21, 22, 23, etc.; 30, 31, 32, etc.; 50, 51, 52, 53, etc.; 100, 101, 102, 103, etc.; 150, 151, 152, 153, etc.; incluindo todos os números

inteiros de 200 a 500; de 500 a 1.000, e análogos.

Os polinucleotídeos da presente invenção, ou fragmentos dos mesmos, independentemente do comprimento da própria seqüência codificante, podem ser combinados com outras seqüências de DNA, como  
5 promotores, sinais de poliadenilação, sítios adicionais de enzimas de restrição, múltiplos sítios de clonagem, outros segmentos codificantes, e análogos, de tal forma que seu comprimento global pode variar consideravelmente. Considera-se, portanto, que é possível empregar um fragmento de ácido nucleico de quase qualquer tamanho, sendo que o comprimento total é  
10 limitado, de preferência, pela facilidade de preparação e uso em um protocolo de DNA recombinante intencionado. Por exemplo, segmentos de DNA ilustrativos com comprimentos totais de cerca de 10.000, cerca de 5000, cerca de 3000, cerca de 2.000, cerca de 1.000, cerca de 500, cerca de 200, cerca de 100, cerca de 50 pares de bases de comprimento, e análogos, (incluindo todos  
15 os comprimentos intermediários) são considerados como sendo úteis em muitas implementações desta invenção.

Além disso, aqueles com prática ordinária na arte haverão de considerar que, como um resultado da degeneração do código genético, há muitas seqüências de nucleotídeos que codificar um polipeptídeo como  
20 descrito aqui. Alguns destes polinucleotídeos portam homologia mínima para a seqüência de nucleotídeos de qualquer gene nativo. No entanto, polinucleotídeos que variam devido a diferenças no uso de códon são considerados especificamente pela presente invenção, por exemplo polinucleotídeos que são otimizados para seleção de códon humano e/ou  
25 primata. Adicionalmente, alelos dos genes que compreendem as seqüências de polinucleotídeos aqui proporcionadas encontram-se dentro do escopo da presente invenção. Alelos são genes endógenos que são alterados como um resultado de uma ou mais mutações, como deleções, adições e/ou substituições de nucleotídeos. A proteína e mRNA resultante pode, embora



não necessariamente, apresentar uma função ou estrutura alterada. Alelos podem ser identificados usando-se técnicas convencionais (como hibridização, amplificação e/ou comparação de seqüências de bancos de dados).

## 5 Identificação e caracterização de polinucleotídeos

Polinucleotídeos podem ser identificados, preparados e/ou manipulados usando-se qualquer uma de uma variedade de técnicas bem estabelecidas. Por exemplo, um polinucleotídeo pode ser identificado, como descrito mais detalhadamente abaixo, por meio de seleção de um micro-  
10 conjunto de cDNAs para expressão associada com tumor (i.e., expressão que é pelo menos duas vezes maior em um tumor do que no tecido normal, como determinado usando-se um ensaio representativo proporcionado aqui). Referidos ensaios podem ser realizados, por exemplo, usando um micro-  
conjunto Synteni (Palo Alto, CA) de acordo com as instruções do fabricante  
15 (e essencialmente como descritas por Schena *et al*, Proc. Natl. Acad. Sci. USA 93:10614-10619 (1996) e Heller *et al*, Proc. Natl. Acad. Sci. USA 94:2150-2155 (1997)). Alternativamente, polinucleotídeos podem ser amplificados de cDNA preparado de células que expressam as proteínas aqui descritas, como células de *M. tuberculosis*. Referidos polinucleotídeos podem  
20 ser amplificados via reação em cadeia de polimerase (PCR). Para esta abordagem, é possível projetar iniciadores específicos para seqüência, com base nas seqüências aqui proporcionadas, e podem ser adquiridos ou sintetizados.

Uma porção amplificada de um polinucleotídeo da presente  
25 invenção pode ser usada para isolar um gene de comprimento pleno de uma biblioteca vantajosa (p. ex., uma biblioteca de cDNA de *M. tuberculosis*) usando-se técnicas bem conhecidas. Entre estas técnicas, uma biblioteca (de cDNA ou genômica) é selecionada usando-se uma ou mais sondas ou iniciadores de polinucleotídeo vantajosos para amplificação. De preferência,

uma biblioteca é selecionado de acordo com o tamanho de forma a incluir moléculas maiores. Bibliotecas iniciadas randomicamente também podem ser preferidas para identificar regiões a 5' e a montante dos genes. Bibliotecas genômicas são preferidas para obtenção de íntrons e extensão de seqüências a 5'.

Para técnicas de hibridização, uma seqüência parcial pode ser marcada (p. ex., por meio de tradução de pequena abrangência [*nick-translation*] ou marcação terminal com  $^{32}\text{P}$ ) usando-se técnicas bem conhecidas. Então, geralmente, uma biblioteca bacteriana ou de bacteriófagos é selecionada com filtros de hibridização contendo colônias bacterianas desnaturadas (ou superfícies parecendo gramado contendo placas de fagos) com a sonda marcada (ver Sambrook *et al*, *Molecular Cloning: A Laboratory Manual* (2000)). Placas ou colônias de hibridização são selecionadas e expandidas, e o DNA é isolado para análise posterior. Clones de cDNA podem ser analisados para se determinar a quantidade de seqüência adicional, por exemplo, por meio de PCR usando um iniciador da seqüência parcial e um iniciador do vetor. Mapas de restrição e seqüências parciais podem ser geradas para se identificar um ou mais clones que se superpõem. Em seguida, é possível determinar a seqüência completa usando-se técnicas convencionais, que podem envolver gerar uma série de clones de deleção. As seqüências superpostas resultantes podem, então, ser montadas formando uma única seqüência contígua. Uma molécula de cDNA de comprimento pleno pode ser gerada ligando-se fragmentos vantajosos, empregando-se técnicas bem conhecidas.

Alternativamente, há numerosas técnicas de amplificação para se obter uma seqüência codificante de comprimento pleno a partir de uma seqüência de cDNA parcial. Entre estas técnicas, a amplificação é geralmente realizada via PCR. É possível usar qualquer um de uma variedade de kits comercialmente obteníveis para realizar a etapa de amplificação. Iniciadores

podem ser projetados usando-se, por exemplo, software bem conhecido na arte. Iniciadores têm comprimentos, de preferência, de 22 a 30 nucleotídeos, apresentam um teor de GC de pelo menos 50% e recombina com a seqüência-alvo a temperaturas de cerca de 68°C a 72°C. A região amplificada  
5 pode ser seqüenciada como descrito acima, e seqüências que se superpõem são montadas formando uma seqüência contígua.

Uma técnica de amplificação do tipo referido é PCR invertida (ver Triglia *et al*, *Nucl. Acids Res.* 16:8186 (1988)), que usa enzimas de restrição para gerar um fragmento na região conhecida do gene. O fragmento  
10 é então circularizado por meio de ligação intramolecular e usado como um modelo para PCR com iniciadores divergentes derivados daquela região conhecida. Em uma abordagem alternativa, seqüências adjacentes a uma seqüência parcial podem ser recuperadas por meio de amplificação com um iniciador a uma seqüência ligante e um iniciador específico para uma região  
15 conhecida. As seqüências amplificadas são submetidas tipicamente a uma segunda operação de amplificação com o mesmo iniciador ligante e um segundo iniciador específico para a região conhecida. Uma variação deste procedimento, que emprega dois iniciadores que iniciam extensão em direções opostas da seqüência conhecida, é descrita no WO 96/38591. Outra  
20 técnica do tipo referido é conhecida como "amplificação rápida de extremidades de cDNA" ou RACE. Esta técnica envolve o uso de um iniciador interno e um iniciador externo, que hibrida com uma região de poliA região ou seqüência de vetor, para identificar seqüências que se encontram a 5' e 3' de uma seqüência conhecida. Técnicas adicionais incluem PCR de  
25 captura (Lagerstrom *et al*, *PCR Methods Applic.* 1:111-19 (1991)) e percurso de PCR (Parker *et al*, *Nucl. Acids. Res.* 19:3055-60 (1991)). Outros métodos que empregam amplificação também podem ser empregados para se obter uma seqüência de cDNA de comprimento pleno.

Em determinados casos, é possível obter uma seqüência de

cDNA de comprimento pleno por meio de análise de seqüências proporcionadas em um banco de dados de marcador de seqüência expresso (EST), como aquele obtível junto ao GenBank. Buscas por ESTs que se superpõem podem ser realizadas geralmente usando-se programas bem conhecidos (p. ex., buscas NCBI BLAST), e referidos ESTs podem ser usados para gerar uma seqüência contígua de comprimento pleno. Seqüências de DNA de comprimento pleno também podem ser obtidas por meio de análise de fragmentos genômicos.

#### Expressão de polinucleotídeo em células hospedeiras

10 Em outras concretizações da invenção, seqüências de polinucleotídeos ou fragmentos das mesmas que codificam polipeptídeos da invenção, ou proteínas de fusão ou seus equivalentes funcionais, podem ser usadas em moléculas de DNA recombinante para dirigir a expressão de um polipeptídeo em células hospedeiras apropriadas. Devido à degeneração  
15 inerente do código genético, é possível produzir outras seqüências de DNA que codificam substancialmente a mesma seqüência de aminoácidos ou uma seqüência de aminoácidos equivalente, e estas seqüências podem ser usadas para clonar e expressar um dado polipeptídeo.

20 Como o compreenderão aqueles com prática na arte, pode ser vantajoso, em alguns casos, produzir seqüências de nucleotídeos que codificam polipeptídeo apresentando códons não-naturalmente ocorrentes. Por exemplo, códons preferidos por um hospedeiro eucariótico ou procariótico particular podem ser selecionados de forma a incrementar a taxa de expressão de proteína ou produzir uma transcrição de RNA recombinante  
25 apresentando propriedades desejáveis, como uma meia-vida que é mais longa do que aquela de uma transcrição gerada de uma seqüência naturalmente ocorrente.

Além disso, as seqüências de polinucleotídeos da presente invenção podem ser manipuladas usando-se métodos geralmente conhecidos

na arte para alterar seqüências codificantes de polipeptídeo por diversas razões, incluindo, embora sem limitação, alterações que modificam a clonagem, o processamento e/ou a expressão do produto gênico. Por exemplo, é possível usar embaralhamento de DNA por meio de fragmentação 5 randômica e remontagem de PCR de fragmentos de genes e oligonucleotídeos sintéticos para manipular as seqüências de nucleotídeos. Adicionalmente, é possível usar mutagênese direcionada para sítio para inserir novos sítios de restrição, alterar padrões de glicosilação, alterar preferência de códon, produzir variantes de recomposição, ou introduzir mutações, e assim por 10 diante.

Em outra concretização da invenção, é possível ligar seqüências de ácido nucleico naturais, modificadas, ou recombinantes a uma seqüência heteróloga para codificar uma proteína de fusão. Por exemplo, para 15 selecionar bibliotecas de peptídeos quanto a inibidores da atividade de polipeptídeo, pode ser útil codificar uma proteína quimérica que pode ser reconhecida por um anticorpo comercialmente obtível. Uma proteína de fusão também pode ser manipulada de forma a conter um sítio de clivagem localizado entre a seqüência codificante de polipeptídeo e a seqüência de proteína heteróloga, de tal forma que o polipeptídeo pode ser clivado e 20 purificado sendo removido da porção heteróloga.

Seqüências que codificam um polipeptídeo desejado podem ser sintetizadas, no todo ou em parte, usando-se métodos químicos bem conhecidos na arte (ver Caruthers, M. H. *et al*, *Nucl. Acids Res. Symp. Ser.* pp. 215-223 (1980), Horn *et al*, *Nucl. Acids Res. Symp. Ser.* pp. 225-232 (1980)).

25 Alternativamente, a proteína propriamente dita pode ser produzida usando-se métodos químicos para sintetizar a seqüência de aminoácidos de um polipeptídeo, ou uma porção do mesmo. Por exemplo, é possível realizar síntese de peptídeo empregando-se várias técnicas de estado sólido (Roberge *et al.*, *Science* 269:202-204 (1995)), e é possível obter síntese

automatizada, por exemplo, usando-se o sintetizador de peptídeo ABI 431A (Perkin Elmer, Palo Alto, CA).

Um peptídeo recentemente sintetizado pode ser purificado substancialmente por meio de cromatografia líquida preparativa de alto desempenho (p. ex., Creighton, *Proteins, Structures and Molecular Principles* (1983)) ou outras técnicas comparáveis na arte. A composição dos peptídeos sintéticos pode ser confirmada por meio de análise ou seqüenciamento de aminoácidos (p. ex., o procedimento de degradação de Edman). Adicionalmente, a seqüência de aminoácidos de um polipeptídeo, ou qualquer parte da mesma, pode ser alterada durante síntese direta e/ou combinada usando-se métodos químicos com seqüências de outras proteínas, ou qualquer parte das mesmas, para produzir um polipeptídeo variante.

Para expressar um polipeptídeo desejado, as seqüências de nucleotídeos que codificam o polipeptídeo, ou equivalentes funcionais, podem ser inseridas em vetor de expressão apropriado, i.e., um vetor que contém os elementos necessários para a transcrição e tradução da seqüência codificante inserida. Métodos que são bem conhecidos por aqueles com prática na arte podem ser usados para construir vetores de expressão contendo seqüências que codificam um polipeptídeo de interesse e elementos de controle transcricional e de tradução apropriados. Estes métodos incluem técnicas de DNA recombinante *in vitro*, técnicas sintéticas, e recombinação genética *in vivo*. Referidas técnicas são descritas em Sambrook *et al*, *Molecular Cloning, A Laboratory Manual* (2000), e Ausubel, *et al.*, *Current Protocols in Molecular Biology* (atualizado anualmente).

Uma variedade de sistemas de hospedeiro/vetor de expressão pode ser usada para conter e expressar seqüências de polinucleotídeos. Estes incluem, embora sem limitação, microorganismos, como bactérias transformadas com vetores de expressão de DNA recombinante de bacteriófago, plasmídeo, ou cosmídeo; levedura transformada com vetores de

expressão de levedura; sistemas de células de inseto com vetores de expressão de vírus (p. ex., baculovírus); sistemas de células de plantas com vetores de expressão de vírus (p. ex., vírus do mosaico da couve-flor, CaMV; vírus do mosaico do tabaco, TMV) ou com vetores de expressão bacterianos (p. ex., plasmídeos Ti ou pBR322); ou sistemas de células animais.

Os "elementos de controle" ou "seqüências reguladoras" presentes em um vetor de expressão são aquelas regiões não-traduzidas dos acentuadores de vetor, promotores, regiões não traduzidas a 5' e 3' que interagem com proteínas celulares do hospedeiro para realizar transcrição e tradução. Referidos elementos podem variar quanto à intensidade e especificidade. Dependendo do sistema de vetor e hospedeiro usado, é possível usar qualquer número de elementos de transcrição e tradução vantajosos, incluindo promotores constitutivos e induzíveis. Por exemplo, quando se clona em sistemas bacterianos, é possível usar promotores induzíveis, como o promotor lacZ híbrido do fagomídeo PBLUESCRIPT (Stratagene, La Jolla, Calif.) ou plasmídeo PSPORT1 (Gibco BRL, Gaithersburg, MD) e análogos. Em sistemas de células mamíferas, prefere-se geralmente promotores de genes mamíferos ou de vírus mamíferos. É necessário gerar uma linha de células que contém múltiplas cópias da seqüência que codifica um polipeptídeo, vetores baseados em SV40 ou EBV podem ser usados vantajosamente com um marcador selecionável apropriado.

Em sistemas bacterianos, é possível selecionar uma quantidade de vetores de expressão dependendo do uso intencionado para o polipeptídeo expresso. Por exemplo, quando se necessita grandes quantidades, por exemplo para a indução de anticorpos, é possível usar vetores que dirigem a expressão de alto nível de proteínas de fusão que são facilmente purificadas. Referidos vetores incluem, embora sem limitação, os vetores de expressão e clonagem de *E. coli* multifuncional, como BLUESCRIPT (Stratagene), em que a seqüência que codifica o polipeptídeo de interesse pode ser ligada no vetor

dentro da matriz com seqüências para o Met amino-terminal e os 7 radicais subseqüentes de  $\beta$ -galactosidase de forma a produzir uma proteína híbrida; vetores pIN (Van Heeke & Schuster, *J. Biol. Chem.* 264:5503-5509 (1989)); e análogos. Também é possível usar vetores pGEX (Promega, Madison, Wis.) para expressar polipeptídeos estranhos como proteínas de fusão com glutathione S-transferase (GST). De uma maneira geral, referidas proteínas de fusão são solúveis e podem ser purificadas facilmente a partir de células lisadas por meio de adsorção em pérolas de glutathione-agarose, seguido de eluição na presença de glutathione livre. Proteínas preparadas nesses sistemas podem ser projetadas de forma a incluir heparina, trombina, ou sítios de clivagem de protease de fator XA de tal forma que o polipeptídeo clonado de interesse possa ser liberado da porção GST conforme desejado.

Na levedura, *Saccharomyces cerevisiae*, é possível usar vários dos vetores contendo promotores constitutivos ou induzíveis, como fator alfa, oxidase de álcool, e PGH. Para revisões, ver Ausubel *et al* (supra) e Grant *et al*, *Methods Enzymol.* 153:516-544 (1987).

Em casos em que se usa vetores de expressão de planta, a expressão de seqüências que codificam polipeptídeos pode ser impelida por qualquer um de uma variedade de promotores. Por exemplo, promotores virais, como os promotores 35S e 19S do CaMV, podem ser usados sozinhos ou em combinação com a seqüência líder ômega do TMV (Takamatsu, *EMBO J.* 6:307-311 (1987)). Alternativamente, é possível usar promotores de planta, como os promotores de choque com calor ou de pequena subunidade de RUBISCO (Coruzzi *et al.*, *EMBO J.* 3:1671-1680 (1984); Broglie *et al*, *Science* 224:838-843 (1984); e Winter *et al*, *Results Probl. Cell Differ.* 17:85-105 (1991)). Estas construções podem ser introduzidas em células de plantas por meio de transformação direta de DNA ou transfecção mediada com patógeno. Referidas técnicas encontram-se descritas em uma variedade de revistas geralmente obteníveis (ver, p. ex., Hobbs em McGraw Hill *Yearbook*



*of Science e Technology*, pp. 191-196 (1992)).

Também é possível usar um sistema de inseto para expressar um polipeptídeo de interesse. Por exemplo, em um sistema do tipo referido, usa-se vírus da poliedrose nuclear de *Autographa californica* (AcNPV) como um vetor para expressar genes estranhos em células de *Spodoptera frugiperda* ou em larvas de *Trichoplusia*. As seqüências que codificam o polipeptídeo podem ser clonadas em uma região não-essencial do vírus, como o gene de poliedrina, e colocado sob o controle do promotor de poliedrina. A inserção bem sucedida da seqüência codificante de polipeptídeo tornará o gene de poliedrina inativo e produzirá vírus recombinante que não apresenta proteína de revestimento. Os vírus recombinantes podem ser então usados para infectar, por exemplo, células de *S. frugiperda* ou larvas de *Trichoplusia* em que o polipeptídeo de interesse pode ser expresso (Engelhard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 91:3224-3227 (1994)).

Em células hospedeiras mamíferas, diversos sistemas de expressão à base de vírus encontram-se geralmente disponíveis. Por exemplo, em casos em que se usa um adenovírus como um vetor de expressão, seqüências que codificam um polipeptídeo de interesse podem ser ligadas em um complexo de tradução/transcrição de adenovírus consistindo do promotor tardio e seqüência líder tripartite. Inserção em uma região não-essencial E1 ou E3 do genoma viral pode ser usada para se obter um vírus viável que é capaz de expressar o polipeptídeo em células hospedeiras infectadas (Logan & Shenk, *Proc. Natl. Acad. Sci. U.S.A.* 81:3655-3659 (1984)). Adicionalmente, é possível usar acentuadores de transcrição, como o acentuador do vírus do sarcoma de Rous (RSV, *Rous Sarcoma Virus*), para incrementar a expressão em células hospedeiras mamíferas. Métodos e protocolos para operar com vetores de adenovírus são revistos em Wold, *Adenovirus Methods and Protocols*, 1998. Referência adicional com relação ao uso de vetores de adenovírus pode ser encontrada em *Adenovirus: A Medical Dictionary*,

*Bibliography, e Annotated Research Guide to Internet References, 2004.*

Também é possível usar sinais de iniciação específicos para obter tradução mais eficiente de seqüências codificando um polipeptídeo de interesse. Referidos sinais incluem o códon de iniciação ATG e seqüências adjacentes. Em casos em que seqüências que codificam o polipeptídeo, seu códon de iniciação, e seqüências a montante são inseridos no vetor de expressão apropriado, podem não ser necessários sinais de controle de transcrição ou tradução adicionais. No entanto, em casos em que apenas a seqüência codificante, ou uma porção da mesma, é inserida, deveria-se proporcionar sinais de controle de tradução exógenos incluindo o códon de iniciação ATG. Adicionalmente, o códon de iniciação deveria encontrar-se na matriz de leitura aberta para assegurar a tradução do inserto inteiro. Códon de iniciação e elementos de tradução exógenos podem ser de várias origens, tanto naturais como sintéticas. A eficiência de expressão pode ser incrementada por meio da inclusão dos acentuadores que são apropriados para o sistema de células particular usado, como aqueles descritos na literatura (Scharf. *et al, Results Probl. Cell Differ.* 20:125-162 (1994)).

Adicionalmente, é possível selecionar uma cepa de células hospedeiras quanto a sua capacidade de modular a expressão das seqüências inseridas, ou de processar a proteína expressa da maneira desejada. Referidas modificações do polipeptídeo incluem, embora sem limitação, acetilação, carboxilação, glicosilação, fosforilação, lipidação, e acilação. Processamento pós-tradução que cliva uma forma "prepro" da proteína também pode ser usado para facilitar a correta inserção, dobragem e/ou função. É possível selecionar diferentes células hospedeiras, como CHO, HeLa, MDCK, HEK293, e WI38, que apresentam maquinaria celular específica e mecanismos característicos para referidas atividades pós-tradução, para assegurar a modificação e o processamento correto da proteína estranha.

Para a produção de longo prazo, alto desempenho de proteínas

recombinantes, prefere-se geralmente a expressão estável. Por exemplo, linhas de células que expressam de forma estável um polinucleotídeo de interesse podem ser transformadas usando-se vetores de expressão que podem  
5 um gene marcador selecionável no mesmo vetor ou em um vetor separado. Após a introdução do vetor, células podem ser deixadas desenvolver durante de 1 a 2 dias em um meio enriquecido antes de serem trocadas para meio seletivo. A finalidade do marcador selecionável consiste em conferir resistência à seleção, e sua presença permite o crescimento e a recuperação de  
10 células que expressam com êxito as seqüências introduzidas. Clones resistentes de células transformadas de forma estável podem ser proliferados usando-se técnicas de cultura de tecidos apropriadas para o tipo de célula.

É possível usar qualquer quantidade de sistemas de seleção para recuperar linhas de células transformadas. Estas incluem, embora sem  
15 limitação, genes da timidina quinase do vírus do herpes simplex (Wigler *et al*, *Cell* 11:223-32 (1977)) e adenina fosforribosiltransferase (Lowy *et al*, *Cell* 22:817-23 (1990)) que podem ser empregados em células tk.sup. ou aprt.sup., respectivamente. Da mesma forma, é possível usar a resistência a herbicidas, antibióticos ou antimetabólitos como a base para seleção; por exemplo, dhfr  
20 que confere resistência ao metotrexato (Wigler *et al*, *Proc. Natl. Acad. Sci. U.S.A.* 77:3567-70 (1980)); npt, que confere resistência a aminoglicosídeos, neomicina e G-418 (Colbere-Garapin *et al*, *J. Mol. Biol.* 150:1-14 (1981)); e als ou pat, que conferem resistência ap clorossulfuron e fosfinotricina acetiltransferase, respectivamente (Murry, *supra*). Genes selecionáveis  
25 adicionais já foram descritos, por exemplo, trpB, que permite às células usar indol em lugar de triptofano, ou hisD, que permite às células usar histinol em lugar de histidina (Hartman & Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* 85:8047-51 (1988)). Recentemente, o uso de marcadores visíveis ganhou em popularidade com marcadores, como antocianinas,  $\beta$ -glucuronidase e seu

substrato GUS, e luciferase e seu substrato luciferina, sendo amplamente usados não só para identificar transformantes, mas também para quantificar a quantidade de expressão de proteína transiente ou estável atribuível a um sistema de vetor específico (Rhodes *et al*, *Methods Mol. Biol.* 55:121-131 (1995)).

