

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5139618号  
(P5139618)

(45) 発行日 平成25年2月6日(2013.2.6)

(24) 登録日 平成24年11月22日(2012.11.22)

(51) Int.Cl.		F I	
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/00 Z N A A
C 0 7 K	14/35 (2006.01)	C 0 7 K	14/35
C 0 7 K	19/00 (2006.01)	C 0 7 K	19/00
C 1 2 N	1/21 (2006.01)	C 1 2 N	1/21
A 6 1 K	39/04 (2006.01)	A 6 1 K	39/04

請求項の数 29 (全 103 頁) 最終頁に続く

(21) 出願番号 特願2002-504609 (P2002-504609)  
 (86) (22) 出願日 平成13年6月20日 (2001.6.20)  
 (65) 公表番号 特表2004-526405 (P2004-526405A)  
 (43) 公表日 平成16年9月2日 (2004.9.2)  
 (86) 国際出願番号 PCT/US2001/019959  
 (87) 国際公開番号 W02001/098460  
 (87) 国際公開日 平成13年12月27日 (2001.12.27)  
 審査請求日 平成20年6月5日 (2008.6.5)  
 (31) 優先権主張番号 09/597,796  
 (32) 優先日 平成12年6月20日 (2000.6.20)  
 (33) 優先権主張国 米国 (US)  
 (31) 優先権主張番号 60/265,737  
 (32) 優先日 平成13年2月1日 (2001.2.1)  
 (33) 優先権主張国 米国 (US)

(73) 特許権者 397069329  
 コリクサ コーポレーション  
 アメリカ合衆国 19808 デラウェア  
 州, ウィルミントン, センターヴィル ロ  
 ード 2711 ザ ユナイテッド ステ  
 イツ コーポレーション, シーエスシー  
 (74) 代理人 100091096  
 弁理士 平木 祐輔  
 (74) 代理人 100096183  
 弁理士 石井 貞次  
 (74) 代理人 100118773  
 弁理士 藤田 節

最終頁に続く

(54) 【発明の名称】 Mycobacterium tuberculosis の融合タンパク質

(57) 【特許請求の範囲】

【請求項 1】

以下：

(i) M T B 3 9 抗原 (配列番号 1 2 または配列番号 1 4 )、該 M T B 3 9 抗原と少なくとも95%の同一性を有する M T B 3 9 抗原の変異体または該 M T B 3 9 抗原の50%以上からなるフラグメント；および

(ii) M T B 3 2 A 抗原 (配列番号 2 または配列番号 4 )、該 M T B 3 2 A 抗原と少なくとも95%の同一性を有する M T B 3 2 A 抗原の変異体または該 M T B 3 2 A 抗原の50%以上からなるフラグメント、を含む融合ポリペプチドであって、該 M T B 3 2 A 抗原、その該変異体またはその該フラグメントが、別のアミノ酸によって置換されている、配列番号 4 の 1 8 3 位のアミノ酸または配列番号 2 の 2 0 8 位のアミノ酸に対応するセリン残基を含むことを特徴とする、融合ポリペプチド。

【請求項 2】

M T B 3 2 A 抗原 (配列番号 2 または配列番号 4 ) の N 末端から少なくとも 1 9 5 個のアミノ酸を含むか、それと少なくとも95%の同一性を有する該 M T B 3 2 A 抗原の変異体を含む、請求項 1 に記載の融合ポリペプチド。

【請求項 3】

M T B 3 2 A 抗原 (配列番号 2 または配列番号 4 ) の N 末端から少なくとも 1 9 5 個のアミノ酸を含む、請求項 2 に記載の融合ポリペプチド。

【請求項 4】

配列番号4の183位のアミノ酸または配列番号2の208位のアミノ酸に対応するセリン残基が、アラニンによって置換されている、請求項1～3のいずれか1項に記載の融合ポリペプチド。

【請求項5】

MTB32A抗原（配列番号2または配列番号4）のC末端から少なくとも132個のアミノ酸を含むポリペプチドまたはそれと少なくとも95%の同一性を有するMTB32A抗原の変異体をさらに含む、請求項1～4のいずれか1項に記載の融合ポリペプチド。

【請求項6】

MTB32A抗原（配列番号2または配列番号4）のC末端から少なくとも132個のアミノ酸を含む、請求項5に記載の融合ポリペプチド。

10

【請求項7】

MTB39抗原（配列番号12または配列番号14）または該MTB39抗原と少なくとも95%の同一性を有するMTB39抗原の変異体を含む、請求項1～6のいずれか1項に記載の融合ポリペプチド。