Embora a presença/ausência de expressão de gene marcador sugira que o gene de interesse também esteja presente, sua presença e expressão pode ainda precisar ser confirmada. Por exemplo, se a seqüência que codifica um polipeptídeo é inserida no interior de uma seqüência de gene marcador, células recombinantes contendo seqüências podem ser identificadas através da ausência de função de gene marcador. Alternativamente, um gene marcador pode ser colocado em tandem com uma seqüência codificante de polipeptídeo sob o controle de um único promotor. A expressão do gene marcador em resposta à indução ou seleção também indica usualmente a expressão do gene em tandem.

Alternativamente, células hospedeiras que contêm e expressam uma seqüência de polinucleotídeos desejada podem ser identificadas por meio de uma variedade de procedimentos conhecidos por aqueles com prática na arte. Estes procedimentos incluem, embora sem limitação, hibridizações DNA-DNA ou DNA-RNA e bio-ensaio de proteína ou técnicas de ensaio imunológico que incluem tecnologias baseadas em membrana, solução, ou chip para a detecção e/ou quantificação de ácido nucleico ou proteína.

Conhece-se na arte uma variedade de protocolos para detectar e medir a expressão de produtos codificados por polinucleotídeos, usando-se anticorpos policlonais ou monoclonais específicos para o produto. Exemplos incluem ensaio imunossorvente ligado a enzima (ELISA), ensaio rádio-imunológico (RIA), e seleção de células ativada com fluorescência (FACS). Um ensaio imunológico de dois sítios, em base monoclonal, empregando anticorpos monoclonais reativos para dois epitopos não-interferentes em um

5 dado polipeptídeo pode ser preferido para algumas aplicações, mas também é possível empregar um ensaio de ligação competitiva. Estes e outros ensaios encontram-se descritos, entre outros locais, em Hampton *et al*, *Serological Methods, a Laboratory Manual* (1990) e Maddox *et al*, *J. Exp. Med.* 158:1211-1216 (1983).

10 Aquelles com prática na arte conhecem uma ampla variedade de marcadores e técnicas de conjugação que podem ser usadas em vários ensaios de ácido nucleico e aminoácidos. Meios para produzir hibridização marcada ou sondas de PCR para detectar seqüências relacionadas com polinucleotídeos incluem oligo- marcação, tradução de pequena abrangência [*nick-translation*] tradução, marcação terminal ou amplificação com PCR usando-se um nucleotídeo marcado. Alternativamente, as seqüências, ou quaisquer porções das mesmas podem ser clonadas em um vetor para a produção de uma sonda de mRNA. Referidos vetores são conhecidos na arte, 15 são comercialmente obteníveis, e podem ser usados para sintetizar sondas de RNA *in vitro* por meio de adição de uma RNA polimerase apropriada, como T7, T3, ou SP6 e nucleotídeos marcados. Estes procedimentos podem ser conduzidos empregando-se uma variedade de kits comercialmente obteníveis. Marcadores ou moléculas repórter vantajosas que podem ser usadas incluem 20 radionuclídeos, enzimas, agentes fluorescentes, quimioluminescentes, ou cromogênicos, e também substratos, co-fatores, inibidores, partículas magnéticas, e análogos.

25 Células hospedeiras transformadas com uma seqüência de polinucleotídeos de interesse podem ser cultivadas em condições vantajosas para a expressão e a recuperação da proteína de cultura de células. A proteína produzida por uma célula recombinante pode ser secretada ou contida intracelularmente dependendo da seqüência e/ou do vetor usado. Como o compreenderão aqueles com prática na arte, vetores de expressão contendo polinucleotídeos da invenção podem ser projetados de forma a contar

seqüências-sinal que dirigem a secreção do polipeptídeo codificado através de uma membrana de células eucarióticas ou procarióticas. É possível usar outras construções recombinantes para conjugar seqüências que codificam um polipeptídeo de interesse a seqüência de nucleotídeos que codifica um domínio de polipeptídeo que facilitará a purificação de proteínas solúveis. Referidos domínios facilitadores de purificação incluem, embora sem limitação, peptídeos quelantes de metal, como módulos de histida-triptofano que permitem purificação em metais imobilizados, domínios de proteína A que permitem purificação em imunoglobulina imobilizada, e o domínio usado no sistema de purificação por afinidade/extensão FLAGS (Immunex Corp., Seattle, Wash.). A inclusão de seqüências de ligantes cliváveis, como aquelas específicas para Fator XA ou enteroquinase (Invitrogen. San Diego, Calif.) entre o domínio de purificação e o polipeptídeo codificado pode ser usada para facilitar a purificação. Um vetor de expressão do tipo referido proporciona a expressão de uma proteína de fusão contendo um polipeptídeo de interesse e um ácido nucleico codificando 6 radicais histidina precedendo uma tiorredoxina ou um sítio de clivagem de enteroquinase. Os radicais histidina facilitam a purificação em EvIIAC (cromatografia de afinidade de íon de metal imobilizado) como descrito em Porath *et al*, *Prot. Exp. Purif.* 3:263-281 (1992) enquanto que o sítio de clivagem de enteroquinase proporciona um meio para purificar o polipeptídeo desejado da proteína de fusão. Uma discussão de vetores que contêm proteínas de fusão é proporcionada em Kroll *et al*, *DNA Cell Biol.* 12:441-453 (1993)).

Adicionalmente a métodos de produção recombinante, polipeptídeos da invenção, e fragmentos dos mesmos, podem ser produzidos por meio de síntese direta de peptídeos usando-se técnicas de fase sólida (Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963)). A síntese de proteína pode ser realizada empregando-se técnicas manuais ou automação. A síntese automatizada pode ser obtida, por exemplo, usando-se o sintetizador de

peptídeos Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternativamente, diversos fragmentos podem ser sintetizados separadamente por via química e combinados usando-se métodos químicos para produzir a molécula de comprimento pleno.

## 5 Técnicas de fornecimento de polinucleotídeo *in vivo*

Em concretizações adicionais, introduz-se construções genéticas compreendendo um ou mais dos polinucleotídeos da invenção nas células *in vivo*. Isto pode ser obtido empregando-se qualquer uma de uma variedade de abordagens bem conhecidas, sendo que várias destas encontram-se delineadas abaixo com o objetivo de ilustração.

### 1. Adenovírus

Um dos métodos preferidos para fornecimento *in vivo* de uma ou mais seqüências de ácido nucleico envolve o uso de um vetor de expressão de adenovírus. "Vetor de expressão de adenovírus" deve incluir aquelas construções contendo seqüências de adenovírus suficientes para (a) suportar empacotamento da construção e (b) expressar um polinucleotídeo que foi clonado ali em uma orientação no sentido ou anti-sentido. Evidentemente, no contexto de uma construção anti-sentido, a expressão não exige que o produto gênico seja sintetizado.

O vetor de expressão compreende uma forma geneticamente manipulada de um adenovírus. Conhecimento da organização genética de adenovírus, um vírus de DNA de filamento duplo, linear, de 36 kb, permite a substituição de grandes peças de DNA adenoviral com seqüências estranhas com até 7 kb (Grunhaus & Horwitz, 1992). Em contraste ao retrovírus, a infecção adenoviral de células hospedeiras não resulta em integração cromossômica porque DNA adenoviral pode replicar em integração cromossômica porque DNA adenoviral pode replicar de uma maneira epissômica sem genotoxicidade potencial. Da mesma forma, adenovírus são estruturalmente estáveis, e não se detectou rearranjo do genoma após

amplificação extensiva. Adenovírus pode infectar virtualmente todas as células epiteliais independentemente do seu estágio de ciclo celular. Até aqui, a infecção adenoviral parece estar relacionada apenas com doença branda, como doença respiratória aguda em humanos.

5 Adenovírus é particularmente vantajoso para uso como um vetor de transferência de gene devido a seu genoma de tamanho médio, facilidade de manipulação, alta titulação, ampla faixa de células-alvo e elevada infectividade. Ambas as extremidades do genoma viral contêm repetições invertidas (ITRs) com de 100 a 200 pares de bases, que são  
10 elementos *cis* necessários para replicação e empacotamento de DNA viral. As regiões precoce (E) e tardia (L) do genoma contêm diferentes unidades de transcrição que são divididas pelo início da replicação de DNA viral. A região E1 (E1A e E1B) codifica proteínas responsáveis pela regulação de transcrição do genoma viral e alguns genes celulares. A expressão da região E2 (E2A e  
15 E2B) resulta na síntese das proteínas para replicação de DNA viral. Estas proteínas estão envolvidas em replicação de DNA, expressão tardia de gene e desativação da célula hospedeira (Renan, 1990). Os produtos dos genes tardios, incluindo a maior parte das proteínas de capsídeo viral, são expressas apenas após processamento significativo de uma única transcrição primária  
20 provocada pelo promotor tardio principal (MLP). O MLP, (localizado em 16.8 m.u.) é particularmente eficiente durante a fase tardia da infecção, e todos os mRNAs provenientes deste promotor apresentam uma seqüência líder tripartite a 5' (TPL) que os torna mRNAs preferidos para tradução.

25 Em um sistema corrente, adenovírus recombinante é gerado a partir de recombinante homóloga entre vetor de transporte e vetor de pró-vírus. Devido à possível recombinação entre dois vetores pró-virais, adenovírus de tipo selvagem pode ser gerado com este processo. Portanto, é crítico isolar um único clone de vírus de uma placa individual e examinar sua estrutura genômica.



A geração e propagação dos vetores de adenovírus correntes, que são deficientes de replicação, depende de uma única linha de células auxiliares, designada 293, que foi transformada de células de rim embrionário humano por meio de fragmentos de DNA Ad5 e expressa constitutivamente proteínas E1 (Graham *et al.*, 1977). Como a região E3 é dispensável do genoma de adenovírus (Jones & Shenk, 1978), os correntes vetores de adenovírus, com o auxílio de células 293, carregam DNA estranho na região E1, na região D3 ou ambas as regiões (Graham & Prevec, 1991). Na natureza, o adenovírus pode empacotar aproximadamente 105% do genoma de tipo selvagem (Ghosh-Choudhury *et al.*, 1987), fornecendo capacidade para cerca de 2 kB extras de DNA. Combinada com os aproximadamente 5,5 kB de DNA que é substituível nas regiões E1 e E3, a capacidade máxima do vetor de adenovírus corrente é inferior a 7,5 kB, ou cerca de 15% do comprimento total do vetor. Mais de 80% do genoma viral de adenovírus permanecem na espinha dorsal do vetor e é a fonte de citotoxicidade transmitida por vetor. Da mesma forma, a deficiência de replicação do vírus apresentando deleção de E1 é incompleta. Por exemplo, observou-se vazamento da expressão de gene viral com os vetores correntemente disponíveis com altas multiplicidades de infecção (MOI) (Mulligan, 1993).

Linhas de células auxiliares podem ser derivadas de células humanas, como células de rim embrionário humano, células musculares, células hematopoiéticas ou outras células epiteliais ou mesenquimais embrionárias humanas. Alternativamente, as células auxiliares podem ser derivadas das células de outras espécies mamíferas que são permissivas para adenovírus humano. Referidas células incluem, p. ex., células Vero ou outras células epiteliais ou mesenquimais embrionárias de macaco. Como indicado acima, a linha de células auxiliares correntemente preferida é 293.

Recentemente, Racher *et al.* (1995) divulgaram métodos aperfeiçoados para cultivar células 293 e propagar adenovírus. Em um

formato, agregados de células naturais são desenvolvidos por meio de inoculação de células individuais em frascos de centrifugação siliconados com volume de 1 litro (Techne, Cambridge, UK) contendo de 100 a 200 ml de meio. Após agitação a 40 rpm, a viabilidade de células é estimada com azul  
5 tripano. Em outro formato, emprega-se micro-veículos Fibra-Cel (Bibby Sterlin, Stone, UK) (5 g/l) como a seguir. Adiciona-se um inóculo de células, ressuspensas em 5 ml de meio, ao veículo (50 ml) em um frasco de Erlenmeyer de 250 ml e deixados estacionários, com agitação ocasional, durante de 1 a 4 h. Em seguida, o meio é substituído por 50 ml de meio fresco  
10 e inicia-se a agitação. Para a produção de vírus, células são deixadas desenvolver a cerca de 80% de confluência, sendo que após isto o meio é substituído (a 25% do volume final) e adiciona-se adenovírus a uma MOI de 0,05. Culturas são deixadas estacionárias de um dia para o outro, após o que o volume é incrementado a 100% e inicia-se a agitação durante mais 72 h.

15 Diferentemente da exigência de que o vetor de adenovírus vetor seja defectivo para replicação, ou pelo menos condicionalmente defectivo, não se acredita que a natureza do vetor de adenovírus seja crucial para a prática bem sucedida da invenção. O adenovírus pode ser qualquer um dos 42 diferentes sorotipos conhecidos ou subgrupos de A a F. Adenovírus de  
20 tipo 5 de subgrupo C é o material de partida preferido para se obter um vetor de adenovírus de replicação defectiva condicional para uso na presente invenção, porque Adenovírus de tipo 5 é uma adenovírus humano sobre o qual se conhece grande quantidade de informação bioquímica e genética, e tem sido usado historicamente para a maior parte das construções empregando  
25 adenovírus como um vetor.

Como indicado acima, o vetor típico de acordo com a presente invenção é defectivo para replicação e não apresentará uma região de adenovírus E1. Assim, será muito conveniente introduzir o polinucleotídeo que codifica o gene de interesse na posição de que se removeu as seqüências

codificantes de E1. No entanto, a posição de inserção da construção no interior das seqüências de adenovírus não é crítica para a invenção. O polinucleotídeo que codifica o gene de interesse também pode ser inserido em lugar da região E3 deletada nos vetores de substituição de E3 como descrito por Karlsson *et al.* (1986) ou na região E4 em que uma linha de células auxiliares ou vírus auxiliar complementa o defeito de E4.

Adenovírus é fácil de desenvolver e manipular e apresenta ampla faixa de hospedeiros *in vitro* e *in vivo*. Este grupo de vírus pode ser obtido com altas titulações, p. ex., de  $10^9$  a  $10^{11}$  unidades formadoras de placa por ml, e podem ser altamente infectivos. O ciclo de vida de adenovírus não requer integração no genoma da célula do hospedeiro. Os genes estranhos fornecidos por vetores de adenovírus são epissômicos e, portanto, apresentam baixa genotoxicidade com relação a células hospedeiras. Não se reportou efeitos colaterais em estudos de vacinação com adenovírus de tipo selvagem (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrando sua segurança e potencial terapêutico como vetores de transferência de genes *in vivo*.

Usou-se vetores de adenovírus na expressão de genes eucarióticos (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) e desenvolvimento de vacina (Grunhaus & Horwitz, 1992; Graham & Prevec, 1992). Recentemente, estudos animais sugeriram que seria possível usar adenovírus recombinante para terapia gênica (Stratford-Perricaudet & Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Estudos envolvendo a administração de adenovírus recombinante para tecidos diferentes incluem instilação na traquéia (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), injeção muscular (Ragot *et al.*, 1993), injeções intravenosas periféricas (Herz & Gerard, 1993) e inoculação estereotáctica no cérebro (Le Gal La Salle *et al.*, 1993).

Vetores de adenovírus podem originar-se de adenovírus humano. Alternativamente, eles podem originar-se de adenovírus de outras

espécies, p. ex. chimpanzés, que podem apresentar a vantagem de que os vetores virais não são neutralizados por anticorpos contra adenovírus humano circulante em muitos sujeitos humanos (ver, p. ex.: Tatsis N *et al* (2005) *Gene Ther.* Dec 1; [publicado eletronicamente enquanto ainda no prelo]).

## 5 2. Retrovírus

Os retrovírus são um grupo de vírus de RNA de filamento simples caracterizados por uma capacidade de converter seu RNA em DNA de filamento simples em células infectadas por meio de um processo de transcrição invertida (Coffin, 1990). O DNA resultante integra-se então de forma estável em cromossomos celulares como um pró-vírus e direciona a síntese de proteínas virais. A integração resulta na retenção das seqüências de gene viral na célula recipiente e seus descendentes. O genoma retroviral contém três genes, gag, pol, e env que codificam para proteínas de capsídeo, enzima polimerase, e componentes de envelope, respectivamente. Uma seqüência encontrada a montante do gene gag contém um sinal para empacotamento do genoma em vírions. Duas seqüências de repetição terminal longas (LTR) estão presentes nas extremidades 5' e 3' do genoma viral. Estas contêm seqüências acentuadora e promotora fortes e também são necessárias para integração no genoma da célula hospedeira (Coffin, 1990).

20 Para construir um vetor retroviral, um ácido nucleico que codifica uma ou mais seqüências de oligonucleotídeos ou polinucleotídeos de interesse é inserido no genoma viral em lugar de determinadas seqüências virais para produzir um vírus que é defectivo para replicação. Para produzir vírions, uma linha de células de empacotamento contendo os genes gag, pol, e 25 env, mas sem a LTR e componentes de empacotamento (Mann *et al*, 1983). Quando um plasmídeo recombinante contendo um cDNA, juntamente com a LTR retroviral e seqüências de empacotamento é introduzido nesta linha de células (por exemplo, por meio de precipitação com fosfato de cálcio), a seqüência de empacotamento permite que a transcrição de RNA do plasmídeo

recombinante seja empacotada em partículas virais, que então são secretadas no meio de cultura (Nicolas & Rubenstein, 1988; Temin, 1986; Mann *et al*, 1983). O meio contendo os retrovírus recombinantes é então coletado, opcionalmente concentrado, e usado para transferência de genes. Vetores retrovirais são capazes de infectar uma ampla variedade de tipos de células. No entanto, a integração e expressão estável exigem a divisão de células hospedeiras (Paskind *et al*, 1975).

Desenvolveu-se recentemente uma abordagem inédita projetada para permitir objetivação específica de vetores retrovirais com base na modificação química de um retrovírus por meio da adição química de radicais lactose no envelope viral. Esta modificação poderia permitir a infecção específica de hepatócitos via receptores de sialoglicoproteína.

Projetou-se uma abordagem diferente para a objetivação de retrovírus recombinantes em que se usa anticorpos biotinilados contra uma proteína de envelope retroviral e contra um receptor de células específico. Os anticorpos foram acoplados via os componentes de biotina por meio do uso de estreptavidina (Roux *et al*, 1989). Usando-se anticorpos contra antígenos de complexo de histocompatibilidade principal de classe I e de classe II, eles demonstraram a infecção de uma variedade de células humanas que portaram aqueles antígenos de superfície com um vírus ecotrópico *in vitro* (Roux *et al*, 1989).

### 3. Vírus adeno-associados

AAV (Ridgeway, 1988; Hermonat & Muzycska, 1984) é um parovírus, descoberto como uma contaminação de estoques adenovirais. É um vírus ubíquo (anticorpos estão presentes em 85% da população humana dos E.U.A.) que ainda não foi ligado a qualquer doença. Ele também é classificado como um dependovírus, porque suas replicações são dependentes da presença de um vírus auxiliar, como adenovírus. Isolou-se cinco sorotipos, dos quais AAV-2 é o melhor caracterizado.

AAV apresenta um DNA linear de filamento simples que é encapsidado em proteínas de capsídeo VP1, VP2 e VP3 para formar um vírion icosaédrico com de 20 a 24 nm de diâmetro (Muzyczka & McLaughlin, 1988).

5 O DNA de AAV tem um comprimento de aproximadamente 4,7 kilobases. Ele contém duas matrizes de leitura aberta e é flanqueado por duas repetições invertidas (ITRs, *inverted repeats*). Há dois genes principais no genoma do AAV: rep e cap. O gene rep codifica para proteínas responsáveis por replicações virais, sendo que cap codifica para proteína de capsídeo VP 1-3. Cada ITR forma uma estrutura de grampo-de-cabelo em forma de T. Estas repetições terminais são os únicos componentes *cis* essenciais do AAV para integração cromossômica. Portanto, o AAV pode ser usado como um vetor com todas as seqüências codificantes virais removidas e substituídas pela cassete de genes para fornecimento. Três promotores virais foram identificados e denominados p5, p19, e p40, de acordo com sua posição map. A transcrição de p5 e p19 resulta na produção de proteínas rep, e transcrição de p40 produz as proteínas de capsídeo (Hermonat & Muzyczka, 10 15 1984).

Há vários fatores que instaram pesquisadores a estudarem a possibilidade de usar rAAV como um vetor de expressão. Uma é que as exigências para fornecimento de um gene para integrar no cromossomo hospedeiro são surpreendentemente poucas. É necessário ter as ITRs com 145 bp, que são apenas 6% do genoma de AAV. Isto deixa espaço no vetor para montar uma inserção de DNA de 4,5 kb. Enquanto esta capacidade de suportar possa impedir o AAV de fornecer genes grandes, ele é amplamente adequado para fornecer as construções anti-sentido da presente invenção. 20 25

AAV também é uma boa escolha de veículos de fornecimento em virtude de sua segurança. Há um mecanismo de recuperação relativamente complicado: não só adenovírus de tipo selvagem, mas também genes de AAV

são necessários para mobilizar rAAV. Da mesma forma, AAV não é patogênico e não está associado com qualquer doença. A remoção de seqüências codificantes virais minimiza reações imunes à expressão de genes virais, e, portanto, rAAV não evoca uma resposta inflamatória.

#### 5 4. Outros vetores virais como construções de expressão

É possível empregar outros vetores virais como construções de expressão na presente invenção para o fornecimento de seqüências de oligonucleotídeos e de polinucleotídeos a uma célula hospedeira. É possível usar vetores derivados de vírus, como vírus *vaccinia* (Ridgeway, 1988; 10 Coupar *et al*, 1988), lentivírus, pólio vírus e herpes vírus. Eles oferecem várias características atraentes para várias células mamíferas (Friedmann, 1989; Ridgeway, 1988; Coupar *et al*, 1988; Horwich *et al*, 1990).

Com o recente reconhecimento de vírus de hepatite B defectivos, obteve-se nova compreensão sobre a relação estrutura-função de 15 diferentes seqüências virais. Estudos *in vitro* mostraram que o vírus poderia conservar a capacidade de empacotamento dependente de auxiliar e de transcrição reversa apesar da deleção de até 80% de seu genoma (Horwich *et al*, 1990). Isto sugeriu que grandes porções do genoma poderiam ser substituídas por material genético estranho. O hepatotropismo e persistência 20 (integração) constituíram propriedades particularmente atraentes para transferência de gene direcionada para o fígado. Chang *et al* (1991) introduziram o gene da cloranfenicol acetil transferase (CAT) no genoma do vírus da hepatite B do pato no local das seqüências codificantes de polimerase, superfície e pré-superfície. Este foi co-transfectado com vírus de 25 tipo selvagem em uma linha de células de hepatoma aviário. Usou-se meio de cultura contendo altas titulações do vírus recombinante para infectar hepatócitos primários de filhotes de pato. A expressão estável do gene de CAT foi detectada durante pelo menos 24 dias após a transfecção (Chang *et al*, 1991).

## 5. Vetores não-virais

Para efetuar a expressão das seqüências de oligonucleotídeos ou polinucleotídeos da presente invenção, a construção de expressão precisa ser fornecida em uma célula. Este fornecimento pode ser realizado *in vitro*,  
5 como em procedimentos de laboratório para transformar linha de células, ou *in vivo* ou *ex vivo*, como no tratamento de determinados estados de doença. Como descrito acima, um mecanismo preferido para fornecimento é via infecção viral, sendo que a construção de expressão é encapsulada em uma partícula viral infecciosa.