【請求項8】

MTB39抗原（配列番号12または配列番号14）を含む、請求項7に記載の融合ポリペプチド。

【請求項9】

前記抗原が化学的リンカーを介して共有結合される、請求項1～8のいずれか1項に記載の融合ポリペプチド。

20

【請求項10】

前記化学的リンカーがアミノ酸リンカーである、請求項9に記載の融合ポリペプチド。

【請求項11】

前記融合ポリペプチドが、MTB72FMutSAのアミノ酸配列（配列番号18）と少なくとも95%の同一性を有する該MTB72FMutSAのアミノ酸配列（配列番号18）の変異体を含む、請求項10に記載の融合ポリペプチド。

【請求項12】

前記融合ポリペプチドが、MTB72FMutSAのアミノ酸配列（配列番号18）を含む、請求項11に記載の融合ポリペプチド。

【請求項13】

前記融合ポリペプチドが、MTB72FMutSAのアミノ酸配列（配列番号18）を有する、請求項12に記載の融合ポリペプチド。

30

【請求項14】

*Mycobacterium*種由来の少なくとも1つのさらなる異種ポリペプチドをさらに含む、請求項1～13のいずれか1項に記載の融合ポリペプチド。

【請求項15】

さらなる抗原が、MTB8.4抗原（配列番号22）、MTB9.8抗原（配列番号24）、MTB9.9抗原（配列番号27）、MTB40抗原（配列番号29）、MTB41抗原（配列番号31）、38-1（配列番号35）、TbRa3（配列番号37）、38kD（配列番号39）、DPEP（配列番号41）、TbH4（配列番号43）、DPPD（配列番号45）、Erd14、ESAT-6抗原（配列番号33）、MTB85複合体抗原もしくは - クリスタリン抗原、またはそれらの免疫原性フラグメントからなる群より選択される、請求項14に記載の融合ポリペプチド。

40

【請求項16】

アジュバントをさらに含有する、請求項1～15のいずれか1項に記載の融合ポリペプチドを含むワクチン組成物。

【請求項17】

前記アジュバントがQS21およびMPLを含む、請求項16に記載のワクチン組成物。

【請求項18】

50

前記アジュバントがQS21および3D-MPLを含む、請求項16に記載のワクチン組成物。

【請求項19】

前記アジュバントが、AS2、ENHANCYN(商標)、MPL、3D-MPL、IFA、QS21、CWS、TDM、AGP、CPG、Leif、サボニン、およびサボニン模倣物からなる群より選択される、請求項16に記載のワクチン組成物。

【請求項20】

請求項1～15のいずれか1項に記載の融合ポリペプチドをコードするポリヌクレオチド。

【請求項21】

MTB72FmutSA(配列番号18)をコードする、請求項20に記載のポリヌクレオチド。

【請求項22】

配列番号17に示される配列を含む、請求項21に記載のポリヌクレオチド。

【請求項23】

薬剤として使用するための請求項1～15のいずれか1項に記載の融合ポリペプチド。

【請求項24】

薬剤として使用するための請求項20～22のいずれか1項に記載のポリヌクレオチド。

【請求項25】

哺乳動物において免疫応答を誘発するための医薬の製造における、請求項1～15のいずれか1項に記載の融合ポリペプチドの使用。

【請求項26】

哺乳動物において免疫応答を誘発するための医薬の製造における、請求項20～22のいずれか1項に記載のポリヌクレオチドの使用。

【請求項27】

請求項1～15のいずれか1項に記載の融合ポリペプチドをその細胞表面上に発現するか、該ポリペプチドを分泌する組換え細菌。

【請求項28】

細菌がBacillus-Calmette-Guerrinである、請求項27に記載の組換え細菌。

【請求項29】

適当な宿主細胞においてポリペプチドの発現を導く組換えDNA分子の製造における請求項1～15のいずれか1項に記載のポリペプチドをコードするポリヌクレオチドの使用。

【発明の詳細な説明】

【0001】

米国特許出願第09/597、796号(2000年6月20日出願)および米国特許出願第60/265,737号(2001年2月1日)は、本明細書中で各々その全体が参考として援用される。

【0002】

本出願は、本明細書中で各々その全体が参考として援用される、米国特許出願第09/056,556号(1998年4月7日出願);米国特許出願第09/223,040号(1998年12月30日出願);米国特許出願第09/287,849号(1999年4月7日);公開PCT出願第WO99/51748号(1999年4月7日出願(PCT/US99/07717))、米国特許出願第60/158,338号(1999年10月7日出願)および米国出願第60/158,425号(1999年10月7日出願);米国出願第09/688,672号(2000年10月10日出願);ならびに公開PCT出願第WO01/24820(2000年10月10日出願(PCT/US00/28095))に関連する。

10

20

30

40

50

## 【0003】

(連邦政府によって支援された研究および開発の下で行われた発明に対する権利に関する宣言)

該当なし。

## 【0004】

(発明の背景)

本発明は、少なくとも2つの *Mycobacterium* sp. 抗原を含む融合タンパク質に関する。特に、本発明は、結核に感染した個体由来の血清の血清学的感受性を増加させる2つ以上の別々の *M. tuberculosis* 抗原を含む融合タンパク質をコードする核酸、ならびに結核感染の診断、処置、および予防におけるそれらの使用のための方法に関する。