10 Uma vez que a construção de expressão foi fornecida na célula, o ácido nucleico codificando as desejadas seqüências de oligonucleotídeos ou polinucleotídeos podem ser posicionadas e expressas em sítios diferentes. Em determinadas concretizações, o ácido nucleico codificando a construção pode ser integrada de forma estável no genoma da  
15 célula. Esta integração pode ser no local e orientação específicas via recombinação homóloga (substituição de gene) ou pode ser integrada em um local randômico, não específico (aumento do gene). Em outras concretizações adicionais, o ácido nucleico pode ser mantido de forma estável na célula como um segmento de DNA separado, epissômico. Referidos segmentos de  
20 ácido nucleico ou "epissomas" codificam seqüências suficientes para permitir a manutenção e replicação independente ou sincronizada com o ciclo da célula hospedeira. Como a construção de expressão é fornecida em uma célula, e onde, na célula, o ácido nucleico permanece depende do tipo de construção de expressão empregada.

25 Em determinadas concretizações da invenção, a construção de expressão compreendendo uma ou mais seqüências de oligonucleotídeos ou polinucleotídeos pode consistir simplesmente de plasmídeos ou DNA recombinante nu. A transferência da construção pode ser realizada por meio de qualquer um dos métodos mencionados acima, que permeabilizam



fisicamente ou quimicamente a membrana da célula. Isto é particularmente aplicável para transformação *in vitro*, mas também pode ser aplicado no uso *in vivo*. Dubensky *et al.* (1984) injetaram com êxito DNA de poliomavírus na forma de precipitados de fosfato de cálcio no fígado e no baço de camundongos adultos e recém-nascido demonstrando replicação viral ativa e infecção aguda. Benvenisty & Reshef (1986) também demonstraram que a injeção intraperitoneal direta de plasmídeos precipitados com fosfato de cálcio resulta na expressão dos genes transfectados. Conjeturamos que o DNA que codifica um gene de interesse também pode ser transferido de maneira similar e expressar o produto gênico.

Outra concretização da invenção para transferir uma construção de expressão de DNA nu para células pode envolver bombardeio de partículas. Este método depende da capacidade de acelerar micro-projéteis revestidos com DNA a uma alta velocidade, permitindo-lhes perfurar membranas celulares e entrar nas células sem matá-las (Klein *et al.*, 1987). Vários dispositivos para acelerar partículas pequenas foram desenvolvidos. Um dispositivo do tipo referido baseia-se numa descarga de alta voltagem para gerar uma corrente elétrica, que, por sua vez, proporciona a força impulsora (Yang *et al.*, 1990). Os micro-projéteis usados consistiram de substâncias biologicamente inertes, como pérolas de tungstênio ou ouro.

Órgãos selecionados incluindo o fígado, pele, e tecido muscular de ratos e camundongos foram bombardeados *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). Isto pode exigir exposição cirúrgica do tecido ou das células, para eliminar qualquer tecido interveniente entre o canhão e o órgão-alvo, i.e., tratamento *ex vivo*. Novamente, o DNA codificando um gene particular pode ser fornecido por meio deste método e ainda ser incorporado pela presente invenção.

#### Composições de polipeptídeos

Em outros aspectos, a presente invenção proporciona

composições de polipeptídeo. De uma maneira geral, um polipeptídeo da invenção será um polipeptídeo isolado (ou um epítipo, variante, ou fragmento ativo do mesmo) derivado de uma espécie mamífera. De preferência, o polipeptídeo é codificado por uma seqüência de polinucleotídeos revelada aqui ou uma seqüência que hibrida sob condições moderadamente estridentes a uma seqüência de polinucleotídeos revelada aqui. Alternativamente, o polipeptídeo pode ser definido como um polipeptídeo que compreende uma seqüência de aminoácidos contígua de uma seqüência de aminoácidos revelada aqui, ou cujo polipeptídeo compreende uma seqüência de aminoácidos inteira aqui revelada.

Porções imunogênicas podem ser identificadas geralmente usando-se técnicas bem conhecidas, como aquelas resumidas em Paul, *Fundamental Immunology*, 3<sup>a</sup> ed., 243-247 (1993) e referências ali indicadas. Referidas técnicas incluem seleção de polipeptídeos quanto à capacidade de reagir com anticorpos específicos para antígeno, anti-soros e/ou linha de células T ou clones. Como usado aqui, anti-soros e anticorpos são "específicos para antígeno" se eles se ligarem especificamente a um antígeno (i.e., eles reagem com a proteína em um ELISA ou outro ensaio imunológico, e não reagem detectavelmente com proteínas não-relacionadas). Referidos anti-soros e anticorpos podem ser preparados como descrito aqui, e usando-se técnicas bem conhecidas. Uma porção imunogênica de uma proteína de *Mycobacterium* sp. é uma porção que reage com referidos anti-soros e/ou células T em um nível que não é substancialmente inferior à reatividade do polinucleotídeo de comprimento pleno (p. ex., em um ensaio ELISA e/ou de reatividade de células T). Referidas porções imunogênicas podem reagir em referidos ensaios em um nível que é similar ou maior do que a reatividade do polipeptídeo de comprimento pleno. Referidas seleções podem ser realizadas geralmente usando-se métodos bem conhecidos por aqueles com prática ordinária na arte, como aqueles descritos em Harlow & Lane, *Antibodies: A*

*Laboratory Manual* (1988) e *Using Antibodies: A Laboratory Manual* (1998).

Por exemplo, um polipeptídeo pode ser imobilizado sobre um suporte sólido e contatado com soros de pacientes de forma a permitir a ligação de anticorpos nos soros do polipeptídeo imobilizado. Soros não-ligados podem então ser  
5 removidos e é possível detectar anticorpos ligados usando-se, por exemplo, proteína A marcada com  $^{125}\text{I}$ .

Polipeptídeos podem ser preparados usando-se qualquer uma dentre várias técnicas bem conhecidas. Polipeptídeos recombinantes codificados por seqüências de DNA como descrito acima podem ser  
10 preparados facilmente a partir das seqüências de DNA usando-se qualquer um de uma variedade de vetores de expressão conhecidos por aqueles com prática ordinária na arte. Expressão pode ser obtida em qualquer célula hospedeira apropriada que foi transformada ou transfectada com um vetor de expressão contendo uma molécula de DNA que codifica um polipeptídeo recombinante.  
15 Células hospedeiras vantajosas incluem procariontes, levedura, e células procarióticas superiores, como células mamíferas e células de planta. De preferência, as células hospedeiras empregadas são de *E. coli*, levedura ou uma linha de células mamífera, como COS ou CHO. Sobrenadantes de sistemas de hospedeiro/vetor vantajosos que secretam polipeptídeo ou  
20 proteína recombinante em meio de cultura podem ser concentrados primeiro usando-se um filtro comercialmente obténível. Após a concentração, o concentrado pode ser aplicado em uma matriz de purificação vantajosa, como uma matriz de afinidade ou uma resina de troca de íon. Finalmente, é possível empregar uma ou mais etapas de HPLC de fase invertida para purificar  
25 adicionalmente um polipeptídeo recombinante.

Polipeptídeos da invenção, fragmentos imunogênicos dos mesmos, e outras variantes apresentando menos do que cerca de 100 aminoácidos, e geralmente menos de cerca de 50 aminoácidos, também podem ser gerados por meios sintéticos, usando-se técnicas bem conhecida

por aqueles com prática ordinária na arte. Por exemplo, referidos polipeptídeos podem ser sintetizados usando-se qualquer uma das técnicas de fase sólida comercialmente obteníveis, como o método de síntese de fase sólida de Merrifield, em que aminoácidos são adicionados seqüencialmente em uma cadeia de aminoácidos em desenvolvimento. Ver Merrifield, *J. Am. Chem. Soc.* 85:2149-2146 (1963). Equipamento para síntese automatizada de polipeptídeos é comercialmente obtível de fornecedores, como Perkin Elmer/ Applied BioSystems Division (Foster City, CA), e pode ser operado de acordo com as instruções do fabricante.

10                   Em determinadas concretizações específicas, um polipeptídeo pode ser uma proteína de fusão que compreende múltiplos polipeptídeos como descrito aqui, ou que compreende pelo menos um polipeptídeo como descrito aqui e uma seqüência não-relacionada, como uma proteína de tumor conhecida. Por exemplo, um parceiro de fusão pode contribuir para proporcionar epítopos auxiliares T (um parceiro de fusão imunológico), de preferência, epítopos de auxiliares T reconhecidos por humanos, ou pode contribuir para a expressão da proteína (um acentuador de expressão) com rendimentos maiores do que a proteína recombinante. Determinados parceiros de fusão preferidos são parceiros de fusão acentuadores de expressão e imunológicos. É possível selecionar outros parceiros de fusão de forma a incrementar a solubilidade da proteína ou permitir objetivar a proteína para compartimentos intracelulares desejados. Outros parceiros de fusão adicionais incluem marcadores de afinidade, que facilitam a purificação da proteína.

25                   Proteínas de fusão podem ser preparadas geralmente usando-se técnicas convencionais, incluindo conjugação química. De preferência, uma proteína de fusão é expressa como uma proteína recombinante, permitindo a produção de níveis incrementados, relativamente a uma proteína não-fundida, em um sistema de expressão. Em resumo, seqüências de DNA codificando os componentes de polipeptídeo podem ser montadas separadamente, e ligadas

em um vetor de expressão apropriado. A extremidade 3' da seqüência de DNA que codifica um componente de polipeptídeo é ligada, com ou sem um ligante de peptídeo, na extremidade 5' de uma seqüência de DNA codificando o segundo componente de polipeptídeo de tal forma que as matrizes de leitura das seqüências se encontrem em fase. Isto permite a tradução numa única proteína de fusão que conserva a atividade biológica de ambos os componentes de polipeptídeos.

É possível empregar uma seqüência de ligante de peptídeo para separar o primeiro e segundo componentes de polipeptídeo a uma distância suficiente para assegurar que cada polipeptídeo se dobre em suas estruturas secundárias e terciárias. Uma seqüência de ligante de peptídeo do tipo referido é incorporada na proteína de fusão usando-se técnicas padrão bem conhecidas na arte. Seqüências de ligante de peptídeo vantajosas podem ser selecionadas com base nos seguintes fatores: (1) sua capacidade de adotar uma conformação estendida flexível; (2) sua incapacidade de adotar uma estrutura secundária que poderia interagir com epitopos funcionais no primeiro e segundo polipeptídeos; e (3) a falta de radicais hidrofóbicos ou carregados que poderiam reagir com os epitopos funcionais de polipeptídeo. Seqüências de ligante de peptídeo preferidas contêm radicais Gly, Asn e Ser. Outros aminoácidos quase neutros, como Thr e Ala, também podem ser usados na seqüência de ligante. Seqüências de aminoácidos que podem ser empregadas de forma útil como ligantes incluem aquelas divulgadas em Maratea *et al*, *Gene* 40:39-46 (1985); Murphy *et al*, *Proc. Natl. Acad. Sci. USA* 83:8258-8262 (1986); Patente dos Estados Unidos nº 4.935.233 e Patente dos Estados Unidos nº 4.751.180. De uma maneira geral a seqüência de ligante pode ter comprimento de 1 a cerca de 50 aminoácidos. Seqüências de ligante não são exigidas quando o primeiro e segundo polipeptídeos apresentam regiões de aminoácidos N-terminais não-essenciais que podem ser usadas para separar os domínios funcionais e prevenir interferência estérica.

As seqüências de DNA ligadas são ligadas operacionalmente a elementos reguladores de transcrição ou tradução vantajosos. Os elementos reguladores responsáveis pela expressão de DNA encontram-se localizados apenas a 5' da seqüência de DNA que codifica os primeiros polipeptídeos. De maneira análoga, são necessários códon de interrupção para parar os sinais de terminação de tradução e transcrição estão presentes apenas a 3' da seqüência de DNA que codifica o segundo polipeptídeo.

Proporciona-se também proteínas de fusão. Referidas proteínas compreendem um polipeptídeo como descrito aqui em conjunto com uma proteína imunogênica não-relacionada. De preferência, a proteína imunogênica é capaz de elicitar uma resposta de recuperação. Exemplos de referidas proteínas incluem proteínas do tétano, tuberculose e hepatite (ver, p. ex., Stoute *et al*, *New Engl J. Med.* 336:86-91 (1997)).

Em concretizações preferidas, um parceiro de fusão imunológico deriva-se da proteína D, uma proteína de superfície da bactéria gram-negativa *Haemophilus influenza B* (WO 91/18926). De preferência, um derivado de proteína D compreende aproximadamente o primeiro terço da proteína (p. ex., os primeiros de 100 a 110 aminoácidos N-terminais), e um derivado de proteína D pode ser lipidado. Em determinadas concretizações preferidas, os primeiros 109 radicais de um parceiro de fusão de lipoproteína D são incluídos na extremidade N para dotar o polipeptídeo com epitopos de células T exógenos adicionais e incrementar o nível de expressão em *Escherichia coli* (desta forma funcionando como um acentuador de expressão). A cauda de lipídeo assegura ótima apresentação do antígeno para células apresentadoras de antígeno. Outros parceiros de fusão incluem a proteína não-estrutural do vírus da influenza, NS1 (hemaglutinina). Tipicamente usa-se os 81 aminoácidos N-terminais, embora seja possível usar diferentes fragmentos que incluem epitopos de auxiliares T.

Em outra concretização, o parceiro de fusão imunológico é a

proteína conhecida como LYTA, ou uma porção da mesma (de preferência uma porção C-terminal). LYTA é derivada de *Streptococcus pneumoniae*, que sintetiza uma N-acetil-L-alanina amidase conhecida como amidase LYTA (codificada pelo gene *LytA*; *Gene* 43:265-292 (1986)). LYTA é uma autolisina que degrada especificamente determinadas ligações na espinha dorsal de peptidoglicano. O domínio C-terminal da proteína LYTA é responsável pela afinidade com a colina ou alguns análogos de colina, como DEAE. Esta propriedade foi explorada para o desenvolvimento de C-LYTA de *Escherichia coli* expressando plasmídeos úteis para expressão de proteínas de fusão. A purificação de proteínas híbridas contendo fragmento de C-LYTA na extremidade amino já foi descrita (ver *Biotechnology* 10:795-798 (1992)). Em uma concretização preferida, é possível incorporar uma porção de repetição de LYTA em uma proteína de fusão. Uma porção de repetição é encontrada na região C-terminal que começa no radical 178. Uma porção de repetição particularmente preferida incorpora radicais de 188 a 305.

De uma maneira geral, polipeptídeos (incluindo proteínas de fusão) e polinucleotídeos como descrito aqui são isolados. Um polipeptídeo ou polinucleotídeo "isolado" é um que é removido de seu ambiente original. Por exemplo, uma proteína naturalmente ocorrente é isolada se ela for separada de parte dos materiais coexistentes, ou todos os materiais coexistentes, no sistema natural. De preferência, referidos polipeptídeos são pelo menos cerca de 90% puros, mais preferivelmente, pelo menos cerca de 95% puros e, da forma mais preferível, pelo menos cerca de 99% puros. Um polinucleotídeo é considerado como sendo isolado se, por exemplo, ele for clonado em um vetor que não é parte do ambiente natural.

### Células T

Composições imunoterápicas também podem compreender, alternativamente, células T específicas para um antígeno de *Mycobacterium*. Referidas células podem ser preparadas geralmente *in vitro* ou *ex vivo*,

usando-se procedimentos convencionais. Por exemplo, células T podem ser isoladas da medula óssea, sangue periférico, ou uma fração de medula óssea ou sangue periférico de um paciente, usando-se um sistema de separação de células comercialmente obténível, como o sistema Isolex™, obténível da  
5 Nexell Therapeutics, Inc. (Irvine, CA; ver também Patente dos Estados Unidos nº 5.240.856; Patente dos Estados Unidos nº 5.215.926; WO 89/06280; WO 91/16116 e WO 92/07243). Alternativamente, células T podem ser derivadas de humanos relacionados ou não-relacionados, mamíferos não-humanos, linhas de células ou culturas.

10 Células T podem ser estimuladas com um polipeptídeo da invenção, polinucleotídeo codificando um referido polipeptídeo, e/ou uma célula que apresenta antígeno (APC) que expressa um polipeptídeo do tipo referido. Referida estimulação é realizada em condições e durante um tempo suficiente para permitir a geração de células T que são específicas para o  
15 polipeptídeo. De preferência, o polipeptídeo ou polinucleotídeo está presente em um veículo para fornecimento, como uma microesfera, para facilitar a geração de células T específicas.

Células T são consideradas específicas para um polipeptídeo da invenção se as células T especificamente proliferarem, secretarem  
20 citocinas ou eliminarem células-alvo revestidas com o polipeptídeo ou expressando um gene que codifica o polipeptídeo. A especificidade de células T pode ser avaliada usando-se qualquer uma dentre várias técnicas convencionais. Por exemplo, em um ensaio de liberação de cromo ou ensaio de proliferação, um índice de estimulação maior que duas vezes o aumento da  
25 lise e/ou proliferação, em comparação com controles negativos, indica especificidade de células T. Referidos ensaios podem ser realizados, por exemplo, como descrito em Chen *et al.*, *Cancer Res.* 54:1065-1070 (1994)). Alternativamente, a detecção da proliferação de células T pode ser realizada por meio de uma variedade de técnicas conhecidas. Por exemplo, a



proliferação de células T pode ser detectada medindo-se uma taxa incrementada de síntese de DNA (p. ex., por meio de marcação-com-pulso de culturas de células T com timidina tritiada e medindo-se a quantidade de timidina tritiada incorporada no DNA). Contato com um polipeptídeo da invenção (de 100 ng/ml a 100 µg/ml, de preferência, de 200 ng/ml a 25 µg/ml) durante de 3 a 7 dias deveria resultar em um incremento de pelo menos duas vezes na proliferação das células T. Contato como descrito acima durante de 2 a 3 horas deveria resultar na ativação das células T, conforme medido usando-se ensaios de citocina convencionais em que um incremento de duas vezes no nível de liberação de citocina (p. ex., TNF ou IFN-γ) é indicativo da ativação de células T (ver Coligan *et al*, *Current Protocols in Immunology*, vol. 1 (1998)). Células que foram ativadas em resposta a uma APC expressando polipeptídeo, polinucleotídeo ou polipeptídeo pode ser CD4<sup>+</sup> e/ou CD8<sup>+</sup>. células T específicas para proteína podem ser expandidas usando-se técnicas convencionais. Em concretizações preferidas, as células T são derivadas de um paciente, um doador relacionado ou de um doador não relacionado, e são administradas ao paciente após estimulação e expansão.

Para fins terapêuticos, células T CD4<sup>+</sup> ou CD8<sup>+</sup> que proliferam em resposta a um polipeptídeo, polinucleotídeo ou APC podem ser expandidas em número, tanto *in vitro* ou *in vivo*. A proliferação de referidas células T *in vitro* pode ser realizada de diversas maneiras. Por exemplo, as células T podem ser novamente expostas a um polipeptídeo, ou um peptídeo curto correspondendo a uma porção imunogênica de um referido polipeptídeo, com ou sem a adição de fatores de crescimento de células T, como interleucina-2, e/ou células estimuladoras que sintetizam um polipeptídeo. Alternativamente, uma ou mais células T que proliferam na presença de uma proteína podem ser expandidas em número por meio de clonagem. Métodos para clonagem de células são bem conhecidos na arte, e incluem diluição limitante.

### Composições farmacêuticas

Em concretizações adicionais, a presente invenção refere-se à formulação de um ou mais do polinucleotídeo, polipeptídeo, célula T, anticorpo, e composições quimioterápicas aqui reveladas, em soluções  
5 farmacêuticamente aceitáveis para administração em uma célula ou animal, quer sozinhas, ou em combinação com um ou mais outras modalidades de terapia.

Também se compreenderá que, se desejado, o segmento de ácido nucleico (p. ex., RNA ou DNA) que expressa um polipeptídeo como  
10 divulgado aqui, também pode ser administrado em combinação com outros agentes, como, p. ex., outras proteínas ou polipeptídeos ou vários agentes farmacêuticamente ativos, incluindo agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis*. Eficazmente, virtualmente não há limite para outros componentes que também podem ser incluídos, dado que os  
15 agentes adicionais não causam um efeito adverso significativo quando do contato com as células-alvo ou tecidos hospedeiros. Assim, as composições podem ser fornecidas juntamente com vários outros agentes conforme necessário no caso particular. Referidas composições podem ser purificadas de células hospedeiras ou outras fontes biológicas, ou, alternativamente,  
20 podem ser sintetizadas quimicamente como descrito aqui. Da mesma forma, referidas composições podem compreender composições de RNA ou DNA substituídos ou derivatizados.

Formulação de excipientes farmacêuticamente aceitáveis e soluções veículo farmacêuticamente aceitáveis é bem conhecida por bem  
25 conhecida por aqueles com prática na arte, como o é o desenvolvimento de regimes vantajosos de dosagem e tratamento para uso das composições particulares aqui descritas em uma variedade de regimes de tratamento, incluindo, p. ex., administração e formulação oral, parenteral, intravenosa, intranasal, e intramuscular. Tipicamente, formulações compreendendo uma

quantidade terapeuticamente eficaz fornece cerca de 2 µg a cerca de 50 µg de polipeptídeo de Mtb72f por administração, mais tipicamente de cerca de 5 µg a cerca de 40 µg de polipeptídeo de Mtb72f por administração.

### 1. Fornecimento oral

5                    Em determinadas aplicações, as composições farmacêuticas reveladas aqui podem ser fornecidas via administração oral a um animal. Desta forma, estas composições podem ser formuladas com um diluente inerte ou com um veículo comestível assimilável, ou elas podem ser embutidas em cápsula de gelatina de capa dura ou de capa mole, ou elas  
10 podem ser comprimidas em tabletes, ou elas podem ser incorporada diretamente com o alimento da dieta.

Os compostos ativos podem ser incorporados com excipientes e usados em forma de tabletes ingeríveis, tabletes bucais, pastilhas, cápsulas, elixires, suspensões, xaropes, wafers, e análogos (Matiowitz *et al*, 1997;  
15 Hwang *et al*, 1998; Patente dos Estados Unidos nº 5.641.515; Patente dos Estados Unidos nº 5.580.579 e Patente dos Estados Unidos nº 5.792.451, cada uma incorporada aqui especificamente e integralmente por referência). Os tabletes, pastilhas, pílulas, cápsulas e análogos também podem conter o seguinte: um aglutinante, como goma tragacanto, acácia, amido de milho, ou  
20 gelatina; excipientes, como fosfato de dicálcio; um agente desintegrante, como amido de milho, amido de batata, ácido algínico e análogos; um lubrificante, como estearato de magnésio; e é possível adicionar um agente adoçante, como sacarose, lactose ou sacarina, ou um agente aromatizante, como menta, óleo de gaultéria, ou aroma de cereja. Quando a forma de  
25 dosagem é uma cápsula, ela pode conter, adicionalmente a materiais do tipo acima, um veículo líquido. Diversos outros materiais podem estar presentes como revestimentos ou, de outra forma, para modificar a forma física da unidade de dosagem. Por exemplo, tabletes, pílulas, ou cápsulas podem ser revestidos com *shellac*, açúcar, ou ambos. Um xarope de elixir pode conter o

composto ativo sacarose como um agente adoçante, metil e propil parabenos como conservantes, um corante e aromatizante, como aroma de cereja ou laranja. Evidentemente, qualquer material usado na preparação de qualquer forma de unidade de dosagem deveria ser farmacêuticamente puro e substancialmente não-tóxico nas quantidades empregadas. Adicionalmente, os compostos ativos podem ser incorporados em formulações e preparações de liberação sustentada.

Tipicamente, estas formulações contêm usualmente entre de 2  $\mu\text{g}$  a 50  $\mu\text{g}$  de polipeptídeo de Mtb72f. Naturalmente, a quantidade de composto(s) ativo(s) em cada composição terapeuticamente útil pode ser preparada de tal modo que seja possível obter uma dosagem vantajosa em qualquer dose unitária do composto. Fatores, como solubilidade, biodisponibilidade, meia-vida biológica, via de administração, vida-de-prateleira do produto, e também outras considerações farmacológicas serão consideradas por alguém com prática na arte da preparação de referidas formulações farmacêuticas, e como tais, pode ser desejável uma variedade de dosagens e regimes de tratamento.

Para administração oral, as composições da presente invenção podem ser incorporadas alternativamente com um ou mais excipientes em forma de um líquido para limpeza bucal, dentifrício, tablete bucal, spray oral, ou formulação sublingual administrada oralmente. Por exemplo, um líquido para limpeza bucal pode ser preparado incorporando-se o ingrediente ativo na quantidade requerida em um solvente apropriado, como uma solução de borato de sódio (solução de Dobell's). Alternativamente, o ingrediente ativo pode ser incorporado em uma solução oral, como uma contendo borato de sódio, glicerina e bicarbonato de potássio, ou dispersado em um dentifrício, ou adicionado em uma quantidade terapeuticamente eficaz a uma composição que pode incluir água, ligantes, abrasivos, agentes aromatizantes, agentes de espumação, e umectantes. Alternativamente, as composições pode ser

configuradas em forma de tablete ou solução que pode ser colocada sob a língua ou, de outra forma, dissolvida na boca.

## 2. Fornecimento injetável

Em determinados casos, será desejável fornecer as composições farmacêuticas aqui reveladas por via parenteral, intravenosa, intramuscular, ou mesmo intravenosa, como descrito na Patente dos Estados Unidos nº 5.543.158; Patente dos Estados Unidos nº 5.641.515 e Patente dos Estados Unidos nº 5.399.363 (cada uma incorporada especificamente e integralmente aqui por referência). Soluções dos compostos ativos como base livre ou sais farmacologicamente aceitáveis podem ser preparadas em água misturada vantajosamente com um tensoativo, como hidroxipropilcelulose. Também é possível preparar dispersões em glicerol, polietileno glicóis líquidos, e misturas dos mesmos e em óleos. Em condições ordinárias de armazenamento e uso, estas preparações contêm um conservante para prevenir o crescimento de microorganismos.

As formas farmacêuticas vantajosas para uso injetável incluem dispersões ou soluções aquosas estéreis e pós estéreis para a preparação extemporânea de dispersões ou soluções injetáveis estéreis (Patente dos Estados Unidos nº 5.466.468, incorporada especificamente e integralmente aqui por referência). Em todos os casos, a forma precisa ser estéril e precisa ser fluida até o ponto de oferecer fácil seringabilidade. Ela precisa ser estável nas condições de condições de fabricação e armazenamento e precisa ser conservada contra a ação contaminadora de microorganismos, como bactérias e fungos. O veículo pode ser um solvente ou meio de dispersão contendo, por exemplo, água, etanol, polioliol (p. ex., glicerol, propileno glicol, e polietileno glicol líquido, e análogos), misturas vantajosas dos mesmos, e/ou óleos vegetais. A fluidez apropriada pode ser mantida, por exemplo, com o uso de um revestimento, como lecitina, por meio da manutenção do tamanho de partículas desejado no caso de dispersão, e por meio do uso de tensoativos. A

prevenção da ação de microorganismos pode ser facilitada por meio de vários agentes antibacterianos e antifúngicos, por exemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal, e análogos. Em muitos casos, será preferível incluir agentes isotônicos, por exemplo, açúcares ou cloreto de sódio. A absorção prolongada das composições injetáveis pode ser proporcionada com o uso, nas composições, de agentes retardadores de absorção, por exemplo, monoesterato de alumínio e gelatina.

Para administração parenteral em uma solução aquosa, por exemplo, de preferência, a solução deveria ser tamponada, se necessário, e o diluente líquido primeiramente tornado isotônico com suficiente solução salina ou glicose. Estas soluções aquosas particulares são particularmente vantajosas para administração intravenosa, intramuscular, subcutânea e intraperitoneal. Com relação a isto, um meio aquoso estéril que pode ser empregado será de conhecimento daqueles com prática na arte, à luz da presente descrição. Por exemplo, uma dosagem pode ser dissolvida em 1 ml de solução isotônica de NaCl e, ou adicionada a 1000 ml de fluido de hipodermoclise ou injetada no sítio de infusão proposto (ver, p. ex., Remington 's Pharmaceutical Sciences, 15<sup>a</sup> edição, pp. 1035-1038 e 1570-1580). Será necessário ocorrer alguma variação na dosagem, dependendo da condição do sujeito que está sendo tratado. A pessoa responsável pela administração determinará, em qualquer caso, a dose apropriada para o sujeito individual. Além disso, para administração humana, preparações deveriam atender os padrões de esterilidade, pirogenicidade, e os padrões de pureza e segurança gerais conforme exigidos pelos padrões do departamento de biológicas do FDA [administração de alimentos e drogas do governo norte-americano].

Soluções injetáveis estéreis são preparadas incorporando-se os compostos ativos na quantidade requerida no solvente apropriado com numerosos dos outros ingredientes enumerados acima, conforme necessário,

seguido de esterilização filtrada. De uma maneira geral, prepara-se dispersões incorporando os diversos ingredientes ativos esterilizados em um veículo estéril que contém o meio de dispersão básico e os outros ingredientes exigidos dentre aqueles enumerados acima. No caso de pós estéreis para a  
5 preparação de soluções injetáveis estéreis, os métodos preferidos de preparação são técnicas de secagem a vácuo e de secagem com congelamento que proporcionam um pó do ingrediente ativo acrescido de qualquer ingrediente adicional desejado a partir de uma solução do mesmo previamente filtrada de forma estéril.

10 As composições aqui reveladas podem ser formuladas em uma forma neutra ou de sal. Sais farmacologicamente aceitáveis incluem os sais de adição de ácido (formados com os grupos amino livres da proteína) e que são formados com ácidos inorgânicos, como, por exemplo, ácido clorídrico ou fosfórico, ou ácidos orgânicos, como acético, oxálico, tartárico, mandélico, e  
15 análogos. Sais formados com os grupos carboxila livres também podem ser derivados de bases inorgânicas, como, por exemplo, hidróxidos de sódio, potássio, amônio, cálcio, ou férrico, e bases orgânicas como isopropilamina, trimetilamina, histidina, procaína e análogos. Quando da formulação, soluções serão administradas de uma maneira compatível com o formulação de  
20 dosagem e em tal quantidade que seja terapêuticamente eficaz. As formulações são facilmente administradas em uma variedade de formas de dosagem, como soluções injetáveis, cápsulas de liberação de droga, e análogos.

25 Como usado aqui, "veículo" inclui qualquer um e todos os solventes, meios de dispersão, veículos, revestimentos, diluentes, agentes antibacterianos e antifúngicos, agentes isotônicos e retardadores de absorção, tampões, soluções veículo, suspensões, colóides, e análogos. O uso de tais meios e agentes para substâncias ativas farmacêuticas é bem conhecido na arte. Exceto se, qualquer meio ou agente convencional for incompatível com

o ingrediente ativo, seu uso nas composições terapêuticas será considerado. Também é possível incorporar ingredientes ativos suplementares nas composições.

A expressão "farmaceuticamente aceitável" refere-se a entidades moleculares e composições que não produzem uma reação alérgica ou similar indesejável quando administradas a um humano. A preparação de uma composição aquosa que contém uma proteína como um ingrediente ativo é bem compreendida na arte. Tipicamente, referidas composições são preparadas como injetáveis, seja como soluções ou dispersões líquidas; também é possível preparar formas sólidas adequadas para dissolução ou suspensão em líquido antes da injeção. A preparação também pode ser emulsificada.

### 3. Fornecimento nasal e bucal

Em determinadas concretizações, as composições farmacêuticas podem ser fornecidas por meio de sprays intranasais, sprays bucais, inalação e/ou outros veículos de fornecimento de aerossol. Métodos para fornecimento de genes, ácidos nucleicos, e composições de peptídeos diretamente nos pulmões, p. ex. via sprays de aerossol nasal e bucal, foram descritos, p. ex., na Patente dos Estados Unidos nº 5.756.353 e Patente dos Estados Unidos nº 5.804.212 (cada uma incorporada especificamente e integralmente aqui por referência). De maneira análoga, o fornecimento de drogas usando resinas de micropartículas intranasais (Takenaga *et al.*, 1998) e compostos de lisofosfatidil-glicerol (Patente dos Estados Unidos nº 5.725.871, incorporada especificamente e integralmente aqui por referência) também é bem conhecido nas artes farmacêuticas. Da mesma forma, o fornecimento de droga transmucosal em forma de uma matriz de suporte de politetrafluoroetileno encontra-se descrito na Patente dos Estados Unidos nº 5.780.045 (incorporada especificamente e integralmente aqui por referência).

### 4. Fornecimento mediado com lipossomas, nanocápsulas e micropartículas



Em determinadas concretizações, os inventores consideram o uso de lipossomas, nanocápsulas, micropartículas, microesferas, partículas lipídicas, vesículas, e análogos, para a introdução das composições da presente invenção em células hospedeiras vantajosas. Em particular, as  
5 composições da presente invenção podem ser formuladas para fornecimento, seja encapsuladas em uma partícula de lipídeo, um lipossoma, uma vesícula, uma nanoesfera, ou uma nanopartícula ou análogo.

Referidas formulações podem ser preferidas para a introdução de formulações farmacêuticamente aceitáveis dos ácidos nucleicos ou  
10 construções aqui reveladas. A formação e uso de lipossomas é de conhecimento geral por parte daqueles com prática na arte (ver por exemplo, Couvreur *et al*, 1977; Couvreur, 1988; Lasic, 1998; que descreve o uso de lipossomas e nanocápsulas na terapia antibiótica objetivada para doenças e infecções bacterianas intracelulares). Recentemente desenvolveu-se  
15 lipossomas com qualidades aperfeiçoadas de estabilidade no soro e meios-tempos de circulação (Gabizon & Papahadjopoulos, 1988; Allen e Choun, 1987; Patente dos Estados Unidos nº 5.741.516, incorporada especificamente e integralmente aqui por referência). Adicionalmente, foram revistos vários métodos de lipossomas e preparações semelhantes a lipossoma como veículos  
20 potenciais para droga (Takakura, 1998; Chandran *et al*, 1997; Margalit, 1995; Patente dos Estados Unidos nº 5.567.434; Patente dos Estados Unidos nº 5.552.157; Patente dos Estados Unidos nº 5.565.213; Patente dos Estados Unidos nº 5.738.868 e Patente dos Estados Unidos nº 5.795.587, cada uma incorporada especificamente e integralmente aqui por referência).

25 Lipossomas foram usados com êxito numa variedade de tipos de células que normalmente são resistentes a transfecção por meio de outros procedimentos incluindo suspensões de células T, culturas de hepatócitos primários e células PC 12 (Renneisen *et al*, 1990; Muller *et al*, 1990). Adicionalmente, lipossomas são isentos de restrições de comprimento de

DNA, que são típicas de sistemas de fornecimento à base de vírus. Lipossomas foram usados eficazmente para introduzir genes, drogas (Heath & Martin, 1986; Heath *et al*, 1986; Balazsovits *et al*, 1989; Fresta & Puglisi, 1996), agentes radioterápicos (Pikul *et al*, 1987), enzimas (Imaizumi *et al*, 1990a; Imaizumi *et al*, 1990b), vírus (Faller & Baltimore, 1984), fatores de transcrição e efetores alostéricos (Nicolau & Gersonde, 1979) em uma variedade de animais e linhas de células cultivadas. Adicionalmente, completou-se vários caminhos clínicos bem sucedidos que examinam a efetividade de fornecimento de droga mediada com lipossoma (Lopez-Berestein *et al*, 1985a; 1985b; Coune, 1988; Sculier *et al*, 1988). Adicionalmente, vários estudos sugerem que o uso de lipossomas não é associado com respostas autoimunes, toxicidade ou localização gonadal após fornecimento sistêmico (Mori & Fukatsu, 1992).

Lipossomas são formados de fosfolipídeos que são dispersados em um meio aquoso e formam esextremidadeneamente vesículas de camada dupla concêntricas multilamelares (também denominadas vesículas multilamelares (MLVs)). MLVs apresentam geralmente diâmetros de 25 nm a 4  $\mu$ m. A sonificação de MLVs resulta na formação de pequenas vesículas unilamelares (SUVs) com diâmetros na faixa de 200 a 500 Å, contendo uma solução aquosa no núcleo.

Lipossomas apresentam similaridade com membranas celulares e são considerados para uso em conexão com a presente invenção como veículos para as composições peptídicas. Eles são amplamente vantajosos, porque substâncias solúveis em água e substâncias solúveis em lipídeos podem ser aprisionadas, i.e. nos espaços aquosos e no interior da própria camada dupla, respectivamente. É possível que os lipossomas portadores de droga podem ser empregados para o fornecimento específico-para-sítio de agentes ativos por meio de modificação seletiva da formulação lipossômica.

Adicionalmente aos ensinamentos de Couvreur *et al.* (1977; 1988), utilizar-se-á a informação a seguir para gerar formulações lipossômicas. Fosfolipídeos podem formar uma variedade de estruturas diferentes de lipossomas quando dispersas em água, dependendo da relação molar de lipídeo e água. A baixas taxas, o lipossoma é a estrutura preferida. As características físicas dos lipossomas dependem do pH, intensidade iônica e da presença de cátions divalentes. Lipossomas podem apresentar baixa permeabilidade para substâncias iônicas e polares, mas, a temperaturas elevadas, sofrem uma transição de fase que altera marcadamente sua permeabilidade. A transição de fase envolve uma alteração de uma estrutura menos ordenada, frouxamente empacotada, conhecida como o estado fluido. Isto ocorre a uma temperatura de transição de fase característica e resulta em um aumento da permeabilidade a íons, açúcares e drogas.

Adicionalmente à temperatura, a exposição a proteínas pode alterar a permeabilidade de lipossomas. determinadas proteínas solúveis, como citocromo c, ligam, deformam e penetram a camada dupla, causando com isso alterações na permeabilidade. O colesterol inibe esta penetração de proteínas, aparentemente por meio de empacotamento dos fosfolipídeos de maneira mais apertada. Considera-se que as formações de lipossomas mais úteis para fornecimento de antibióticos e inibidores conterão colesterol.

A capacidade de aprisionar solutos varia entre diferentes tipos de lipossomas. Por exemplo, MLVs são moderadamente eficientes para aprisionar solutos, mas SUVs são extremamente ineficientes. SUVs oferecem a vantagem de homogeneidade e reprodutibilidade na distribuição de tamanhos, no entanto, um equilíbrio entre tamanho e eficiência de aprisionamento é oferecido por grandes vesículas unilamelares (LUVs, *large unilamellar vesículas*). Estas são preparadas por meio de evaporação de éter e são de três a quatro vezes mais eficientes no aprisionamento de soluto do que MLVs.

Adicionalmente a características de lipossoma, uma importante determinante para aprisionar compostos compreende as propriedades físico-químicas do composto propriamente dito. Compostos polares são aprisionados nos espaços aquoso, e compostos não-polares ligam-se à camada dupla lipídica da vesícula. Compostos polares são liberados por meio de permeação ou quando a camada dupla é quebrada, mas compostos não-polares permanecem afiliados à camada dupla exceto quando são disrompidos por temperatura ou exposição a lipoproteínas. Ambos os tipos apresentam taxas de efluxo máximas na temperatura de transição de fase.

10 Lipossomas interagem com células por meio de quatro mecanismos diferentes: endocitose por células fagocíticas dos sistema reticuloendotelial, como macrófagos e neutrófilos; adsorção na superfície da célula, seja por meio de forças eletrostáticas ou hidrofóbicas fracas não-específicas, ou por meio de interações específicas com componentes de superfície celular; fusão com a membrana plasmática da célula por meio de inserção da camada dupla de lipídeo do lipossoma na membrana plasmática, com liberação simultânea de conteúdos lipossômicos no citoplasma; e por meio de transferência dos lipídeos lipossômicos para membranas celulares ou subcelulares, ou vice versa, sem qualquer associação dos conteúdos do lipossoma. Frequentemente é difícil determinar que organismo é operativo, e mais do que um podem operar ao mesmo tempo.

25 O destino e a disposição de lipossomas injetados intravenosamente dependem de suas atividades físicas, como tamanho, fluidez, e carga superficial. Eles podem persistir em tecidos durante horas ou dias, dependendo de sua composição, e meias-vidas no sangue compreendem de minutos até várias horas. Lipossomas maiores, como MLVs e LUVs, são recolhidos rapidamente por células fagocíticas do sistema reticulo endotelial, porém a fisiológica do sistema circulatório restringe a saída de espécies tão grandes na maior parte dos sítios. Eles só podem sair em locais em que

existem grandes aberturas ou poros no endotélio capilar, como os sinusóides do fígado ou do baço. Assim, estes órgãos são o sítio de absorção predominante. Por outro lado, SUVs apresentam uma distribuição mais ampla no tecido, mas ainda são seqüestradas em grau elevado no fígado e no baço.

5 De uma maneira geral, este comportamento *in vivo* limita a objetivação potencial de lipossomas apenas naqueles órgãos e tecidos acessíveis a seu grande tamanho. Estes incluem o sangue, fígado, baço, medula óssea, e órgãos linfóides.

10 De uma maneira geral, a objetivação não é uma limitação em termos da presente invenção. No entanto, desejando-se objetivação específica, estão disponíveis métodos para que isto seja realizado. Anticorpos podem ser usados para ligar a superfície do lipossoma e dirigir o anticorpo e seu conteúdo de droga para receptores antigênicos específicos localizados em uma superfície de um tipo de célula particular. Também é possível usar  
15 determinantes de carboidrato (componentes de superfície celular de glicoproteína ou glicolípido que desempenham um papel no reconhecimento, interação e adesão célula-a-célula) como sítios de reconhecimento porque possuem o potencial de dirigir lipossomas para tipos de células particulares. Em grande parte, considera-se a possibilidade de usar injeção intravenosa de  
20 preparações lipossômicas, mas também é possível considerar outras vias de administração.

Alternativamente, a invenção proporciona formulações de nanocápsulas farmacologicamente aceitáveis das composições da presente invenção. De uma maneira geral nanocápsulas podem aprisionar compostos  
25 de uma maneira estável e reproduzível (Henry-Michelland *et al*, 1987; Quintanar-Guerrero *et al*, 1998; Douglas *et al*, 1987). Para evitar efeitos colaterais devido a supercarregamento polimérico intracelular, referidas partículas ultrafinas (com tamanhos em torno de 0,1  $\mu\text{m}$ ) deveriam ser projetadas empregando-se polímeros que podem ser degradados *in vivo*.

Considera-se para uso na presente invenção nanopartículas de polialquilcianoacrilato biodegradáveis que atendem a estas exigências.

Referidas partículas podem ser preparadas com facilidade, como descrito (Couvreur *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.*, 1998; Pinto-Alphandry *et al.*, 1995 e Patente dos Estados Unidos nº 5.145.684, incorporada especificamente e integralmente aqui por referência).

### Vacinas

Em determinadas concretizações preferidas da presente invenção proporciona-se vacinas. As vacinas compreendem geralmente uma ou mais composições farmacêuticas, como aquelas discutidas acima, em combinação com um imuno-estimulante. Um imuno-estimulante pode ser qualquer substância que acentua ou potencializa uma resposta imunológica (mediada com anticorpo e/ou célula) para um antígeno exógeno. Exemplos de imuno-estimulantes incluem adjuvantes, microesferas biodegradáveis (p. ex., galactídeo poliláctico) e lipossomas (nos quais o composto é incorporado; ver, p. ex., Fullerton, Patente dos Estados Unidos nº 4.235.877). Preparação de vacina é descrita de uma maneira geral, por exemplo, em *Vaccine Design* (a abordagem de subunidade e adjuvante) por Powell & Newman, eds., (1995). Composições farmacêuticas e vacinas que se enquadram no escopo da presente invenção também podem conter outros compostos, que podem ser biologicamente ativos ou inativos. Por exemplo, uma ou mais porções imunogênicas de outros antígenos de tumor podem estar presentes, seja incorporadas em um polipeptídeo de fusão ou como um composto separado, na composição ou vacina.

Vacinas ilustrativas podem conter DNA que codifica um ou mais dos polipeptídeos como descrito acima, de tal forma que o polipeptídeo é gerado *in situ*. Como indicado acima, o DNA pode estar presente em qualquer um de uma variedade de sistemas de fornecimento conhecidos por aqueles com prática ordinária na arte, incluindo sistemas de expressão de

ácido nucleico, sistemas de expressão bacterianos e virais. Numerosas técnicas de fornecimento de gene são bem conhecidas na arte, como aquelas descritas por Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198 (1998), e referências ali indicadas. Sistemas de expressão de ácido nucleico

5 apropriados contêm as seqüências de DNA necessárias para expressão no paciente (como um promotor e sinal de terminação vantajosos). Sistemas de fornecimento bacterianos envolvem a administração de uma célula hospedeira de bactéria (por exemplo, uma cepa de *Mycobacterium*, *Bacillus* ou *Lactobacillus*, incluindo Bacillus-Calmette-Guerrin ou *Lactococcus lactis*)

10 que expressa uma porção imunogênica do polinucleotídeo sobre sua superfície celular ou que secreta um epitopo do tipo referido (ver, por exemplo, Ferreira, *et al.*, *An Acad Bras Cienc* (2005) 77:113-124; e Raha, *et al.*, *Appl Microbiol Biotechnol* (2005) PubMedID 15635459). Em uma concretização preferida, o DNA pode ser introduzido usando-se um sistema

15 de expressão viral (p. ex., vaccinia ou outro vírus da varíola, retrovírus, ou adenovírus), que podem envolver o uso de um vírus não-patogênico (defectivo), competente para replicação. Sistemas vantajosos são divulgados, por exemplo, por Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA* 86:317-321 (1989); Flexner *et al.*, *Ann. N Y. Acad. Sci.* 569:86-103 (1989); Flexner *et al.*,

20 *Vaccine* 8:17-21 (1990); Patentes dos Estados Unidos nums. 4.603.112, 4.769.330, e 5.017.487; WO 89/01973; Patente dos Estados Unidos nº 4.777.127; GB 2.200.651; EP 0.345.242; WO 91/02805; Berkner, *Biotechniques* 6: 616-627 (1988); Rosenfeld *et al.*, *Science* 252: 431-434 (1991); Kolls *et al.*, *Proc. Natl. Acad. Sci. USA* 91:215-219 (1994); Kass-

25 Eisler *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 11498-11502 (1993); Guzman *et al.*, *Circulation* 88: 2838-2848 (1993); e Guzman *et al.*, *Cir. Res.* 73:1202-1207 (1993). Técnicas para incorporar DNA em sistemas de expressão do tipo referido são bem conhecidos por aqueles com prática ordinária na arte. O DNA também pode ser "nu", como descrito, por exemplo, em Ulmer *et al.*,

*Science* 259: 1745-1749 (1993) e revisto por Cohen, *Science* 259: 1691-1692 (1993). A absorção de DNA nu pode ser incrementada por meio de revestimento do DNA sobre pérolas biodegradáveis, que são transportadas eficientemente nas células. Perceber-se-á que uma vacina pode compreender um componente de polinucleotídeo e um componente de polipeptídeo. Referidas vacinas podem proporcionar uma resposta imunológica acentuada.

Perceber-se-á que uma vacina pode conter sais farmacologicamente aceitáveis dos polinucleotídeos e polipeptídeos aqui proporcionados. Referidos sais podem ser preparados a partir de bases não-tóxicas farmacologicamente aceitáveis, incluindo bases orgânicas (p. ex., sais de aminas primárias, secundárias e terciárias e aminoácidos básicos) e bases inorgânicas (p. ex., sais de sódio, potássio, lítio, amônio, cálcio e magnésio).