10

## 【0005】

(発明の背景)

結核は、*M. tuberculosis* および他の *Mycobacterium* 種による感染によって引き起こされる慢性感染疾患である。結核は、発展途上国における主要な疾患であり、そして世界の先進地域においても問題が深刻化しており、毎年約800万の新たな症例および300万人の死者を出している。この感染は、かなりの期間にわたって無症候性であり得るが、この疾患は、最も一般的には、肺の急性炎症として発症し、発熱および非生産性の咳を生じる。未処置の場合、代表的には深刻な合併症および死亡を生じる。

20

## 【0006】

結核は、一般的に、長期の抗生物質治療を用いて制御され得るが、このような処置は、この疾患の伝播を予防するには十分ではない。感染個体は、無症候性であり得るが、一定の期間、伝染性であり得る。さらに、処置レジメンのコンプライアンスが重要であるが、患者の活動はモニタリングすることが困難である。一部の患者は、処置過程を完了せず、これによって、無効な処置および薬物耐性の発生を導き得る。

## 【0007】

結核の伝播を制御するために、有効なワクチン接種およびこの疾患の正確な早期診断が最も重要である。現在、生細菌を用いたワクチン接種が、防御免疫を誘導するために最も効率的な方法である。この目的のために用いられる最も一般的なミコバクテリウムは、*M. bovis* の無発病性株であるカルメット-ゲラン杆菌 (BCG) である。しかし、BCGの安全性および有効性は、論議のもとであり、いくつかの国(例えば、米国)では、この薬剤を一般にワクチン接種していない。

30

## 【0008】

結核の診断は、一般的に、皮膚試験を用いて達成される。この試験は、ツベルクリン PPD (タンパク質精製された誘導体) に対する皮内暴露を含む。抗原特異的 T 細胞応答は、注射後 48 ~ 72 時間までに注射部位に測定可能な硬化を生じる。この硬化は、ミコバクテリア抗原に対する暴露を示す。しかし、感受性および特異性が、この試験の課題であり、そして BCG でワクチン接種した個体は、感染した個体と区別することができない。

## 【0009】

マクロファージは、*Mycobacterium* 免疫の主要なエフェクターとして作用することが示されており、一方、T 細胞は、このような免疫の主なインデューサーである。*Mycobacterium* 感染に対する防御における T 細胞の本質的な役割は、ヒト免疫不全ウイルス (HIV) 感染に関連した CD4<sup>+</sup> T 細胞の枯渇に起因する、AIDS 患者における *Mycobacterium* 感染の高頻度の出現によって説明される。*Mycobacterium* 反応性 CD4<sup>+</sup> T 細胞は、 $\gamma$ -インターフェロン (IFN- $\gamma$ ) の強力なプロデューサーであることが示され、その後、マウスにおいてマクロファージの抗ミコバクテリア効果を誘発することが示された。ヒトにおける IFN- $\gamma$  の役割は、あまり明確ではないが、1, 25-ジヒドロキシ-ビタミン D<sub>3</sub> が、単独でかまたは IFN- $\gamma$  もしくは腫瘍壊死因子- $\alpha$  との組合せてかのいずれかで、ヒトマクロファージを活

40

50

性化し、M . t u b e r c u l o s i s 感染を阻害することが、研究により示された。さらに、IFN -  $\gamma$  がヒトマクロファージの 1 , 2 5 - ジヒドロキシ - ビタミン D 3 産生を刺激することが知られている。同様に、インターロイキン - 1 2 ( I L - 1 2 ) は、M . t u b e r c u l o s i s 感染に対する耐性を刺激する上で役割を果たすことが示された。M . t u b e r c u l o s i s 感染の免疫学の概要については、Chan および Kaufmann, Tuberculosis: Pathogenesis, Protection and Control ( Bloom 編, 1 9 9 4 )、ならびに Harrison ' s Principles of Internal Medicine, 第 1 巻, 1 0 0 4 - 1 0 1 4 頁および 1 0 1 9 - 1 0 2 3 頁 ( 第 1 4 版, Fauci 編, 1 9 9 8 ) を参照のこと。

10

**【 0 0 1 0 】**

従って、改善された診断試薬、ならびに結核を診断、予防、および処置するための改善された方法に対する必要性が存在する。

**【 0 0 1 1 】**

( 発明の要旨 )

従って、本発明は、少なくとも 2 つの異種抗原を含む組成物、これらの抗原を含む融合タンパク質、およびこれらの抗原をコードする核酸を提供し、ここで、これらの抗原は、結核菌由来の Mycobacterium 種および免疫無防備状態の患者において日和見感染を引き起こす他の Mycobacterium 種由来である。本発明はまた、Mycobacterium 感染の診断、処置および予防においてポリペプチドおよびポリヌクレオチドを使用する方法に関する。