Embora seja possível empregar qualquer veículo vantajoso conhecido por aqueles com prática ordinária na arte nas composições de vacina desta invenção, o tipo de veículo variará dependendo do modo de administração. Composições da presente invenção podem ser formuladas para qualquer modo apropriado de administração, incluindo, por exemplo, administração tópica, oral, nasal, intravenosa, intracraniana, intraperitoneal, subcutânea ou intramuscular. Para administração parenteral, como injeção subcutânea, o veículo compreende, de preferência, água, solução salina, álcool, uma gordura, uma cera ou uma tampão. Para administração oral, é possível empregar qualquer um dos veículos acima ou um veículo sólido, como manitol, lactose, amido, estearato de magnésio, sacarina de sódio, talco, celulose, glicose, sacarose, e carbonato de magnésio. Microesferas biodegradáveis (p. ex., poliglicolato de polilactato) também podem ser empregadas como veículos para as composições farmacêuticas desta invenção. Microesferas biodegradáveis vantajosas encontram-se descritas, por exemplo, nas Patentes dos Estados Unidos num. 4.897.268; 5.075.109; 5.928.647; 5.811.128; 5.820.883; 5.853.763; 5.814.344 e 5.942.252. Também



é possível empregar um veículo compreendendo os complexos de proteína-particulado descritos na Patente dos Estados Unidos nº 5.928.647, que são capazes de induzir respostas de linfócitos T citotóxicos restritos a classe I em um hospedeiro.

5                   Referidas composições também podem compreender tampões (p. ex., solução salina tamponada neutra ou solução salina tamponada com fosfato), carboidratos (p. ex., glicose, manose, sacarose ou dextranos), manitol, proteínas, polipeptídeos ou aminoácidos, como glicina, antioxidantes, bacteriostáticos, agentes quelantes, como EDTA ou glutathione, 10 adjuvantes (p. ex., hidróxido de alumínio), solutos que tornam a formulação isotônica, hipotônica ou fracamente hipertônica com o sangue de um recipiente, agentes de suspensão, agentes espessantes e/ou conservantes. Alternativamente, composições da presente invenção podem ser formuladas como um liofilizado. Compostos também podem ser encapsulados em 15 lipossomas com o uso de tecnologia bem conhecida.

                  É possível empregar qualquer um de uma variedade de imunostimulantes nas vacinas desta invenção. Por exemplo, um adjuvante pode ser incluído. A maior parte dos adjuvantes contém uma substância projetada para proteger o antígeno do catabolismo rápido, como hidróxido de alumínio ou 20 óleo mineral, e um estimulador de respostas imunológicas, como lipídeo A, *Bordetella pertussis* ou espécies de *Mycobacterium* ou proteínas derivadas de *Mycobacterium*. Por exemplo, é possível usar *M. vaccae* desglícolipidada, deslipidada, ("pVac"). Adjuvantes vantajosos são comercialmente obteníveis como, por exemplo, Adjuvante Incompleto de Freund e Adjuvante Completo 25 (Difco Laboratories, Detroit, MI); Adjuvante Merck 65 (Merck e Company, Inc., Rahway, NJ); ASO1B, AS02A, AS15, AS-2 e derivados dos mesmos (GlaxoSmithKline, Philadelphia, PA); CWS, TDM, Leif, sais de alumínio, como gel de hidróxido de alumínio (alume) ou fosfato de alumínio; sais de cálcio, ferro ou zinco; uma suspensão insolúvel de tirosina acilada; açúcares

acilados; polissacarídeos derivatizados cationicamente ou anionicamente; polifosfazenos; microesferas biodegradáveis; lipídeo A de monosfosforila e quil A. Citocinas, como GM-CSF ou interleucina-2, -7, ou -12, também podem ser usadas como adjuvantes.

5                    Nas vacinas aqui proporcionadas, a composição de adjuvante é projetada, de preferência, para induzir uma resposta imunológica predominantemente do tipo Th1. Níveis elevados de citocinas de tipo Th1 (p. ex., IFN- $\gamma$ , TNF $\alpha$ , IL-2 e IL-12) tendem a favorecer a indução de respostas imunológicas mediadas com células a um antígeno administrado. Em  
10                    contraste, altos níveis de citocinas de tipo Th2 (p. ex., IL-4, IL-5, IL-6 e IL-10) tendem a favorecer a indução de respostas imunológicas humorais. Após a aplicação de uma vacina como proporcionada aqui, um paciente suportará uma resposta imunológica que inclui respostas de tipo Th1 e Th2. Em uma concretização preferida, em que uma resposta é predominantemente do tipo  
15                    Th1, o nível de citocinas do tipo Th1 aumentará em maior grau do que o nível de citocinas do tipo Th2. Os níveis destas citocinas podem ser facilmente avaliados com o uso de ensaios convencionais. Para uma revisão das famílias de citocinas, ver Janeway, *et al*, *Immunobiology*, 5<sup>a</sup> edição, 2001.

                         Adjuvantes preferidos para uso na elicitação de um resposta  
20                    predominantemente de tipo Th1 incluem, por exemplo, uma combinação de lipídeo A de monosfosforila, de preferência, lipídeo A 3-O-desacilado de monosfosforila (3D-MPL), opcionalmente com um sal de alumínio (ver, por exemplo, Ribi, *et al*, 1986, *Immunology e Immunopharmacology of Bacterial Endotoxins*, Plenum Publ. Corp., NY, pp. 407-419; GB 2122204B; GB  
25                    2220211; e US 4.912.094). Uma forma preferida de 3D-MPL é a de uma emulsão apresentando um pequeno tamanho de partículas que é inferior a 0,2  $\mu$ m de diâmetro, e seu método de fabricação é divulgado no WO 94/21292. Formulações aquosas compreendendo lipídeo A de monosfosforila e um tensoativo foram descritas no WO 98/43670. Adjuvantes preferidos

exemplificados incluem ASO1B (MPL e QS21 em uma formulação de lipossoma), 3D-MPL e QS21 em uma formulação de lipossoma, AS02A (MPL e QS21 e uma emulsão óleo em água), 3D-MPL e QS21 e uma emulsão óleo em água, e AS 15, obtível da GlaxoSmithKline. Adjuvantes de MPL  
5 são obtíveis da GlaxoSmithKline, Seattle, WA (ver Patentes dos Estados Unidos num. 4.436.727; 4.877.611; 4.866.034 e 4.912.094).

Oligonucleotídeos contendo CpG (em que o dinucleotídeo CpG é não-metilado) também induzem uma resposta predominantemente Th1. CpG é uma abreviatura para unidades repetitivas de dinucleotídeo citosina-guanosina presentes no DNA. Referidos oligonucleotídeos são bem conhecidos e encontram-se descritos, por exemplo, no WO 96/02555, WO 99/33488 e Patentes dos Estados Unidos num. 6.008.200 e 5.856.462. Seqüências de DNA imunoestimuladoras também são descritas, por exemplo, por Sato *et al*, *Science* 273:352 (1996). CpG quando formulado em vacinas, é  
10 geralmente administrado em solução livre juntamente com o antígeno livre (WO 96/02555; McCluskie e Davis, *supra*) ou conjugado covalentemente a um antígeno (WO 98/16247), ou formulado com um veículo, como hidróxido de alumínio ((antígeno de superfície da hepatite) Davis *et al*. *supra*; Brazolot-Millan *et al.*, *Proc. Natl. Acad. Sci., USA*, 1998, 95(26), 15553-8). CpG é  
15 conhecido na arte como sendo um adjuvante que pode ser administrado por vias sistêmicas e mucosais (WO 96/02555, EP 468520, Davis *et al*, *J. Immunol.*, 1998, 160(2):870-876; McCluskie e Davis, *J. Immunol.*, 1998, 161(9):4463-6).

Outro adjuvante preferido é uma saponina ou mimético de saponina ou derivados, de preferência, QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), que podem ser usados sozinhos ou combinação com outros adjuvantes. Por exemplo, um sistema incrementado envolve a combinação de um lipídeo A de monosfosforila e derivado de saponina, como a combinação de QS21 e 3D-MPL como descrito em WO 94/00153, ou uma  
25

composição menos reatogênica em que o QS21 é extinto com colesterol, como descrito no WO 96/33739. Outras formulações preferidas compreendem uma emulsão óleo-em-água e tocoferol. Uma formulação adjuvante particularmente potente envolvendo QS21, 3D-MPL e tocoferol em uma emulsão óleo-em-água é descrita no WO 95/17210. Adjuvantes de saponina adicionais de uso na presente invenção incluem QS7 (descrito no WO 96/33739 e WO 96/11711) e QS17 (descrito na Patente dos Estados Unidos nº 5.057.540 e EP 0 362 279 B1).

Outros adjuvantes preferidos incluem Montanide ISA 720 (Seppic, França), SAF (Chiron, Califórnia, Estados Unidos), ISCOMS (CSL), MF-59 (Chiron), as séries SBAS de adjuvantes (p. ex., SBAS-2, AS2\ AS2", SBAS-4, ou SBAS6, obteníveis da GlaxoSmithKline, Rixensart, Bélgica), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) e outros aminoalquil glucosaminida 4-fosfatos (AGPs), como aqueles descritos nos Pedidos de Patentes dos Estados Unidos com números de série 08/853.826 e 09/074.720, cujas revelações são incorporadas aqui integralmente por referência.

Adjuvantes exemplares adicionais incluem MPL sintético e adjuvantes à base de subunidade B de toxina Shiga (ver WO2005/112991).

Qualquer vacina aqui proporcionada pode ser preparada usando-se métodos bem conhecidos que resultam numa combinação de antígeno, acentuador de resposta imunológica e um veículo ou excipiente vantajoso. As composições aqui descritas podem ser administradas como parte de uma formulação de liberação sustentada (i.e., uma formulação, como uma cápsula, esponja ou gel (constituído de polissacarídeos, por exemplo) que efetua uma baixa liberação de composto após a administração). Referidas formulações podem ser preparadas geralmente usando-se tecnologia bem conhecida (ver, p. ex., Coombes *et al*, *Vaccine* 14:1429-1438 (1996)) e administradas, por exemplo, por meio de implantação oral, retal ou

subcutânea, ou por meio de implante no sítio-alvo desejado. Formulações de liberação sustentada podem conter um polipeptídeo, polinucleotídeo ou anticorpo disperso em uma matriz de veículo e/ou contido em um reservatório envolvido por uma membrana controladora de taxa.

5                    Veículos para uso em tais formulações são biocompatíveis, e também podem ser biodegradáveis; de preferência, a formulação proporciona um nível relativamente constante de liberação de composto ativo. Referidos veículos incluem micropartículas de poli(lactídeo-co-glicolídeo), poliAcrilato, látex, amido, celulose, dextrano e análogos. Outros veículos de liberação  
10 retardada incluem biovetores supramoleculares, que compreendem um núcleo hidrofílico não-líquido (p. ex., um oligossarídeo ou polissacarídeo reticulado) e, opcionalmente, uma camada externa compreendendo um composto anfifílico, como um fosfolipídeo (ver, p. ex., Patente dos Estados Unidos nº 5.151.254 e pedidos PCT WO 94/20078, WO/94/23701 e WO 96/06638). A  
15 quantidade de composto ativo aqui contida, em uma formulação de liberação sustentada depende do sítio de implantação, da taxa e da duração esperada de liberação e da natureza da condição a ser tratada ou prevenida.

                    É possível empregar qualquer um de uma variedade de veículos de fornecimento nas composições farmacêuticas e vacinas para  
20 facilitar a produção de uma resposta imunológica específica para antígeno que objetiva células de tumor. Veículos de fornecimento incluem células apresentando antígeno (APCs), como células dendríticas, macrófagos, células B, monócitos e outras células que podem ser manipuladas para serem APCs eficientes. Referidas células podem ser, embora não necessariamente,  
25 modificadas geneticamente para incrementar a capacidade de apresentar o antígeno, para aperfeiçoar a ativação e/ou manutenção da resposta de células T, para apresentar efeitos anti-tumor *per se* e/ou para serem imunologicamente compatíveis com o receptor (i.e., haplótipo de HLA equiparado). De uma maneira geral, APCs podem ser isolados de qualquer um

de uma variedade de fluidos biológicos e órgãos, incluindo tecidos de tumor e peritumorais, e podem ser células autólogas, alogênicas, singênicas ou xenogênicas.

Determinadas concretizações preferidas da presente invenção usam células dendríticas ou progenitores das mesmas como células apresentadoras de antígeno. Células dendríticas são APCs altamente potentes (Banchereau & Steinman, *Nature* 392:245-251 (1998)) e mostrou-se que são eficazes como um adjuvante fisiológico para eliciar imunidade antitumor profilática ou terapêutica (ver Timmerman & Levy, *Ann. Rev. Med.* 50:507-529 (1999)). De uma maneira geral, células dendríticas podem ser identificadas com base em sua forma típica (estrelada *in situ*, com processos citoplasmáticos (dendritos) marcantes visíveis *in vitro*), sua capacidade de absorver, processar e apresentar antígenos com alta eficiência, e sua capacidade de ativar respostas de células T não-tratadas. Evidentemente, células dendríticas podem ser manipuladas de forma a expressar ligantes ou receptores de superfície celular específicos que não são comumente encontrados em células dendríticas *in vivo* ou *ex vivo*, e referidas células dendríticas modificadas são consideradas pela presente invenção. Como uma alternativa às células dendríticas, é possível usar células dendríticas carregadas com antígeno de vesículas secretadas (denominadas exossomas) na vacina (ver Zitvogel *et al.*, *Nature Med.* 4:594-600 (1998)).

Células dendríticas e progenitores podem ser obtidos do sangue periférico, medula óssea, células infiltradoras de tumor, células infiltradoras de tecidos peritumorosos, nódulos linfáticos, baço, pele, sangue do cordão umbilical e qualquer outro fluido ou tecido apropriado. Por exemplo, células dendríticas podem ser diferenciadas *ex vivo* por meio de adição de uma combinação de citocinas, como GM-CSF, IL-4, IL-13 e/ou TNF $\alpha$ , a culturas de monócitos colhidas de sangue periférico. Alternativamente, células positivas para CD34 colhidas de sangue periférico,

sangue do cordão umbilical ou medula óssea podem ser diferenciadas em células dendríticas por meio de adição no meio de cultura de combinações de GM-CSF, IL-3, TNF $\alpha$ , ligante de CD40, LPS, ligante de flt3 e/ou outro(s) composto(s) que induz(em) diferenciação, maturação e proliferação de células dendríticas.

Células dendríticas são categorizadas de maneira vantajosa como células "imaturas" e "maduras", o que fornece uma via simples para discriminar entre dois fenótipos bem caracterizados. No entanto, esta nomenclatura não deveria ser interpretada de forma a excluir todos os possíveis estágios intermediários de diferenciação. Células dendríticas imaturas são caracterizadas como APC com uma grande capacidade de absorção de antígeno e de processamento, o que se correlaciona com a alta expressão de receptor de Fc $\gamma$  e receptor de manose. O fenótipo maduro é caracterizado tipicamente por uma menor expressão destes marcadores, porém a elevada expressão de moléculas de superfície celular responsáveis pela ativação de células T, como MHC de classe I e classe II, moléculas de adesão (p. ex., CD54 e CD11) e moléculas co-estimuladoras (p. ex., CD40, CD80, CD86 e 4-1BB).

De uma maneira, APCs podem ser transfectadas com um polinucleotídeo codificando uma proteína (ou porção ou outra variante da mesma) de tal forma que o polipeptídeo, ou uma porção imunogênica do mesmo, é expresso sobre a superfície da célula. Referida transfecção pode ocorrer *ex vivo*, e, então, é possível usar uma composição ou vacina compreendendo referidas células transfectadas para fins terapêuticos, como descrito aqui. Alternativamente, um veículo de fornecimento de genes que objetiva uma célula dendrítica ou outra célula apresentadora de antígeno pode ser administrada a um paciente, resultando em transfecção que ocorre *in vivo*. Transfecção *in vivo* e *ex vivo* de células dendríticas, por exemplo, pode ser realizada geralmente usando-se métodos conhecidos na arte, como aqueles

descritos no WO 97/24447, ou a abordagem de canhão de genes descrita por Mahvi *et al*, *Immunology and Cell Biology* 75:456-460 (1997). O carregamento de antígeno de células dendríticas pode ser obtido por meio de incubação de células dendríticas ou células progenitoras com o polipeptídeo, DNA (nu ou com um vetor de plasmídeo) ou RNA; ou com vírus ou bactéria recombinante expressando antígeno (p. ex., vetores de *vaccinia*, varíola das aves, adenovírus ou lentivírus). Antes do carregamento, o polipeptídeo pode ser conjugado covalentemente com um parceiro imunológico que proporciona auxílio de células T (p. ex., uma molécula de veículo). Alternativamente, uma célula dendrítica pode ser pulsada com um parceiro imunológico não-conjugado, separadamente ou na presença do polipeptídeo.

Vacinas e composições farmacêuticas podem ser apresentadas em recipientes de dose unitária ou de doses múltiplas, como frascos ou ampolas fechadas hermeticamente. De preferência, referidos recipientes são fechados hermeticamente para preservar a esterilidade da formulação até o uso. De uma maneira geral, formulações podem ser armazenadas como suspensões, soluções ou emulsões em veículos oleosos ou aquosos. Alternativamente, uma vacina ou composição farmacêutica pode ser armazenada em uma condição secada com congelamento requerendo apenas a adição de um veículo líquido estéril imediatamente antes do uso.

Todas as publicações e pedidos de patentes citados nesta descrição são incorporados aqui por referência, como se cada publicação individual ou pedido de patente fosse indicado especificamente e individualmente como sendo incorporado por referência.

Embora a invenção precedente houvesse sido descrita em algum detalhe a título de ilustração e de exemplo objetivando maior clareza e compreensão, alguém com prática ordinária na arte perceberá com facilidade, à luz dos ensinamentos desta invenção, que é possível realizar determinadas alterações e modificações na mesma sem afastar-se do espírito ou escopo das



reivindicações anexas.

### Exemplos

Os exemplos a seguir são fornecidos apenas a título de ilustração e não como limitação. Aqueles com prática na arte reconhecerão facilmente uma variedade de parâmetros não-críticos que poderiam ser alterados ou modificados para fornecer resultados essencialmente similares.

#### Exemplo 1: Preparação de Mtb72f (sem marcador His) (SEO ID No: 6)

##### Construção do vetor de expressão de Mtb72f

Mtb72f é uma proteína de fusão formada de 2 proteínas de *Mycobacterium tuberculosis* Mtb32 e Mtb39. Mtb72f é construída fundindo-se Mtb39 e as porções N e C terminais de Mtb32 como a seguir: Mtb32 extremidade C-terminal - Mtb39 - Mtb32 extremidade N-terminal. Particularmente, a proteína Mtb72f foi gerada por meio da ligação seqüencial, em tandem, das matrizes de leitura aberta (ORFs) codificando o fragmento C-terminal com ~14 kDa de Mtb32 (radicais de 192 a 323; 132 aminoácidos) até a ORF de comprimento pleno de Mtb39 seguido, na extremidade C, da porção N-terminal com ~20 (radicais de 1 a 195) de Mtb32. Isto foi realizado usando-se oligonucleotídeos específicos para seqüência contendo sítios de restrição únicos (EcoRI e EcoRV) e isentos dos códons de interrupção nas extremidades C-terminais (no caso de Mtb32-C e Mtb39) para reação em cadeia de polimerase (PCR) fora do DNA genômico da cepa H37Rv de *M. tuberculosis*. Os detalhes do processo são como a seguir:

Primeiramente, o DNA que codifica a porção C-terminal de Mtb32 (Mtb32C) foi clonado de H37Rv usando-se PCR com os seguintes oligonucleotídeos: 5' (5'-CAA-TTA-CAT-ATG-CAT-CAC-CAT- CAC-CAT-CAC-ACG-GCC-GCG-TCC-GAT-AAC-TTC-3') e 3' (5'-CTA-ATC-GAA-TCC- GGC-CGG-GGG-TCC-CTC-GGC-CAA-3'). O oligonucleotídeo a 5' contido no sítio de restrição NdeI (sublinhado) compreendendo o códon de iniciação ATG. O oligonucleotídeo 3' contido no sítio de restrição EcoRI

(sublinhado). Estes oligonucleotídeos foram usados para amplificar Mtb32C, uma porção de 396 nucleotídeos de Mtb32 e o produto resultante foi subclonado nos sítios NdeI e *EcoRI* de um vetor de expressão. Digestão realizada com *EcoRI* e *EcoRV* linearizou subseqüentemente o plasmídeo de  
5 Mtb32C.

Para Mtb39, usou-se os seguintes oligonucleotídeos para amplificação com PCR e clonagem: 5'- (5'-CTA-ATC-GAA-TTC-ATG-GTG-GAT-TTC-GGG-GCG-TTA-3') e 3' (5'-CTA- ATC-GAT-ATC-GCC-GGC-TGC-CGG-AGA-ATG-CGG-3'). O oligonucleotídeo a 5' continha um  
10 sítio de restrição *EcoRI* (sublinhado) enquanto que o oligonucleotídeo a 3' continha um sítio de restrição *EcoRV* (sublinhado). A seqüência codificante de comprimento pleno de Mtb39 foi amplificada, digerida, e subclonada na matriz a jusante de Mtb32c usando-se o plasmídeo pré-digerido da primeira etapa.

Os oligonucleotídeos a 5' e 3' do fragmento N-terminal de  
15 Mtb32 foram projetados como a seguir: 5'- (5'-CTA-ATC-GAT-ATC-GCC-CCG-CCG-GCC-TTG-TCG-CAG-GAC-3') e 3' (5'- CTA-ATC-GAT-ATC-CTA-GGA-CGC-GGC-CGT-GTT-CAT-AC-3'). Ambos os conjuntos de oligonucleotídeos continham um sítio de restrição *EcoRV* (sublinhado)  
20 enquanto que o oligonucleotídeo a 3' também incluía um códon de interrupção (itálico). Os oligonucleotídeos foram projetados para amplificar uma porção de Mtb32 com 585 bp [pares de bases] codificando o domínio N-terminal predito desta proteína. O produto de PCR resultante foi subclonado no plasmídeo de fusão de Mtb32c-Mtb39. A orientação apropriada de insertos e a  
25 ausência de mutações foi então verificada por meio de seqüenciamento de DNA. Para a construção final, usada para realizar o Banco de Células Mestre [*Master Cell Bank*] e o Banco de Células de Trabalho do Fabricante [*Manufacturer's Working Cell Bank*], o marcador de afinidade 6xHis foi removido por meio de PCR, e a matriz de leitura aberta (ORF) para Mtb72f

foi subclonada em pPDM, um vetor de expressão derivado de pET. A ORF codifica para uma poliproteína com cerca de 72 kDa (Mtb72f), com domínios organizados na ordem linear: Mtb32C-Mtb39-Mtb32N. Este DNA foi então transformado na cepa HMS174 pLysS de *Escherichia coli* e usado para testes, criação de bancos de células, e fabricação.

#### Produção de substância de droga volumétrica de Mtb72f

O processo de fabricação para a produção de Mtb72f é resumido como a seguir:

- Fermentação seguida de colheita por centrifugação, 10  
disrupção de células (microfluidizador) e centrifugação dando um pellet de corpos de inclusão;

- Purificação do pellet de corpos de inclusão por meio de 15  
extração em uréia 8M, seguido de cromatografia Q Sepharose Fast Flow (QFF), cromatografia de Hidroxiapatita Cerâmica (sigla em inglês: CHT), diafiltração, e filtração esterilizante dando a substância de droga volumétrica purificada.

#### Fermentação

Fermentações são realizadas a um volume de trabalho de 10 l. O fermentador é inoculado com 300 ml de uma cultura de frascos agitados 20  
das células de semente de trabalho desenvolvidas a 37°C de um dia para o outro. Ambos, o inóculo e a fermentação, usam um meio semi-definido com glicerol derivado de planta como a fonte primária de carbono. A composição do meio é mostrada na tabela abaixo. Todos os componentes do meio são esterilizados por meio de aquecimento a 121°C durante 20 minutos ou por 25  
meio de filtração esterilizante. Durante a fermentação o fermentador é mantido a uma temperatura de 37°C. Asperge-se ar a uma taxa de 5 litros padrão por minuto (SLPM). O pH do meio é mantido a 7,0 por meio de adição automática de ácido (H<sub>2</sub>SO<sub>4</sub>) ou base (NaOH). O fermentador é programado para controlar o oxigênio dissolvido a 30% por meio de ajuste

automática da agitação, enquanto se mantinha uma agitação mínima de 200 rpm. O controle da espuma no fermentador é obtido por meio da adição automática de 1,05% de antiespumante de silicone SAG-471 (Witco Corp.). Quando a densidade da célula atinge uma densidade óptica (600 nm) de aproximadamente 3,5, adiciona-se isopropil-beta-D-tiogalactopiranosídeo (IPTG) é no fermentador até uma concentração de 1,0 mM. O IPTG induz expressão do gene recombinante que codifica a proteína Mtb72f. Após 3,0 horas da indução, o fermentador é resfriado e as células são colhidas por meio de centrifugação em frascos de centrifugação de 1 litro.

#### 10 Composição do meio de fermentação

Material	Concentração
Extrato de levedura	15 g/l
Glicerol	30 g/l
Sulfato de magnésio, heptaidrato (MgSO <sub>4</sub> )	0,5 g/l
Fosfato de potássio, monobásico (KH <sub>2</sub> PO <sub>4</sub> )	2,4 g/l
Cloreto de amônio (NH <sub>4</sub> Cl)	1,0 g/l
Cloreto de sódio (NaCl)	0,5 g/l
Sulfato de canamicina	30 mg/l
Cloranfenicol	34 mg/l
Silicone antiespumante SAG-471 (Witco Corp.)	0,0005% (v/v) (não incluído)

#### Isolamento de corpos de inclusão

As pelotas de células são suspensas e combinadas em 2,3 l de tampão de lise (50 mM de NaCl, 10 mM de Tris pH 8,0), e usa-se um microfluidizador M-1 10Y Microfluidizer® para disromper as células. As células são passadas através do Microfluidizador cinco vezes a uma pressão de 11.000 ± 1.000 psi [75900 kPa ± 6900 kPa]). A suspensão é centrifugada a 8000 x g em frascos de 500 ml. Nestas condições, os corpos de inclusão (IB, *inclusion bodies*) contendo a proteína Mtb72f são pelotizados, enquanto que a maior parte dos fragmentos de células permanece no sobrenadante. As pelotas de IB são ressupensas em tampão de lavagem (uréia 2 M, 50 mM de NaCl, 10 mM de Tris pH 8,0), seguido de centrifugação a 8.000 g. As frações de sobrenadante são descarregadas e as pelotas de IB são armazenados a de -

70°C a -80°C até necessário para purificação adicional.

### Purificação de poliproteína

As preparações de IB congeladas são orvalhadas a 37°C durante 15 minutos e, depois, ressuspensas em uréia 8 M, 50 mM de NaCl, 20 mM de Bis-tris propano, pH 7,0 (tampão A) usando-se agitação mecânica delicada. Os IBs ressuspensos são então agitados à temperatura ambiente com uma barra de agitação magnética a 300 rpm durante 2 horas. O extrato de IB é então centrifugado a alta velocidade, e a fração de sobrenadante resultante é filtrada através de um filtro de 0,45 µm (Pall, Supor) antes do fracionamento cromatográfico.

O extrato de IB é aplicado numa coluna contendo resina de troca de ânion Q Sepharose Fast Flow (QFF) (10 x 12,5 cm da Amersham/Pharmacia BPG; leito empacotado de 1 litro) previamente sanitizado com NaOH 1 N e então equilibrado com tampão A. A coluna é revelada a uma taxa de fluxo linear de 60 cm/hora com tampão A e o fluxo passante contendo predominantemente menos contaminantes em massa é coletado por referência. O volume da Mtb72f é eluído em uma única etapa usando uréia 8 M, 90 mM de NaCl, 20 mM de Bis-tris propano, pH 7,0 e é coletado como um único pico volumétrico com base na absorbância.

Resinas QFF são resinas de agarose altamente reticuladas com um grupo funcional de amina quaternária que é carregado positivamente nas condições usadas durante a purificação. A matriz carregada permite a ligação dos vários ânions que, então, podem ser eluídos seletivamente usando-se um gradiente de sal. Esta cromatografia de troca de ânion é usada para separar ácidos nucleicos e endotoxina, que se ligam estreitamente à resina, a partir da proteína, que é ligado mais fracamente e elui antes destes contaminantes. Adicionalmente, esta etapa remove contaminantes não-carregados e uma grande parte das impurezas de proteína.

O pico de eluído de 90 mM de NaCl é da coluna QFF é

aplicado numa coluna (2,6 x 12 cm Amersham/Pharmacia XK26/20; 63 ml de leite empacotado) contendo hidroxiapatita cerâmica (CHT) MacroPrep® (tipo I, 40 µM, BioRad) previamente sanitizada usando-se NaOH 1 N e, então, equilibrada com tampão C (uréia 8 M, 250 mM de NaCl, e 20 mM de Bis-tris propano, pH 7,0). O material de fluxo passante (FT1) contendo a maior parte da Mtb72f, livre de contaminantes, é recolhido. A coluna é lavada com tampão C e qualquer material absorvente de UV resultante é recolhido. Por fim, a coluna é eluída em tampão D (8 M de uréia, 200 mM de fosfato de sódio, pH 7,4).

10 CHT MacroPrep® é uma forma macroporosa esférica de hidroxiapatita  $[Ca_5(PO_4)SOH]_2$ . Cromatografia CHT pode ser um método altamente seletivo de purificação se as condições apropriadas de ligação e eluição forem encontradas. Os modos de ligação incluem ligação de tipo troca de íon com íons carregados de cálcio e fosfato e também quelação de moléculas. DNA se ligará a esta resina e é possível obter alta seletividade para proteínas individuais. As condições usadas para a purificação de Mtb72f servem como uma etapa de polimento permitindo remoção virtualmente completa de contaminantes de células hospedeiras detectáveis.

20 Durante separações cromatográficas, monitora-se e registra-se absorbância ultravioleta (UV), condutividade, pressão, pH, taxa reduzida, e temperatura ambiente. Usa-se o material inicial de fluxo passante de CHT (FT1) para processamento adicional a jusante.

#### Diafiltração e filtração estéril

25 Diafiltração é realizada no combinado de CHT FT1 para remover a uréia e substituir o tampão por 20 mM de Tris, pH 7,5. A diafiltração é realizada usando-se um sistema Pall Minim™ com um dispositivo de filtração de fluxo tangencial LV-Centramate™ com uma membrana de ultrafiltração com corte de peso molecular de 30 kDa (sigla em inglês: MWCO). A solução de Mtb72f em 20 mM de Tris, pH 7,5, é

esterilizada por meio de filtração usando-se um filtro esterilizante de 0,2 um (Millipak 40). Cinquenta ml da solução são distribuídos em frascos de 60 ml de meio de PETG estéril (copolímero de polietileno tereftalato), depois congelados e armazenados a -70°C. Este material é a substância de droga volumétrica purificada Mtb72f.

#### Exemplo 2: Preparação de Mtb72f (marcador 6 His) (SEQ ID No: 2)

O método do Exemplo 1 pode ser seguido, exceto pelo fato de que a etapa de subclonagem em pPDM para remover o marcador His é omitida.

#### 10 Exemplo 3: Preparação de M72 (marcador 2 His) (SEQ ID No: 4)

##### Construção do vetor de expressão M72

Material de partida para a construção do antígeno M72 foi o plasmídeo recombinante 6His-Mtb72fmut. 6His-Mtb72fmut foi preparado por meio de mutagênese direcionada para sítio usando-se o plasmídeo recombinante 6his-Mtb72f (ver Exemplo 1) como modelo. Mutagênese direcionada para sítio envolveu substituir o códon para Ser na posição 710 na SEQ ID No: 1 por um códon para Ala.

A deleção de quatro histidinas N-terminais presentes na construção 6His-Mtb72fmut (plasmídeo Corixa) foi obtida com "Gene Tailor Site-Directed Mutagenesis System" (Invitrogen), levando à esperada construção 2His-Mtb72Fmut. Após verificação de seqüência, a seqüência codificante de 2His-Mtb72fmut foi excisada do plasmídeo (por meio de restrição enzimática), purificada em gel e ligada no vetor de expressão pET29a resultando no plasmídeo recombinante final pET29a/ 2His-Mtb72fmut. Após verificação de seqüência, o plasmídeo recombinante recebeu a designação oficial pRIT 15497 e usado para transformar células hospedeiras HMS174(DE3). pRIT15497 codifica para uma proteína com 725 aminoácidos denominada M72.

##### Produção de proteína M72

É possível empregar o mesmo processo de produção como descrito para Mtb72f (ver Exemplo 1), exceto que, para a produção de M72, cloranfenicol está ausente no meio de fermentação.

Exemplo Biológico 1: Um modelo de camundongo de um estado inativo/latente de infecção com *M. tuberculosis*

5 Para estabelecer um modelo de mouse de infecção latente com *M. tuberculosis* usou-se a cepa SWR. Camundongos SWR não são imunocomprometidos, mas são deficientes para secreção de componente de complemento C5 (ver, Ooi e Colten, *Nature* (1979) 282:207-8).

10 Camundongos SWR são incapazes de estabelecer um estado crônico de infecção com Mtb, mas desenvolvem pneumonia granulomatosa difusa caracterizada por grandes macrófagos espumosos e epitelióides com inclusões cristalóides (grânulos derivados de eosinófilos ou neutrófilos que foram submetidos a fagocitose), necrose multifocal, acúmulo escasso de neutrófilos

15 e linfócitos (ver, Turner, *et al.*, *J Submicrosc Cytol Pathol.* (2001) 33(1-2):217-9; e Turner, *et al.*, *Infect Immun.* (2003) 71(9):5266-72). A seguir encontra-se o protocolo para uso da cepa de camundongo Swiss Webster (SWR/J) em um modelo de infecção latente com *M. tuberculosis* para se avaliar a eficácia terapêutica de Mtb72f (SEQ ID No:6) formulada com

20 adjuvante de AS01B. AS01B de intensidade dupla é preparado por meio de adição de QS21 (5 µg) a vesículas unilamelares pequenas (SUV, *small unilamellar vesicles*) de fosfatidilcolina de dioleíla (100 µg) contendo colesterol (25 µg) (WO 96/33739) e lipídeo A de monosfosforila (MPL) (5 µg) na membrana (ver, Publicação de Patente dos E.U.A. nº 2003/0143240).

25 Prepara-se uma fração para injeção (50 µl) por meio de misturação de 4 µg de proteína em tampão (PBS pH 6,8) com 50 µl de AS01B de dupla intensidade. Cada camundongo recebeu duas injeções de 50 µl (i.e. 8 µg de proteína).

Uma linha-do-tempo representativa para estabelecer um modelo de uma infecção latente de *M. tuberculosis* é ilustrada



esquemáticamente na Figura 1.

Dia 1: Infectar via aerossol com de 50 a 100 unidades formadoras de colônia (CFU) de organismos de *M. tuberculosis*

5 Dias de 30 a 90: Tratar um subconjunto de camundongos com 50 mg de rifampina/85 mg de isoniazida por litro de água potável

Dia 61: Todos os camundongos que receberam a vacina candidata 5 deveriam ser imunizados com rMtb72f + AS01B

Dia 82: Todos os camundongos que receberam a vacina candidata deveriam ser imunizados com rMtb72f + AS01B

10 Dia 103: Todos os camundongos que receberam a vacina candidata deveriam ser imunizados com rMtb72f + AS01B

Dia 113: Sangramento para ensaios de IgG

Em vários momentos: Retirar baços e pulmões para enumeração de CFU e imunogenicidade

15 Varição 1 → Tratar com quimioterapia durante 60 dias. Começando no dia 30 → Descansar durante 3, 4, 5 meses → CFU em 2 camundongos em cada momento e deixar de 4 a 7 camundongos para estudos de sobrevida

20 Varição 2 → Tratar com quimioterapia durante 90 dias. Começando no dia 30 → Descansar durante 4, 5 meses → CFU em 2 camundongos em cada momento e deixar 7 camundongos para estudos de sobrevida

25 Varição 3 → Descansar durante 4, 5, 6 meses → CFU em 2 camundongos em cada momento e deixar 4 camundongos para estudos de sobrevida

Varição 4 → Tratar com quimioterapia durante 60 dias. Começando no dia 30 → 3 imunizações com r72F+AS01B intramuscularmente (i.m.) começando no Dia 60 → Descansar durante 3, 4, 5 meses → CFU em 2 camundongos em cada momento e deixar de 4 a 7

camundongos para estudos de sobrevivida

Varição 5 → Tratar com quimioterapia durante 90 dias. Começando no dia 30 → 3 imunizações com r72F+AS01B i.m começando no Dia 60 → Descansar durante 4, 5 meses → CFU em 2 camundongos em cada momento e deixar de 4 a 7 camundongos par estudos de sobrevivida

Análise de respostas de anticorpos após infecção usando rMtb72f para revestir as placas de ELISA revelou que aqueles grupos que receberam uma combinação de quimioterapia e imunização com Mtb72f+AS01B apresentaram uma maior resposta de anticorpos (OD de até 2,0) do que camundongos que não foram tratados ou que receberam tratamento com quimioterapia apenas (OD inferior a 0,5) (Figura 2). Camundongos imunizados com Mtb72f apresentaram uma resposta mensurável de anticorpos específicos para Mtb72f (OD entre 1,5 e 2,5) quer tenham recebido quimioterapia durante 60 ou 90 dias (Figura 3).

Células de baço foram retiradas de camundongos em diversos intervalos após os camundongos terem sido infectados com *M. tuberculosis*. Os esplenócitos foram reestimulados *in vitro* com antígenos recombinantes para medir a secreção de IFN- $\gamma$ . Níveis de IFN- $\gamma$  produzidos por estas células foram uniformemente negligenciáveis nos grupos 1 (não-tratado) e 2 (quimioterapia apenas) no dia 60, com exceção de Mtb39. Estimulações de controle positivo com lisado de conA, PPD e BCG demonstraram que as células foram capazes de sintetizar e secretar IFN- $\gamma$  em resposta a outras moléculas estimuladoras (Figura 4). Níveis de IFN- $\gamma$  foram elevados em grupos que receberam Mtb72f+AS01B, mas foram baixos ou negligenciáveis em grupos que não foram imunizados com Mtb72f+AS01B, quer tenham recebido quimioterapia ou não (Figura 5).

Durante o curso da infecção de tuberculose e o tratamento subsequente, células T específicas respondem. Usando-se manchamento de citocina intracelular para IFN- $\gamma$  mediu-se o percentual de respostas a Mtb72f

por parte de células CD4<sup>+</sup> específicas (Figura 6). Pareceu não haver alteração nas respostas de células T CD4<sup>+</sup> IFN  $\gamma$ <sup>+</sup> específicas para Mtb72F durante o curso de aplicação de quimioterapia apenas em qualquer momento, conforme medido por meio deste ensaio (Figura 7). No dia 120 após infecção com Mtb, uma tendência de reposta CD4<sup>+</sup> IFN $\gamma$ <sup>+</sup> a Mtb72F em grupos que receberam a vacina de Mtb72f mais AS01B pareceu aumentar com a prolongação do tempo em quimioterapia (Figura 7).

Os resultados de nossos experimentos demonstram que camundongos SWR são suscetíveis a infecção com *M. tuberculosis*. Se deixados não-tratados, camundongos SWR morrem em torno de 115 dias após a infecção com Mtb (Figuras 8 e 9). O tempo médio de sobrevivência para camundongos que receberam 60 dias de quimioterapia de combinação foi de 170 dias (Figuras 8 e 9). O tempo médio de sobrevivência para camundongos que receberam 60 dias de quimioterapia de combinação e 3 imunizações de Mtb72f/AS01B foi de 215 dias (Figuras 8 e 9). A sobrevivência para o grupo de camundongos que recebeu quimioterapia é significativamente diferente (95% de intervalo de confiança ( $p=0,0067$ )) daqueles que receberam quimioterapia e a vacina Mtb72f/AS01B.

LISTAGEM DA SEQUÊNCIA

<110> Coler, Rhea N  
 Reed, Steven G  
 Lobet, Yves

<120> MÉTODO INÉDITO PARA PREVENIR OU TRATAR INFECÇÃO  
 COM M TUBERCULOSIS

<130> VB61507

<160> 6

<170> PatentIn versão 3.3

<210> 1

<211> 2287

<212> DNA

<213> Artificial

<220>

<223> Descrição de Sequência Artificial: proteína de tri-fusão Mtb72F  
 (Ra12-TbH9-Ra35 ou Mtb32-Mtb39)

<220>

<221> caract. misc.

<222> (30)..(30)

<223> n é a, c, g, ou t

<220>

<221> caract. misc.

<222> (33)..(33)

<223> n é a, c, g, ou t

<220>

<221> CDs

<222> (42)..(2228)

<220>

<221> caract. misc.

<222> (2270)..(2270)

<223> n é a, c, g, ou t

<400> 1

tctagaata attttgttta cttaagaan ganatatata t atg cat cac cat cac  
 56

Met His His His His  
 1 5

cat cac acg gcc gcg tcc gat aac ttc cag ctg tcc cag ggt ggg cag  
 104

His His Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln  
 10 15 20



tcg tgg ata ggt tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tgg ccg  
 680  
 Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro  
 200 205 210

tat gtg gcg tgg atg agc gtc acc gcg ggg cag gcc gag ctg acc gcc  
 728  
 Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala  
 215 220 225

gcc cag gtc cgg gtt gct gcg gcg gcc tac gag acg gcg tat ggg ctg  
 776  
 Ala Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu  
 230 235 240 245

acg gtg ccc ccg ccg gtg atc gcc gag aac cgt gct gaa ctg atg att  
 824  
 Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile  
 250 255 260

ctg ata gcg acc aac ctc ttg ggg caa aac acc ccg gcg atc gcg gtc  
 872  
 Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val  
 265 270 275

aac gag gcc gaa tac ggc gag atg tgg gcc caa gac gcc gcc gcg atg  
 920  
 Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Met  
 280 285 290

ttt ggc tac gcc gcg gcg acg gcg acg gcg acg gcg acg ttg ctg ccg  
 968  
 Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro  
 295 300 305

ttc gag gag gcg ccg gag atg acc agc gcg ggt ggg ctc ctc gag cag  
 1016  
 Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln  
 310 315 320 325

gcc gcc gcg gtc gag gag gcc tcc gac acc gcc gcg gcg aac cag ttg  
 1064  
 Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu  
 330 335 340

atg aac aat gtg ccc cag gcg ctg caa cag ctg gcc cag ccc acg cag  
 1112  
 Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln  
 345 350 355

ggc acc acg cct tct tcc aag ctg ggt ggc ctg tgg aag acg gtc tgg  
 1160  
 Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser  
 360 365 370

ccg cat cgg tgg ccg atc agc aac atg gtg tgg atg gcc aac aac cac  
 1208  
 Pro His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His  
 375 380 385

atg tgg atg acc aac tgg ggt gtg tgg atg acc aac acc ttg agc tgg  
 1256  
 Met Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser  
 390 395 400 405

atg ttg aag ggc ttt gct ccg gcg gcg gcc gcc cag gcc gtg caa acc  
 1304  
 Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr  
 410 415 420

gcg gcg caa aac ggg gtc cgg gcg atg agc tgg ctg ggc agc tgg ctg  
 1352  
 Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu  
 425 430 435

ggt tct tgg ggt ctg ggc ggt ggg gtg gcc gcc aac ttg ggt cgg gcg  
 1400  
 Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala  
 440 445 450

gcc tgg gtc ggt tgg ttg tgg gtg ccg cag gcc tgg gcc gcg gcc aac  
 1448  
 Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn  
 455 460 465

cag gca gtc acc ccg gcg gcg cgg gcg ctg ccg ctg acc agc ctg acc  
 1496  
 Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr  
 470 475 480 485

agc gcc gcg gaa aga ggg ccc ggg cag atg ctg gcc ggg ctg ccg gtg  
 1544  
 Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val  
 490 495 500

ggg cag atg gcc gcc agg gcc ggt ggt ggg ctc agt ggt gtg ctg cgt  
 1592  
 Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg  
 505 510 515

gtt ccg ccg cga ccc tat gtg atg ccg cat. tct ccg gca gcc ggc gat  
 1640  
 Val Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp  
 520 525 530

atc gcc ccg ccg gcc ttg tgg cag gac cgg ttc gcc gac ttc ccc gcg  
 1688  
 Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala  
 535 540 545

ctg ccc ctc gac cag tcc gcg atg gtc gcc caa gtg ggg cca cag gtg  
 1736  
 Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val  
 550 555 560 565

gtc aac atc aac acc aaa ctg ggc tac aac aac gcc gtg ggc gcc ggg  
 1784  
 Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly  
 570 575 580

acc ggc atc gtc atc gat ccc aac ggt gtc gtg ctg acc aac aac cac  
 1832  
 Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His  
 585 590 595

gtg atc gcg ggc gcc acc gac atc aat gcg ttc agc gtc ggc tcc ggc  
 1880  
 Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly  
 600 605 610

caa acc tac ggc gtc gat gtg gtc ggg tat gac cgc acc cag gat gtc  
 1928  
 Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val  
 615 620 625

gcg gtg ctg cag ctg cgc ggt gcc ggt ggc ctg ccg tcg gcg gcg atc  
 1976  
 Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile  
 630 635 640 645

ggt ggc ggc gtc gcg gtt ggt gag ccc gtc gtc gcg atg ggc aac agc  
 2024  
 Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser  
 650 655 660

ggt ggg cag ggc gga acg ccc cgt gcg gtg cct ggc agg gtg gtc gcg  
 2072  
 Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala  
 665 670 675

ctc ggc caa acc gtg cag gcg tcg gat tcg ctg acc ggt gcc gaa gag  
 2120  
 Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu  
 680 685 690

aca ttg aac ggg ttg atc cag ttc gat gcc gcg atc cag ccc ggt gat  
 2168  
 Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp  
 695 700 705

tcg ggc ggg ccc gtc gtc aac ggc cta gga cag gtg gtc ggt atg aac  
 2216  
 Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn  
 710 715 720 725



acg gcc gcg tcc taggatatcc atcacactgg cggccgctcg agcagatccg  
 2268  
 Thr Ala Ala Ser

gntgtaacaa agcccgaaa  
 2287

<210> 2

<211> 729

<212> PRT

<213> Artificial

<220>

<223> Descrição de Sequência Artificial: proteína de tri-fusão Mtb72F  
 (Ral2-TbH9-Ra35 ou Mtb32-Mtb39)

<400> 2

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu  
 1 5 10 15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala  
 20 25 30

Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile  
 35 40 45

Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn  
 50 55 60

Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu  
 65 70 75 80

Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile  
 85 90 95

Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly  
 100 105 110

Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr  
 115 120 125

Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp  
 130 135 140

Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly  
 145 150 155 160

Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val  
 165 170 175

Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp  
 180 185 190

Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val  
 195 200 205

Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln  
 210 215 220

Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu  
 225 230 235 240

Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg  
 245 250 255

Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr  
 260 265 270

Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln  
 275 280 285

Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr  
 290 295 300

Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly  
 305 310 315 320

Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala  
 325 330 335

Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu  
 340 345 350

Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu

355

360

365

Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn Met Val Ser  
 370 375 380

Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val Ser Met Thr  
 385 390 395 400

Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala  
 405 410 415

Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser  
 420 425 430

Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala  
 435 440 445

Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala  
 450 455 460

Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro  
 465 470 475 480

Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu  
 485 490 495

Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu  
 500 505 510

Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met Pro His Ser  
 515 520 525

Pro Ala Ala Gly Asp Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe  
 530 535 540

Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln  
 545 550 555 560

Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn  
 565 570 575

Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val  
580 585 590

Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe  
595 600 605

Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp  
610 615 620

Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu  
625 630 635 640

Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val  
645 650 655

Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro  
660 665 670

Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu  
675 680 685

Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala  
690 695 700

Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln  
705 710 715 720

Val Val Gly Met Asn Thr Ala Ala Ser  
725

<210> 3

<211> 2178

<212> DNA

<213> Artificial

<220>

<223> Descrição de Sequência Artificial: variante de proteína de tri-  
fusão

Mtb72F (Ra12-TbH9-Ra35 ou Mtb32-Mtb39)

<220>

<221> caract. misc.