20

**【 0 0 1 2 】**

1 つの局面において、本発明は、活性部位の 3 つのアミノ酸 ( ヒスチジン、アスパラギン酸、またはセリン ) のうちの 1 つ、2 つ、または 3 つが異なるアミノ酸に変異した、Ra35 ( MTB32A の N 末端部分 ) または Ra35FL ( 全長 MTB32A ) の変異版を含む組成物および融合タンパク質を提供する。1 つの実施形態において、Ra35FL において、183 位のセリンがアラニン残基に変異し、Ra35FLMutSA を生じる。1 つの実施形態において、Ra35FL をコードする DNA が、T から G への変化により変異し、配列番号 4 のアミノ酸 183 でのセリンからアラニンへの変異を生じる。別の実施形態において、本発明は、融合タンパク質 MTB72FMutSA であって、この融合タンパク質の Ra35 成分が、MTB72F 配列のアミノ酸 710 位においてセリンからアラニンへの変異を有する融合タンパク質を提供する。別の実施形態において、本発明は、融合タンパク質 MTB72F をコードする核酸であって、Ra35 成分をコードする核酸が T から G への変化により変異し、MTB72F 配列のアミノ酸 710 位でのセリンからアラニンへの変異を生じる核酸を提供する。

30

**【 0 0 1 3 】**

本発明は、少なくとも 2 種の異種 M . t u b e r c u l o s i s コード配列または抗原を含む、融合ポリヌクレオチド、融合ポリペプチドまたは組成物が、高度に抗原性であり、そして患者への投与に際して、結核血清の感受性を増大させるという本発明者らの発見に部分的に基づく。さらに、この組成物、融合ポリペプチドおよびポリヌクレオチドは、Mycobacterium に感染したかもしれない患者において診断ツールとして有用である。

40

**【 0 0 1 4 】**

1 つの局面において、本発明の組成物、融合ポリペプチドおよび核酸は、感染の診断のため、または疾患の進行をモニタリングするために、M . t u b e r c u l o s i s に対する体液性抗体または細胞性免疫を検出するために、インビトロおよびインビボでのアッセイにおいて用いられる。例えば、このポリペプチドは、インビボでの診断薬剤として皮内皮膚試験の形態で用いられ得る。このポリペプチドはまた、患者の血清を用いたインビトロ試験 ( 例えば、ELISA ) において用いられ得る。あるいは、この核酸、組成物および融合ポリペプチドを用いて、非ヒト動物において抗 M . t u b e r c u l o s i s 抗体

50

を惹起し得る。この抗体を用いて、インビボおよびインビトロで標的抗原を検出し得る。

【0015】

別の局面において、この組成物、融合ポリペプチドおよび核酸は、患者における防御免疫応答を生じるかまたは誘発するための免疫原として用いられ得る。単離または精製されたポリヌクレオチドを用いて、組換え融合ポリペプチド抗原をインビトロで産生し、次いで、これをワクチンとして投与する。あるいは、このポリヌクレオチドは、DNAワクチンとして被験体に直接投与されて、被験体において抗原の発現を引き起こし得、そしてその後、抗*M. tuberculosis*免疫応答を誘導し得る。従って、本発明の単離または精製された*M. tuberculosis*のポリペプチドおよび核酸は、*M. tuberculosis*感染の予防および/または処置における被験体への投与のための薬学的組成物として処方され得る。この融合タンパク質または抗原の免疫原性は、アジュバントならびに*Mycobacterium*または他の生物由来のさらなる融合ポリペプチド（例えば、細菌性ポリペプチド、ウイルスポリペプチド、哺乳動物ポリペプチド）を含むことによって増強され得る。さらなるポリペプチドはまた、融合ポリペプチドまたは組成物に対して連結されるか、または連結されないかのいずれかで、組成物中に含まれ得る。

10

【0016】

（特定の実施形態の説明）

本発明は、*Mycobacterium*感染の診断および処置に有用な抗原組成物および融合ポリペプチド、そのような抗原をコードするポリヌクレオチドを含む組成物、およびそれらを使用する方法に関する。本発明の抗原は、*Mycobacterium*抗原およびその免疫原性部分のポリペプチドまたは融合ポリペプチドである。より詳細には、本発明の組成物は、結核菌群の*Mycobacterium*種（例えば、*M. tuberculosis*、*M. bovis*、または*M. africanum*種）または環境的もしくは日和見感染（例えば、日和見感染）を引き起こす*Mycobacterium*種（例えば、BCG、*M. avium*、*M. intracellulare*、*M. celatum*、*M. genavense*、*M. haemophilum*、*M. kansasii*、*M. simiae*、*M. vaccae*、*M. fortuitum*、および*M. scrofulaceum*（例えば、Harrison's Principles of Internal Medicine, 第1巻, 1004 - 1014頁および1019 - 1023頁（第14版、Fauciら編、1998）を参照のこと））のうちの少なくとも2種の異種ポリペプチドを含む。本願発明者らは、驚くべきことに、少なくとも2つの異種*Mycobacterium*抗原、またはその免疫原性フラグメントを含む組成物および融合タンパク質が、高度に抗原性であることを発見した。したがって、これらの組成物、融合ポリペプチド、およびそれらをコードする核酸は、患者において防御的応答を誘発するために、および診断的適用に有用である。