<222> (4)..(9)

<223> cat cac' em lugar de 'cat cac cat cac cat cac'

<220>

<221> caract. misc.

<222> (2116)..(2118)

<223> 'gcg' em lugar de 'tcg'

<400> 3

atgcatcaca eggccgcgtc cgataacttc cagctgtccc aggggtgggca gggattcgcc  
60

attccgatcg ggcaggcgat ggcgatecgc gccagatcc gatcgggtgg ggggtcacc  
120

accgttcata tcgggectac cgccttcctc ggcttgggtg ttgtcgaaa caacggcaac  
180

ggcgcacgag tccaacgcgt ggtcgggagc gctcgggcgg caagtctcgg catctccacc  
240

ggcgacgtga tcaccgcggt cgacggcgct ccgatcaact cggccaccgc gatggcggac  
300

gcgcttaacg ggcacatcc cgggtgacgc atctcggtsa cctggcaaac caagtcgggc  
360

ggcacgcgta cagggaacgt gacattggcc gagggacccc cggccgaatt catggtggat  
420

ttcggggcgt taccaccgga gatcaactcc gcgaggatgt acgcccggcc gggttcggcc  
480

tcgctggtgg ccgcggctca gatgtgggac agcgtggcga gtgacctgtt ttcggccgcg  
540

tcggcgtttc agtcgggtgt ctggggtctg acgggtgggt cgtggatagg ttcgtcggcg  
600

ggtctgatgg tggcggcggc ctgcgccgat gtggcgtgga tgagcgtcac cgcggggcag  
660

gccgagctga ccgcccacca ggtccgggtt gctgcccggc cctacgagac ggcgatggg  
720

ctgacggtgc ccccgccggt gatcggcgag aaccgtgctg aactgatgat tctgatagcg  
780

accaacctct tggggcaaaa caccocggcg atcgggtca acgaggcca atacggcgag  
840

atgtgggccc aagacggcgc cgcgatggtt ggtacggccg cggcgacggc gacggcgacg  
900

gcgacgttgc tgccgttcga ggaggcggc gagatgacca gcgcgggtgg gctcctcgag  
960

caggccgccc cggtcgagga ggcctccgac accgcccggg cgaaccagtt gatgaacaat  
1020

gtgccccagg cgctgcaaca gctggcccag cccacgcagg gcaccacgcc ttcttccaag  
1080

ctgggtggcc tgtggaagac ggtctcgccg catcggtcgc cgatcagcaa catgggtgctg  
1140

atggccaaca accacatgtc gatgaccaac tcgggtgtgt cgatgaccaa caccttgagc  
1200

tcgatgttga agggctttgc tccggcggcg gccgcccagg ccgtgcaaac cgcggcgcaa  
1260

aacggggctc gggcgatgag ctcgctgggc agctcgctgg gttcttcggg tctgggagg  
1320

ggggtggccc ccaacttggg tcgggcggcc tcggtcgggt cgttgctggg gccgcaggcc  
1380

tgggcccggc ccaaccaggc agtcaccccg ggggcggggg cgtgcccgt gaccagcctg  
1440

accagcggc cggaaagagg gcccgggcag atgctggggc ggctgccggg ggggcagatg  
1500

ggcgccaggg ccgggtggtg gctcagtggg gtgctgctg ttccgcccgg accctatgtg  
1560

atgccgcatt ctccggcagc cgggatata gccccgcccg ccttgctgca ggaccgggtc  
1620

gccgacttec ccgogctgcc cctcgaccgg tcgctgatgg tcgcccagt ggggccacag  
1680

gtgggtcaaca tcaacaccaa actgggtac aacaacggcg tgggcgcccg gaccggcatc  
1740

gtcatcgatc ccaacgggtg cgtgctgacc aacaaccacg tgatcgggg gccaccgac  
1800

atcaatgcgt tcagcgtcgg ctccggccaa acctacggcg tcgatgtggt cgggtatgac  
1860

cgcaccagg atgtcgggt gctgcagctg cgcgggtgcc gtggcctgcc gtcggcggcg  
1920

atcggaggcg gcgtcgggt tggtagccc gtcgctcgca tgggcaacag cggtagggcag  
1980

ggcggaacgc cccgtcgggt gctggcagg gtggctcgc tcggccaaac cgtgcaggcg  
2040

tccgattcgc tgaccggcgc cgaagagaca ttgaacgggt tgatccagtt cgatgccgcg  
2100

atccagcccg gtgatgcggg cgggcccgtc gtcaacggcc taggacaggt ggtcggstatg  
2160

aacacggccg cgtcctag  
2178

<210> 4

<211> 725

<212> PRT

<213> Artificial

<220>

<223> Descrição de Sequência Artificial: variante de proteína de tri-  
fusão

Mtb72F (Ra12-TbH9-Ra35 ou Mtb32-Mtb39)

<220>

<221> CARACT. MISC.

<222> (2)..(3)

<223> 'His His' em lugar de 'His His His His His His'

<220>

<221> CARACT. MISC.

<222> (706)..(706)

<223> 'Ala' em lugar de 'Ser'

<400> 4

Met His His Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly  
1 5 10 15

Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln  
20 25 30

Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala  
35 40 45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val  
50 55 60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr  
65 70 75 80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr  
85 90 95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser  
100 105 110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr  
115 120 125

Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu  
130 135 140

Pro Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala  
145 150 155 160

Ser Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu  
165 170 175

Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val  
180 185 190

Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser  
195 200 205

Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr  
210 215 220

Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly  
225 230 235 240

Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met  
245 250 255

Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala  
260 265 270

Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala  
275 280 285

Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu  
290 295 300

Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu





Asp Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro  
530 535 540

Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln  
545 550 555 560

Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala  
565 570 575

Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn  
580 585 590

His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser  
595 600 605

Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp  
610 615 620

Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala  
625 630 635 640

Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn  
645 650 655

Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val  
660 665 670

Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu  
675 680 685

Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly  
690 695 700

Asp Ala Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met  
705 710 715 720

<210> 5

<211> 2172

<212> DNA

<213> Artificial

<220>

<223> Descrição de Seqüência Artificial: Mtb72F-IND

<220>

<221> caract. misc.

<222> (4)..(4)

<223> deleção de marcador 'cat cac cat cac cat cac'

<400> 5

gccccagacg ccgcccgat gtttggctac gccgcggcga cggcgacggc gacggcgacg  
900

ttgctgccgt tcgaggaggc gccggagatg accagcgcgg gtgggctcct cgagcaggcc  
960

gccgcggtcg aggaggcctc cgacaccgcc gcggcgaacc agttgatgaa caatgtgccc  
1020

caggcgctgc aacagctggc ccagcccacg cagggcacca cgccttcttc caagctgggt  
1080

ggcctgtgga agaecgtctc gccgcacggt tcgocgatca gcaacatggt gtcgatggcc  
1140

aacaaccaca tgtcgatgac caactcgggt gtgtcgatga ccaacacctt gagctcgatg  
1200

ttgaagggtt ttgctccggc ggcggccgcc caggccgtgc aaaccgcggc gcaaaacggg  
1260

gtccgggcga tgagctcgct gggcagctcg ctgggttctt cgggtctggg cggtcggggtg  
1320

gccgccaact tgggtcgggc ggcctcggtc ggttcgttgt cggtgccgca ggcctgggcc  
1380

gcggccaacc aggcagtcac cccggcggcg cgggcgctgc cgtgaccag cctgaccagc  
1440

gccgcggaaa gagggcccgg gcagatgctg ggcgggctgc cggtcgggca gatgggcgcc  
1500

agggccgggt gtgggctcag tgggtgtgctg cgtgttccgc cgcgacctta tgtgatgccg  
1560

cattctccgg cagccggcga tatgccccg ccggccttgt cgcaggaccg gttcgcgac  
1620

ttcccgcgc tgcacctcga cccgtccgcg atggtcgccc aagtggggcc acaggtggtc  
1680

aacatcaaca ccaaactggg ctacaacaac gccgtgggcg ccgggaccgg catcgtcacc  
1740

gateccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat  
1800

gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc  
1860

caggatgtcg cggtcctgca gctgcgcggt gccgggtggc tgcctcggc ggcgatcggc  
1920

ggcggcgctcg cggttggtga gcccgtcgtc gcgatgggca acagcgggtgg gcagggcgga  
1980

acgccccgtg cggtgctgg cagggtggtc gcgctcggcc aaaccgtgca ggcgtcggat  
2040

tcgctgaccg gtgcccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag  
2100

cccggtgatt cgggcgggccc cgtcgtcaac ggcctaggac aggtggtcgg tatgaacacg  
2160

gccgcgtcct ga  
2172

<210> 6

<211> 723

<212> PRT

<213> Artificial

<220>

<223> Descrição de Sequência Artificial; Mtb72F-IND

<220>

<221> CARACT. MISC.

<222> (2)..(2)

<223> deleção de marcador 'His His His His His His'

<400> 6

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly  
1 5 10 15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg  
20 25 30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu  
35 40 45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg  
50 55 60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp  
65 70 75 80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met  
85 90 95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr  
100 105 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala  
115 120 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro  
130 135 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu  
145 150 155 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser  
165 170 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser  
180 185 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr  
195 200 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala  
210 215 220

Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr  
225 230 235 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu  
245 250 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn  
260 265 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe  
275 280 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe  
290 295 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala  
305 310 315 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met  
325 330 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly  
340 345 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro  
355 360 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met  
370 375 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met  
385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala  
405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly  
420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala  
435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln  
450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser  
465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly  
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val  
500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile  
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu

530	535	540													
Pro	Leu	Asp	Pro	Ser	Ala	Met	Val	Ala	Gln	Val	Gly	Pro	Gln	Val	Val
545					550					555					560
Asn	Ile	Asn	Thr	Lys	Leu	Gly	Tyr	Asn	Asn	Ala	Val	Gly	Ala	Gly	Thr
				565					570					575	
Gly	Ile	Val	Ile	Asp	Pro	Asn	Gly	Val	Val	Leu	Thr	Asn	Asn	His	Val
			580					585					590		
Ile	Ala	Gly	Ala	Thr	Asp	Ile	Asn	Ala	Phe	Ser	Val	Gly	Ser	Gly	Gln
		595					600					605			
Thr	Tyr	Gly	Val	Asp	Val	Val	Gly	Tyr	Asp	Arg	Thr	Gln	Asp	Val	Ala
	610						615				620				
Val	Leu	Gln	Leu	Arg	Gly	Ala	Gly	Gly	Leu	Pro	Ser	Ala	Ala	Ile	Gly
625					630					635					640
Gly	Gly	Val	Ala	Val	Gly	Glu	Pro	Val	Val	Ala	Met	Gly	Asn	Ser	Gly
				645					650					655	
Gly	Gln	Gly	Gly	Thr	Pro	Arg	Ala	Val	Pro	Gly	Arg	Val	Val	Ala	Leu
			660					665					670		
Gly	Gln	Thr	Val	Gln	Ala	Ser	Asp	Ser	Leu	Thr	Gly	Ala	Glu	Glu	Thr
		675					680					685			
Leu	Asn	Gly	Leu	Ile	Gln	Phe	Asp	Ala	Ala	Ile	Gln	Pro	Gly	Asp	Ser
690						695					700				
Gly	Gly	Pro	Val	Val	Asn	Gly	Leu	Gly	Gln	Val	Val	Gly	Met	Asn	Thr
705					710					715					720
Ala	Ala	Ser													

## REIVINDICAÇÕES

1. Composição farmacêutica, caracterizada pelo fato de compreender uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose e um adjuvante, para uso no tratamento ou prevenção da reativação de tuberculose em um mamífero já infectado com *Mycobacterium tuberculosis*.
2. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que o mamífero apresenta uma infecção ativa de *M. tuberculosis*.
3. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que o mamífero apresenta uma infecção latente de *M. tuberculosis*.
4. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que o mamífero é infectado com uma cepa de *M. tuberculosis* resistente a múltiplas drogas.
5. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que o mamífero foi imunizado previamente com Bacillus Calmette-Guerin (BCG).
6. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que a Mtb72f é de *Mycobacterium tuberculosis*.
7. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que a Mtb72f é um polipeptídeo compreendendo radicais de 8 a 729 da SEQ ID NO:2.
8. Composição farmacêutica de acordo com a reivindicação 7, caracterizada pelo fato de que a Mtb72f é um polipeptídeo que consiste de radicais 1 e de 8 a 729 da SEQ ID NO:2 opcionalmente com um marcador His inserido após o radical Met inicial.
9. Composição farmacêutica de acordo com a reivindicação 7, caracterizada pelo fato de que a Mtb72f é um polipeptídeo da SEQ ID NO:2.



10. Composição farmacêutica de acordo com a reivindicação 7, caracterizada pelo fato de que a Mtb72f é um polipeptídeo da SEQ ID NO:6.

5 11. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que a Mtb72f é um polipeptídeo compreendendo radicais de 4 a 725 da SEQ ID NO:4.

10 12. Composição farmacêutica de acordo com a reivindicação 11, caracterizada pelo fato de que a Mtb72f é um polipeptídeo que consiste de radicais 1 e de 4 a 725 da SEQ ID NO:4 opcionalmente com um marcador His inserido após o radical Met inicial.

13. Composição farmacêutica de acordo com a reivindicação 11, caracterizada pelo fato de que a Mtb72f é um polipeptídeo da SEQ ID NO:4.

15 14. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que o mamífero é um humano.

15. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que o adjuvante é selecionado do grupo que consiste de 3D-MPL e QS21 em uma formulação de lipossoma e 3D-MPL e QS21 e uma emulsão óleo em água.

20 16. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que é administrada juntamente com um ou mais agentes quimioterápicos eficazes para tratar uma infecção com *M. tuberculosis*.

25 17. Composição farmacêutica de acordo com a reivindicação 16, caracterizada pelo fato de que o um ou mais agentes quimioterápicos é selecionado dentre isoniazida e rifampina.

18. Composição farmacêutica de acordo com a reivindicação 16, caracterizada pelo fato de que o mamífero recebe administração inicial de um ou mais agentes quimioterápicos durante um determinado período e,

depois, recebe a administração da composição farmacêutica como definida na reivindicação 1.

19. Composição farmacêutica de acordo com a reivindicação 16, caracterizada pelo fato de que o mamífero recebe administração inicial da  
5 composição farmacêutica como definida na reivindicação 1, e, depois, recebe a administração de um ou mais agentes quimioterápicos durante um determinado período.

20. Composição farmacêutica de acordo com a reivindicação 16, caracterizada pelo fato de que a administração do um ou mais agentes  
10 quimioterápicos e da composição farmacêutica como definida na reivindicação 1 é iniciada ao mesmo tempo.

21. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que é administrada uma ou mais vezes  
subseqüentes.

22. Composição farmacêutica de acordo com a reivindicação 1, caracterizada pelo fato de que é administrada em um método de iniciar e  
15 reforçar mediante administração subseqüente um ácido nucleico que codifica uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose.

23. Composição farmacêutica, caracterizada pelo fato de  
compreender um ácido nucleico que codifica uma proteína de fusão Mtb72f  
ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium*  
do complexo de tuberculose, para uso no tratamento ou prevenção da  
reativação de tuberculose em um mamífero já infectado com *Mycobacterium*  
25 *tuberculosis*.

24. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que o ácido nucleico é SEQ ID NO: 1.

25. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que o ácido nucleico compreende nucleotídeos

de 63 a 2222 da SEQ ID NO:1.

26. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que o ácido nucleico é SEQ ID NO:3.

5 27. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que o ácido nucleico compreende nucleotídeos de 10 a 2175 da SEQ ID NO:3.

28. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que o ácido nucleico é fornecido em um vetor de adenovírus.

10 29. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que o ácido nucleico é fornecido em um vetor de célula hospedeira de *Bacillus* ou *Mycobacterium* mutante.

15 30. Composição farmacêutica de acordo com a reivindicação 23, caracterizada pelo fato de que é administrada em um método de iniciar e reforçar mediante administração subsequente de uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose.

20 31. Composição farmacêutica, caracterizada pelo fato de que compreende uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma com uma espécie de *Mycobacterium* do complexo de tuberculose e um adjuvante, para reduzir o decurso da quimioterapia contra uma infecção com *M. tuberculosis*.

25 32. Uso de uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma, caracterizado pelo fato de ser de uma espécie de *Mycobacterium* do complexo de tuberculose, para a preparação de um medicamento para o tratamento ou prevenção da reativação de tuberculose em um mamífero já infectado com *Mycobacterium tuberculosis*.

30 33. Uso de acordo com a reivindicação 32, caracterizado pelo fato de que a Mtb72f é um polipeptídeo compreendendo radicais de 4 a 725 da SEQ ID NO:4.

34. Uso de acordo com a reivindicação 32, caracterizado pelo fato de que a Mtb72f é um polipeptídeo compreendendo radicais de 4 a 725 da SEQ ID NO:6.

5 35. Uso de um ácido nucléico, caracterizado pelo fato de que codifica uma proteína de fusão Mtb72f ou um fragmento imunogênico da mesma de uma espécie de *Mycobacterium* do complexo de tuberculose, para a preparação de um medicamento para o tratamento ou prevenção da reativação de tuberculose em um mamífero já infectado com *Mycobacterium tuberculosis*.

10 36. Uso de acordo com a reivindicação 35, caracterizado pelo fato de que o ácido nucléico compreende nucleotídeos de 10 a 2175 da SEQ ID NO:3.

37. Uso de acordo com a reivindicação 35, caracterizado pelo fato de que o ácido nucléico compreende a SEQ ID NO:5.

Figura 1

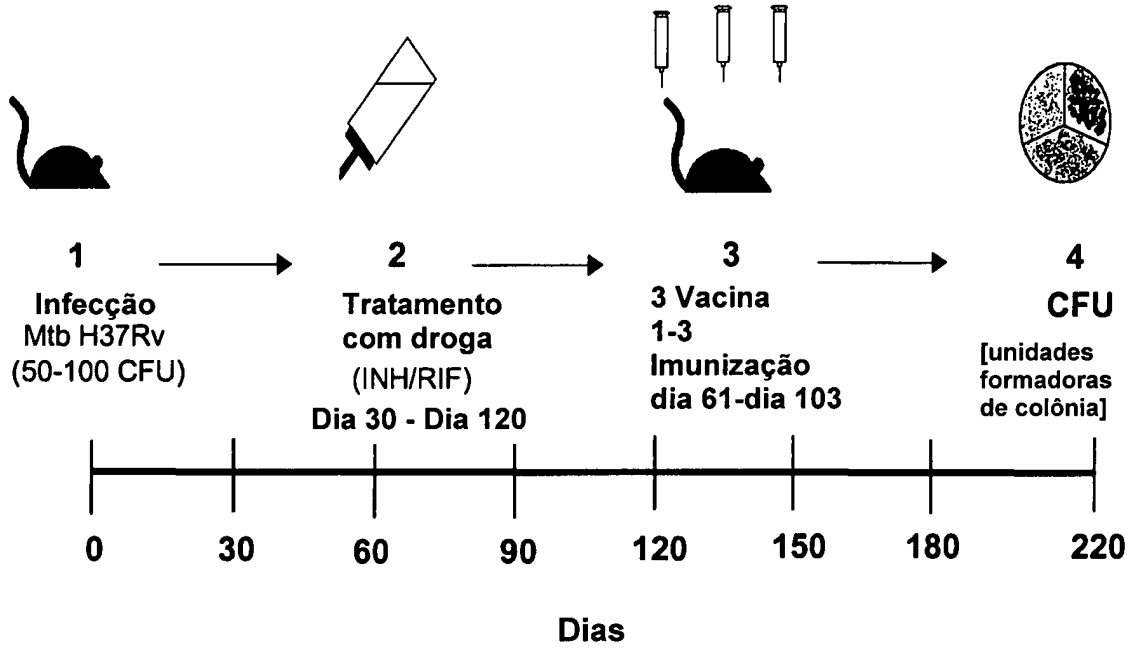


Figura 2

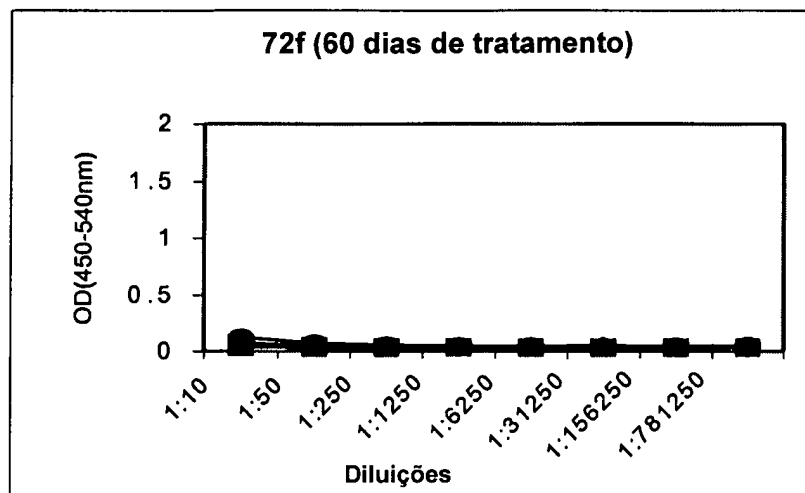
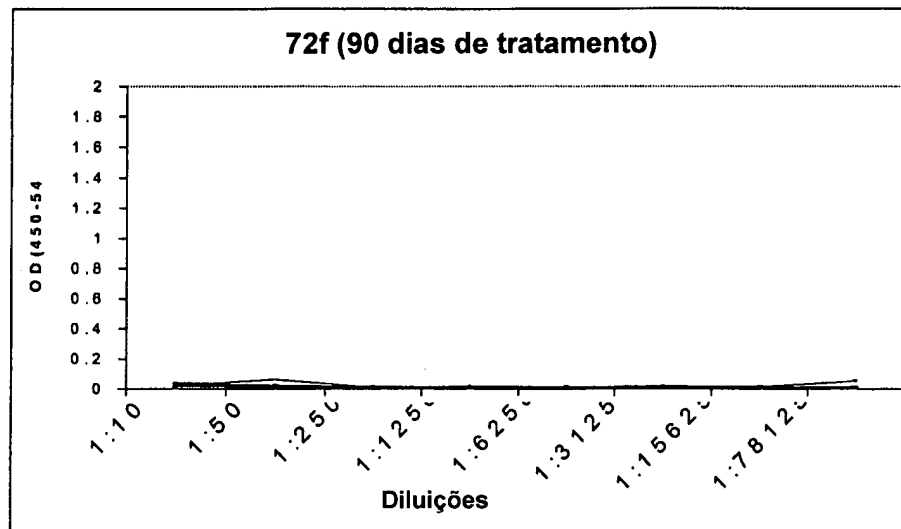
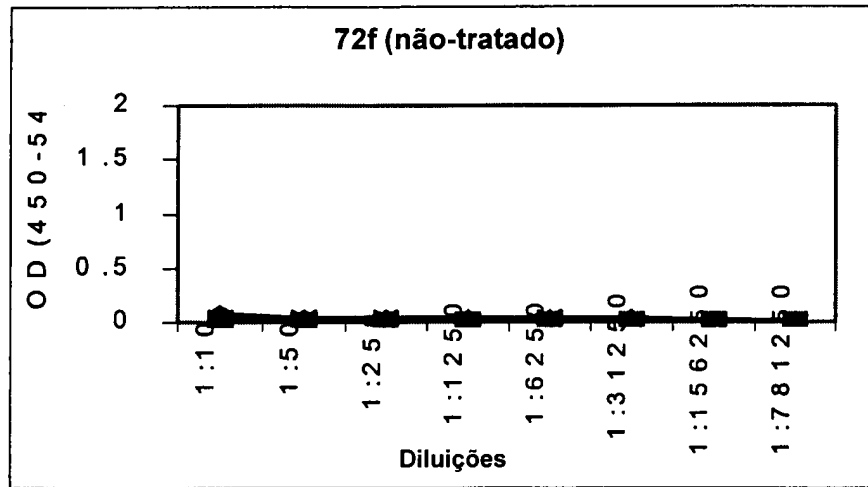