20

30

【0017】

本発明の抗原は、さらに、抗原の抗原性を増強するため、または他の局面（例えば、抗原の一方の末端でのヒスチジン残基のストレッチの付加を通じたこれらの抗原の単離）においてこれらの抗原を改善するために設計された他の成分を含み得る。本発明の組成物、融合ポリペプチド、および核酸は、抗原のさらなるコピー、または*Mycobacterium* sp. 由来のさらなる異種ポリペプチド（例えば、MTB 8.4抗原、MTB 9.8抗原、MTB 9.9抗原、MTB 40抗原、MTB 41抗原、38-1、TbRa3、38kD、DPEP、TbH4、DPPD、ESAT-6抗原、MTB 85複合抗原（例えば、MTB 85b）、または -クリスタリン抗原、およびErd14）を含み得る。本発明の組成物、融合ポリペプチド、および核酸はまた、他の非*Mycobacterium*供給源由来のさらなる異種ポリペプチドを含み得る。例えば、本発明の組成物および融合タンパク質は、ポリペプチドまたはポリペプチドをコードする核酸を含み得、ここで、そのポリペプチドは、抗原、例えば、NS1、インフルエンザウイルスタンパク質、またはその免疫原性部分の発現を増強する（例えば、WO 99/40188およびWO 9

40

50

3 / 0 4 1 7 5 を参照のこと)。本発明の核酸は、選択された種（例えば、ヒト）におけるコドン優先性に基づいて操作され得る。

【 0 0 1 8 】

本発明の組成物は、裸のDNA、または組成物であり得る。例えば、ポリペプチドはまた、アジュバント（例えば、MPL、3D-MPL、IFA、ASアジュバント（例えば、AS2、AS2'、AS2"、AS4、AS6）、ENHANCYIN (Detox)、QS21、CWS、TDM、AGP、CPG、Leif、サポニン、およびサポニン模倣物、ならびにそれらの誘導体）を含み得る。さらに、本発明の組成物は、アジュバントとしてBCGまたはPvacを含み得る。

【 0 0 1 9 】

1つの実施形態において、本発明の組成物および融合タンパク質は、結核菌群のMycobacterium種由来のMTB39抗原またはその免疫原性フラグメント、および結核菌群のMycobacterium種由来のMTB32A抗原もしくはその免疫原性フラグメントからなる群から選択される少なくとも2つの抗原から構成される。

【 0 0 2 0 】

別の実施形態において、この抗原は、結核菌群のMycobacterium種由来のMTB39抗原またはその免疫原性フラグメント、および結核菌群のMycobacterium種由来のMTB32A抗原のN末端の少なくとも205アミノ酸を含むポリペプチドからなる群から選択される。

【 0 0 2 1 】

別の実施形態において、この抗原は、結核菌群のMycobacterium種由来のMTB39抗原またはその免疫原性フラグメント、結核菌群のMycobacterium種由来のMTB32A抗原のN末端の少なくとも約205アミノ酸を含むポリペプチド、および結核菌群のMycobacterium種由来のMTB32A抗原のC末端の少なくとも約132アミノ酸を含むポリペプチドからなる群から選択される。

【 0 0 2 2 】

本出願の命名法において、Ra35は、MTB32A (Ra35FL)のN末端をいい、M.tuberculosis由来のMTB32Aの少なくとも約195~205アミノ酸または別のMycobacterium種由来の対応する領域を含む。Ra12は、MTB32A (Ra32FL)のC末端をいい、M.tuberculosis由来のMTB32Aの少なくとも最後の約132アミノ酸または別のMycobacterium種由来の対応する領域を含む。

【 0 0 2 3 】

以下に、本発明の組成物および融合タンパク質において使用されるいくつかの抗原の配列を提供する。

【 0 0 2 4 】

配列番号1~4: MTB32A (Ra35FLまたはRa35成熟)、この配列はまた、米国特許出願第08/523,436号、同第08/523,435号、同第08/658,800号、同第08/659,683号、同第08/818,112号、同第09/056,556号、および同第08/818,111号ならびにWO97/09428およびWO97/09429の出願において、配列番号17 (cDNA) および配列番号79 (タンパク質) として開示される (例えば、Skeikyら、Infection and Immunity 67:3998-4007 (1999) もまた参照のこと)。用語MTB32Aはまた、三つ組活性部位の3つのアミノ酸 (His、Asp、Ser) のうちのいずれか1つ (例えば、配列番号2におけるアミノ酸208位または配列番号4におけるアミノ酸183位のセリン残基) が、別のアミノ酸 (例えば、アラニン、Ra35FL Mut SA、例えば、図6および配列番号6を参照のこと) に変化したMTB32Aアミノ酸を含む。

【 0 0 2 5 】

配列番号5および6: Ra35FL Mut SA、配列番号4のアミノ酸183位のセリン

10

20

30

40

50

残基がアラニン残基に変化した R A 3 5 F L の成熟版。

【 0 0 2 6 】

配列番号 7 および 8 : R a 3 5 ( M T B 3 2 A ( R a 3 5 F L ) の N 末端 ) は、M . t u b e r c u l o s i s 由来の M T B 3 2 A の N 末端由来の少なくとも約 1 9 5 アミノ酸を含み、このヌクレオチドおよびアミノ酸配列は、図 4 に開示される ( 配列番号 2 のアミノ酸 3 3 - 2 2 7 および配列番号 4 のアミノ酸 8 - 2 0 2 もまた参照のこと ) 。用語 R a 3 5 ( N - t e r m ) はまた、三つ組活性部位 ( すなわち、H i s、A s p、または S e r ) で、3 つのアミノ酸のうちのいずれかの 1 つが上記のように変更されている R a 3 5 アミノ酸配列を含む。

【 0 0 2 7 】

配列番号 9 および 1 0 : M T B R a 1 2 ( M T B 3 2 A ( R a 3 5 F L ) の C 末端 ) は、M . t u b e r c u l o s i s 由来の M T B 3 2 A の C 末端由来の少なくとも約 1 3 2 アミノ酸を含む ( 例えば、配列番号 2 のアミノ酸 2 2 4 - 3 5 5 および配列番号 4 のアミノ酸 1 9 9 - 3 3 0 もまた参照のこと ) 。この配列は、米国特許出願番号第 0 9 / 0 7 2 , 9 6 7 号に配列番号 4 ( D N A ) および配列番号 6 6 ( 推定されるアミノ酸配列 ) として開示される。

【 0 0 2 8 】

配列番号 1 1、1 2、1 3 および 1 4 : M T B 3 9 ( T b H 9 )、この配列は、米国特許出願番号第 0 8 / 6 5 8 , 8 0 0 号、同第 0 8 / 6 5 9 , 6 8 3 号、同第 0 8 / 8 1 8 , 1 1 2 号、同第 0 8 / 8 1 8 , 1 1 1 号ならびに W O 9 7 / 0 9 4 2 8 および W O 9 7 / 0 9 4 2 9 出願に、配列番号 1 0 6 ( c D N A 全長 ) および配列番号 1 0 7 ( タンパク質全長 ) として開示される。この配列はまた、米国特許出願番号第 0 9 / 0 5 6 , 5 5 9 号に配列番号 3 3 ( D N A ) および配列番号 9 1 ( アミノ酸 ) として開示される。

【 0 0 2 9 】

以下は、本発明のいくつかの融合タンパク質の配列を提供する。

【 0 0 3 0 】

配列番号 1 5 および 1 6 : M T B 7 2 F ( R a 1 2 - T b H 9 - R a 3 5 )、この配列は、米国特許出願第 0 9 / 2 2 3 , 0 4 0 号、同第 0 9 / 2 2 3 , 0 4 0 号ならびに P C T / U S 9 9 / 0 7 7 1 7 出願に、配列番号 1 ( D N A ) および配列番号 2 ( タンパク質 ) として開示される。用語 M T B 3 7 2 F はまた、R a 3 5 F L における三つ組活性部位 ( すなわち、H i s、A s p、または S e r ) で、3 つのアミノ酸のうちのいずれかの 1 つが上記のように変更されている M T B 7 2 F アミノ酸配列を含む ( 例えば、M T B 7 2 F M u t S A、図 5 を参照のこと ) 。

【 0 0 3 1 】

配列番号 1 7 および 1 8 : M T B 7 2 F M u t S A ( R a 1 2 - T b H 9 - R a 3 5 M u t S A )、ここで、融合タンパク質の R a 3 5 成分において、7 1 0 位のセリンがアラニンに変更されている。

【 0 0 3 2 】

配列番号 1 9 および 2 0 : T b H 9 - R a 3 5 ( M T B 5 9 F )、この配列は、米国特許出願番号第 0 9 / 2 8 7 , 8 4 9 および P C T / U S 9 9 / 0 7 7 1 7 出願に、配列番号 2 3 ( c D N A ) および配列番号 2 4 ( タンパク質 ) として開示される。