Figura 3

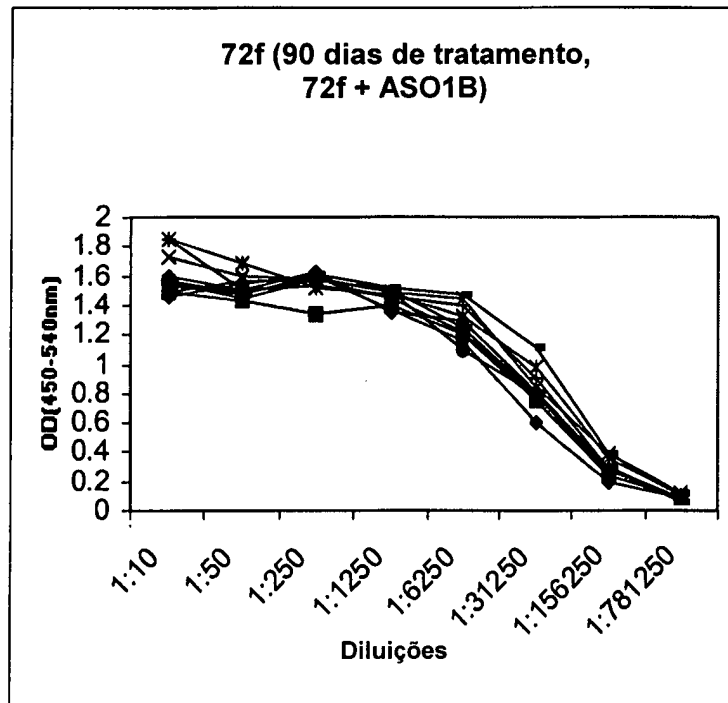
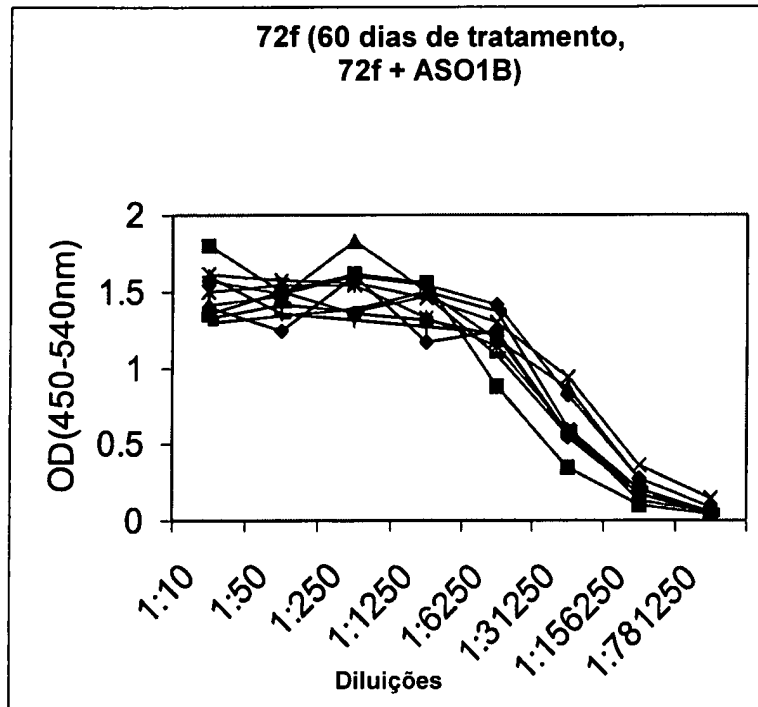


Figura 4

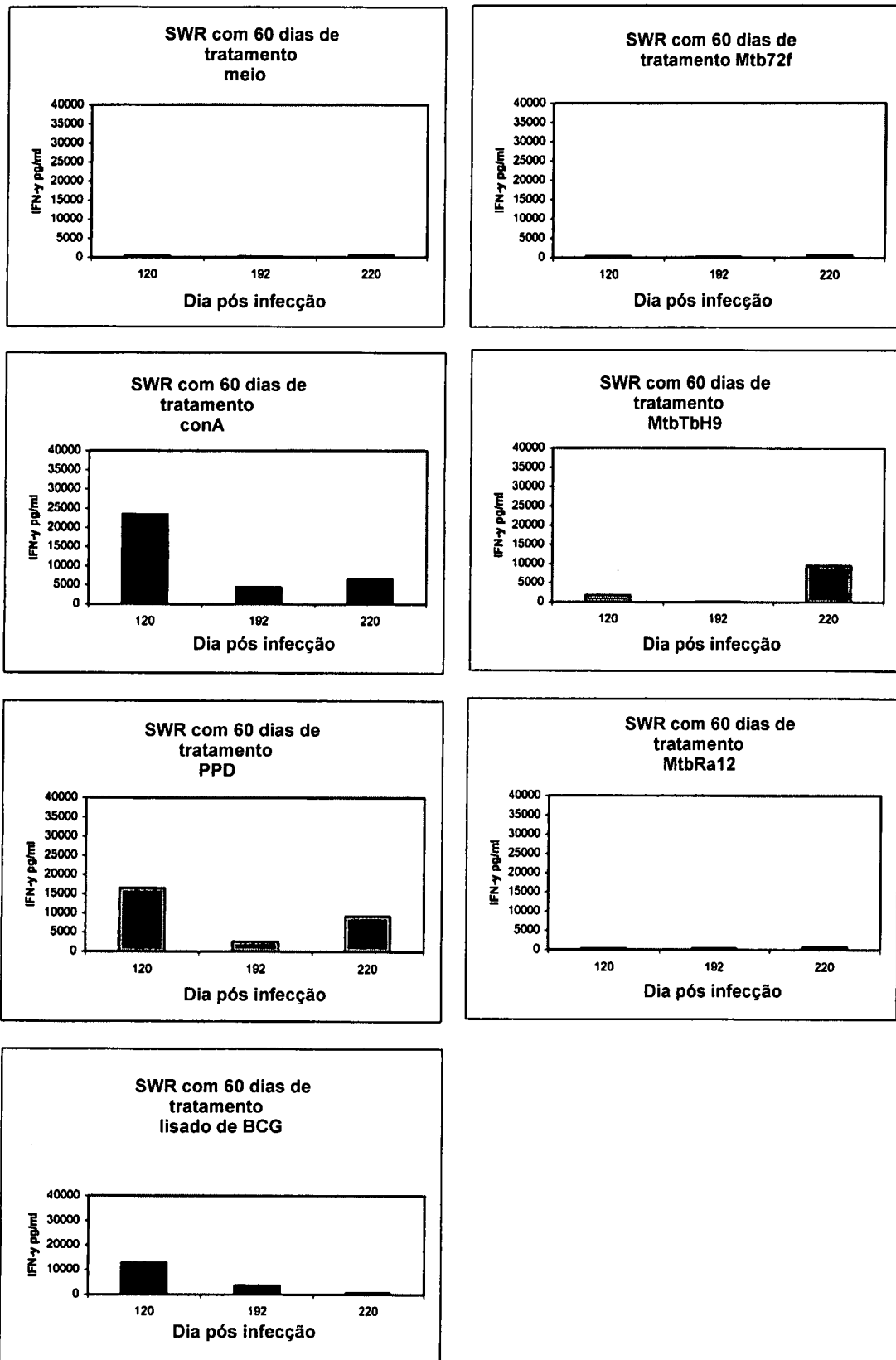




Figura 5

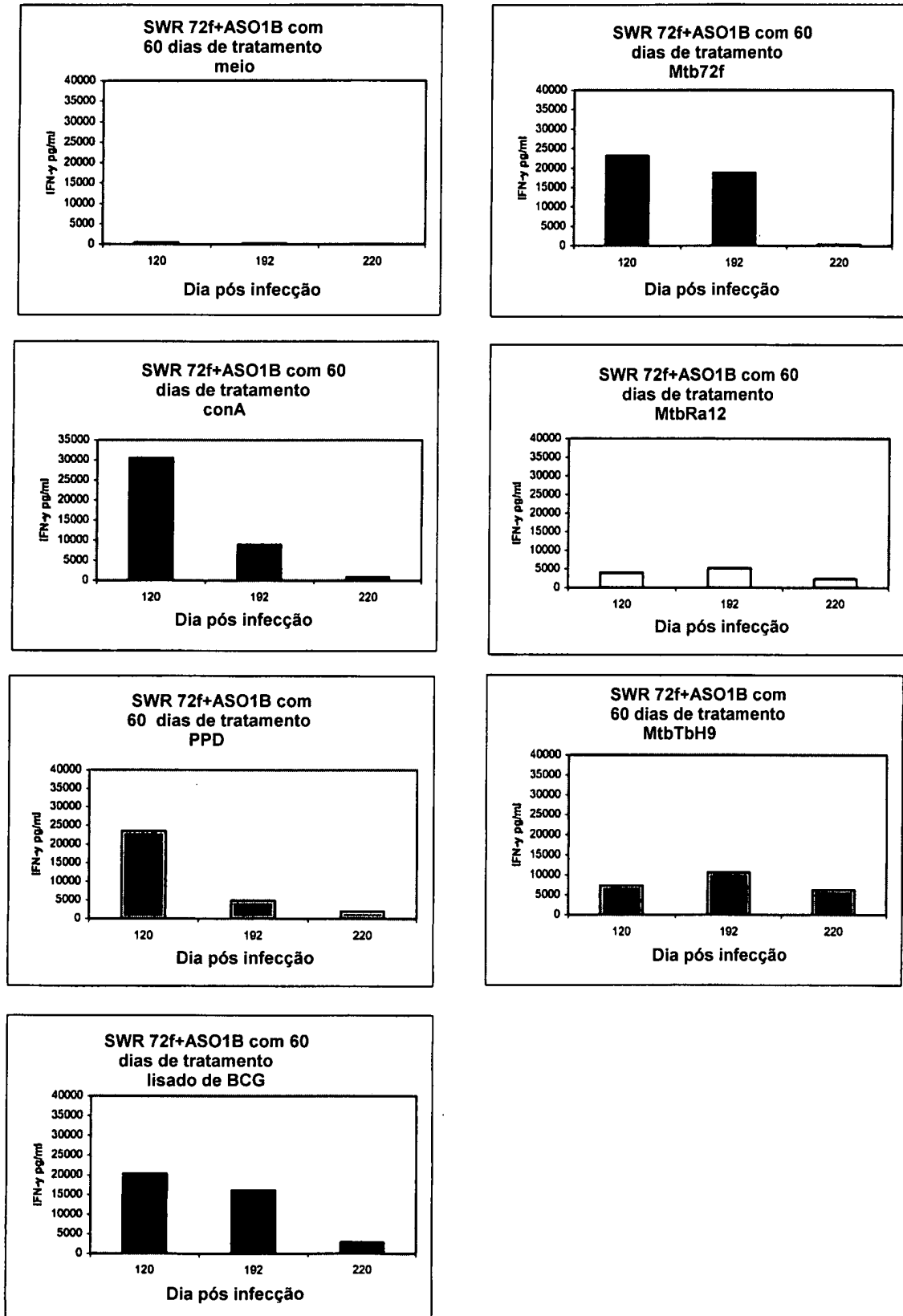
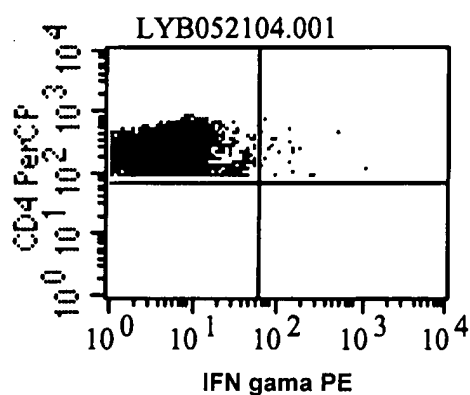
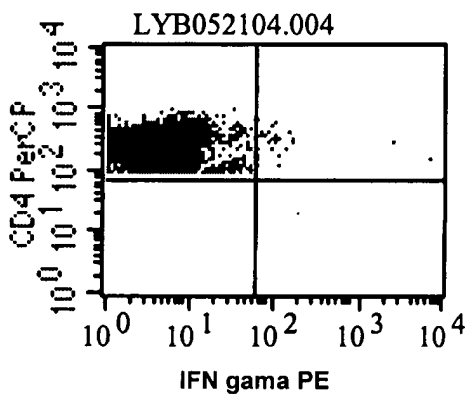


Figura 6A



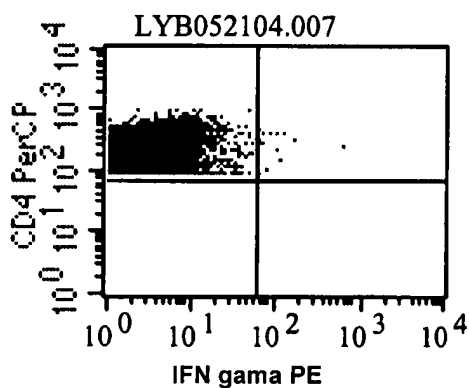
arquivo: LYB052104.001  
 ID da amostra: SWR 1 meio  
 Direcionamento: CD4  
 Eventos direcionados: 10129  
 Total de eventos: 236412

Quad	Eventos	% direcionados	% total
UL	10110	99.81	4.28
UR	19	0.19	0.01
LL	0	0.00	0.00
LR	0	0.00	0.00



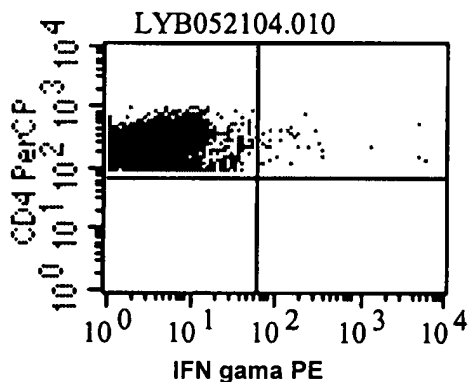
arquivo: LYB052104.004  
 ID da amostra: SWR 2 meio  
 Direcionamento: CD4  
 Eventos direcionados: 15138  
 Total de eventos: 184156

Quad	Eventos	% direcionados	% total
UL	15114	99.84	8.21
UR	24	0.16	0.01
LL	0	0.00	0.00
LR	0	0.00	0.00



arquivo: LYB052104.007  
 ID da amostra: SWR 4 meio  
 Direcionamento: CD4  
 Eventos direcionados: 19494  
 Total de eventos: 180809

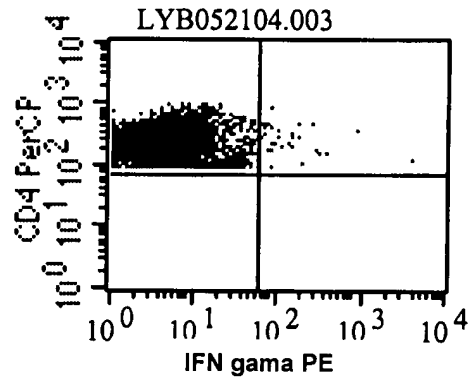
Quad	Eventos	% direcionados	% total
UL	19483	99.94	10.78
UR	11	0.06	0.01
LL	0	0.00	0.00
LR	0	0.00	0.00



arquivo: LYB052104.010  
 ID da amostra: SWR 5 meio  
 Direcionamento: CD4  
 Eventos direcionados: 14295  
 Total de eventos: 193256

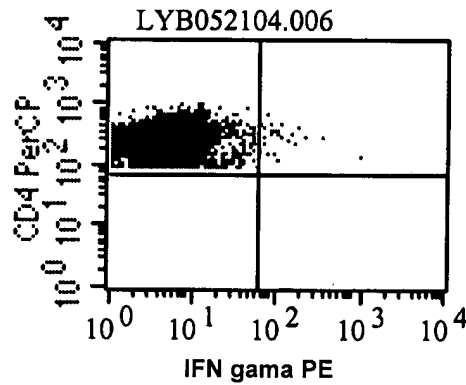
Quad	Eventos	% direcionados	% total
UL	14265	99.79	7.38
UR	30	0.21	0.02
LL	0	0.00	0.00
LR	0	0.00	0.00

Figura 6B



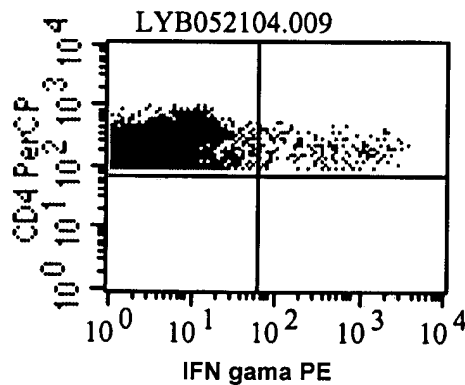
arquivo: LYB052104.003  
 ID da amostra: SWR 1 72F  
 Direcionamento: CD4  
 Eventos direcionados: 10257  
 Total de eventos: 232880

Quad	Eventos	% direcionados	% total
UL	10228	99.72	4.39
UR	29	0.28	0.01
LL	0	0.00	0.00
LR	0	0.00	0.00



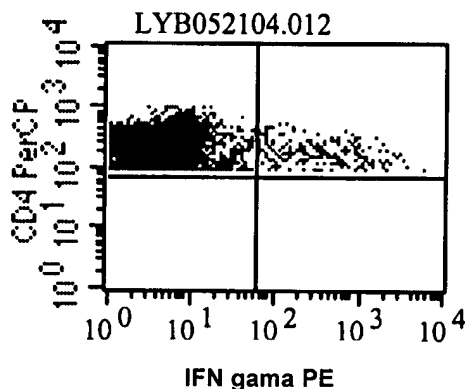
arquivo: LYB052104.006  
 ID da amostra: SWR 2 72F  
 Direcionamento: CD4  
 Eventos direcionados: 15634  
 Total de eventos: 190322

Quad	Eventos	% direcionados	% total
UL	15609	99.84	8.20
UR	25	0.16	0.01
LL	0	0.00	0.00
LR	0	0.00	0.00



arquivo: LYB052104.009  
 ID da amostra: SWR 4 72 F  
 Direcionamento: CD4  
 Eventos direcionados: 19555  
 Total de eventos: 186968

Quad	Eventos	% direcionados	% total
UL	19384	99.13	10.37
UR	171	0.87	0.09
LL	0	0.00	0.00
LR	0	0.00	0.00



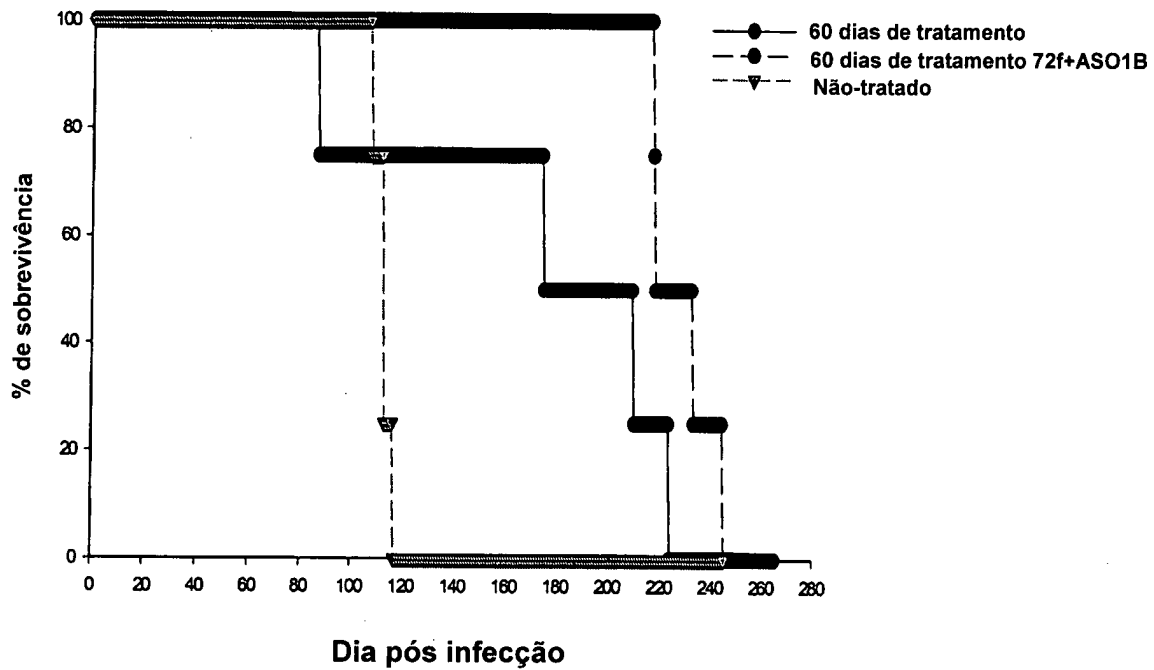
arquivo: LYB052104.012  
 ID da amostra: SWR 5 72 F  
 Direcionamento: CD4  
 Eventos direcionados: 14045  
 Total de eventos: 200028

Quad	Eventos	% direcionados	% total
UL	13849	98.60	6.92
UR	196	1.40	0.10
LL	0	0.00	0.00
LR	0	0.00	0.00

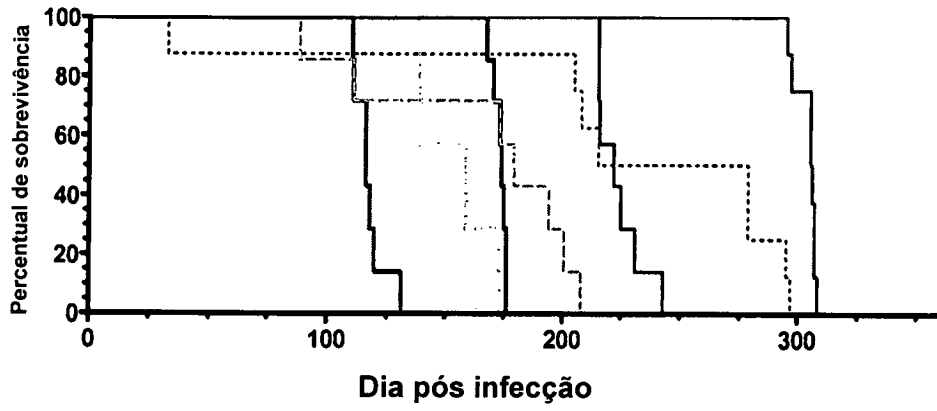
Figura 7

Sims	Grupo 1		Grupo 2		Grupo 3		Grupo 4		Grupo 5	
	Meio	72F	Meio	72F	Meio	72F	Meio	72F	Meio	72F
FNg+ % Gated	0.19	0.28	0.16	0.16	0.31	0.8	0.06	0.87	0.21	1.4
% Total	0.01	0.01	0.01	0.01	0.04	0.1	0.01	0.09	0.02	0.1
FNg+ % Gated	1.78	1.56	0.45	0.56	0.42	2.87	0.34	0.59	0.22	0.61
% Total	0.05	0.04	0.02	0.03	0.03	0.18	0.03	0.05	0.01	0.03

Figura 8



**Figura 9**



— Não-tratado      - - - 30 dias de tratamento 72F+ASO1B      ···· 60 dias de tratamento 72F+ASO1B      - · - · 90 dias de tratamento 72F+ASO1B

RESUMO

“COMPOSIÇÃO FARMACÊUTICA, USOS DE UMA PROTEÍNA DE FUSÃO Mtb72f OU UM FRAGMENTO IMUNOGÊNICO DA MESMA, E, DE UM ÁCIDO NUCLÉICO”

5                   A presente invenção refere-se a métodos de prevenir reativação de infecções ativas e latentes com *M. tuberculosis* por meio de administração de uma composição farmacêutica compreendendo um ácido nucleico codificando para uma proteína de fusão Mtb72f, ou uma proteína de fusão Mtb72f ou um fragmento imunogênico do mesmo, por exemplo, em  
10 conjunto com um adjuvante. O ácido nucleico Mtb72f ou proteína de fusão pode ser administrado com um ou mais agentes quimioterápicos eficazes contra uma infecção com *M. tuberculosis*. Os métodos também proporcionam o encurtamento do período de um regime quimioterápico contra uma infecção com *M. tuberculosis*.