【 0 0 3 3 】

以下は、本発明の組成物および融合タンパク質において使用されるいくつかのさらなる抗原の配列を提供する。

【 0 0 3 4 】

配列番号 2 1 および 2 2 : M T B 8 . 4 ( D P V )、この配列は、米国特許出願番号第 0 8 / 6 5 8 , 8 0 0 号、同第 0 8 / 6 5 9 , 6 8 3 号、同第 0 8 / 8 1 8 , 1 1 2 号および同第 0 8 / 8 1 8 , 1 1 1 号ならびに W O 9 7 / 0 9 4 2 8 および W O 9 7 / 0 9 4 2 9 出願に、配列番号 1 0 1 ( c D N A ) および配列番号 1 0 2 ( タンパク質 ) として開示される。

10

20

30

40

50



## 【 0 0 3 5 】

配列番号 23 および 24 : M T B 9 . 8 ( M S L )、この配列は、米国特許出願番号第 08 / 8 5 9 , 3 8 1 号、同第 0 8 / 8 5 8 , 9 9 8 号、同第 0 9 / 0 7 3 , 0 0 9 号および同第 0 9 / 0 7 3 , 0 1 0 号ならびに P C T / U S 9 8 / 1 0 4 0 7 および P C T / U S 9 8 / 1 0 5 1 4 出願に、配列番号 12 ( D N A )、配列番号 109 ( 推定されるアミノ酸配列 ) および配列番号 110 ~ 124 ( ペプチド ) として開示される。

## 【 0 0 3 6 】

配列番号 25、26 および 27 : M T B 9 . 9 A ( M T I , M T I - A としても公知である )、この配列は、米国特許出願番号第 0 8 / 8 5 9 , 3 8 1 号、同第 0 8 / 8 5 8 , 9 9 8 号、同第 0 9 / 0 7 3 , 0 0 9 号および同第 0 9 / 0 7 3 , 0 1 0 号ならびに P C T / U S 9 8 / 1 0 4 0 7 および P C T / U S 9 8 / 1 0 5 1 4 出願に、配列番号 3 および配列番号 4 ( D N A ) ならびに配列番号 29 および配列番号 51 ~ 66 ( M T I についての O R F ペプチド ) として開示される。2つの他の M T I 改変体もまた存在し、M T I - B および M T I - C と呼ばれる。

10

## 【 0 0 3 7 】

配列番号 28 および 29 : M T B 4 0 ( H T C C # 1 )、この配列は、米国特許出願番号第 0 9 / 0 7 3 , 0 0 9 号、同第 0 9 / 0 7 3 , 0 1 0 号ならびに P C T / U S 9 8 / 1 0 4 0 7 および P C T / U S 9 8 / 1 0 5 1 4 出願に、配列番号 137 ( c D N A ) および 138 ( 推定されるアミノ酸配列 ) として開示される。

## 【 0 0 3 8 】

配列番号 30 および 31 : M T B 4 1 ( M T C C # 2 )、この配列は、米国特許出願番号第 0 9 / 0 7 3 , 0 0 9 号および同第 0 9 / 0 7 3 , 0 1 0 号ならびに P C T / U S 9 8 / 1 0 4 0 7 および P C T / U S 9 8 / 1 0 5 1 4 出願に、配列番号 140 ( c D N A ) および配列番号 142 ( 推定されるアミノ酸配列 ) として開示される。

20

## 【 0 0 3 9 】

配列番号 32 および 33 : E S A T - 6、この配列は、米国特許出願番号第 0 9 / 0 7 2 , 9 6 7 号に、配列番号 103 ( D N A ) および配列番号 104 ( 推定されるアミノ酸配列 ) として開示される。E S A T - 6 の配列はまた、米国特許第 5 , 9 5 5 , 0 7 7 号に開示される。

## 【 0 0 4 0 】

配列番号 34 および 35 : T b 3 8 - 1 または 3 8 - 1 ( M T b 1 1 )、この配列は、米国特許出願番号第 0 9 / 0 7 2 , 9 6 号 ; 同第 0 8 / 5 2 3 , 4 3 6 号 ; 同第 0 8 / 5 2 3 , 4 3 5 号 ; 同第 0 8 / 8 1 8 , 1 1 2 号 ; および同第 0 8 / 8 1 8 , 1 1 1 号 ; ならびに W O 9 7 / 0 9 4 2 8 および W O 9 7 / 0 9 4 2 9 出願に、配列番号 46 ( D N A ) および配列番号 88 ( 推定されるアミノ酸 ) として開示される。

30

## 【 0 0 4 1 】

配列番号 36 および 37 : T b R a 3、この配列は、W O 9 7 / 0 9 4 2 8 および W O 9 7 / 0 9 4 2 9 出願における配列番号 15 ( D N A ) および配列番号 77 ( 推定されるアミノ酸配列 ) に開示される。

## 【 0 0 4 2 】

配列番号 38 および 39 : 3 8 k D、この配列は、米国特許出願番号第 0 9 / 0 7 2 , 9 6 7 号における配列番号 154 ( D N A ) および配列番号 155 ( 推定されるアミノ酸配列 ) に開示される。3 8 k D は、N 末端にシステイン残基を伴う形態および伴わない形態の、2つの代替的形態を有する。

40

## 【 0 0 4 3 】

配列番号 40 および配列番号 41 : D P E P、この配列は、W O 9 7 / 0 9 4 2 8 および W O 9 7 / 0 9 4 2 9 公開における配列番号 52 ( D N A ) および配列番号 53 ( 推定されるアミノ酸配列 ) に開示される。

## 【 0 0 4 4 】

配列番号 42 および 43 : T b H 4、この配列は、W O 9 7 / 0 9 4 2 8 および W O 9 7

50

/ 0 9 4 2 9 公開に、配列番号 4 3 ( DNA ) および配列番号 8 1 ( 推定されるアミノ酸配列 ) として開示される。

【 0 0 4 5 】

配列番号 4 4 および 4 5 : D P P D、この配列は、U S S N 0 9 / 0 7 2 , 9 6 7 ならびに P C T / U S 9 9 / 0 3 2 6 8 および P C T / U S 9 9 / 0 3 2 6 5 出願における配列番号 2 4 0 ( DNA ) および配列番号 2 4 1 ( 推定されるアミノ酸配列 ) に開示される。D P P D の分泌形態は、P C T / U S 0 0 / 2 8 0 9 5 の図 1 2 に示される。

【 0 0 4 6 】

M T b 8 2 ( M T b 8 6 7 )、この配列は、P C T / U S 0 0 / 2 8 0 9 の図 8 ( DNA ) および 9 ( アミノ酸 ) に開示される。

10

【 0 0 4 7 】

E r d 1 4 ( M T b 1 6 )、この c D N A およびアミノ酸配列は、V e r b o n e t a l . , J . B a c t e r i o l o g y 1 7 4 : 1 3 5 2 - 1 3 5 9 ( 1 9 9 2 ) に開示される。

【 0 0 4 8 】

- クリスタリン抗原、この配列は、V e r b o n e t a l . , J . B a c t . 1 7 4 : 1 3 5 2 - 1 3 5 9 ( 1 9 9 2 ) に開示される。

【 0 0 4 9 】

8 5 複合体抗原 ( 例えば、8 5 b 抗原 )、この配列は、C o n t e n t e t a l . , I n f e c t . & I m m u n o l . 5 9 : 3 2 0 5 - 3 2 1 2 ( 1 9 9 1 ) に開示される。

20

【 0 0 5 0 】

以下は、本発明の組成物および融合タンパク質において用いられる、いくつかのさらなる融合タンパク質の配列を提供する。

【 0 0 5 1 】

配列番号 4 6 および 4 7 : D P V - M T I - M S L ( M T b 3 1 F )、この配列は、米国特許出願番号第 0 9 / 2 8 7 , 8 4 9 号 および P C T / U S 9 9 / 0 7 7 1 7 出願の配列番号 1 8 ( c D N A ) および配列番号 1 9 ( タンパク質 ) に開示される。

【 0 0 5 2 】

配列番号 4 8 および 4 9 : D P V - M T I - M S L - M T C C # 2 ( M T b 7 1 F )、この配列は、米国特許出願番号第 0 9 / 2 8 7 , 8 4 9 号 および P C T / U S 9 9 / 0 7 7 1 7 出願に、配列番号 1 5 ( 核酸 ) および配列番号 1 6 ( タンパク質 ) として開示される。

30

【 0 0 5 3 】

上記配列のそれぞれはまた、C o l e e t a l . N a t u r e 3 9 3 : 5 3 7 ( 1 9 9 8 ) に開示され、そして例えば、<http://www.sanger.ac.uk> および <http://www.pasteur.fr/mycdb/> において見出され得る。

【 0 0 5 4 】

上記配列は、米国特許出願番号第 0 8 / 5 2 3 , 4 3 5 号、同第 0 8 / 5 2 3 , 4 3 6 号、同第 0 8 / 6 5 8 , 8 0 0 号、同第 0 8 / 6 5 9 , 6 8 3 号、同第 0 8 / 8 1 8 , 1 1 1 号、同第 0 8 / 8 1 8 , 1 1 2 号、同第 0 8 / 9 4 2 , 3 4 1 号、同第 0 8 / 9 4 2 , 5 7 8 号、同第 0 8 / 8 5 8 , 9 9 8 号、同第 0 8 / 8 5 9 , 3 8 1 号、同第 0 9 / 0 5 6 , 5 5 6 号、同第 0 9 / 0 7 2 , 5 9 6 号、同第 0 9 / 0 7 2 , 9 6 7 号、同第 0 9 / 0 7 3 , 0 0 9 号、同第 0 9 / 0 7 3 , 0 1 0 号、同第 0 9 / 2 2 3 , 0 4 0 号、同第 0 9 / 2 8 7 , 8 4 9 号、同第 0 9 / 5 9 7 , 7 9 6 号 ; ならびに P C T 特許出願 P C T / U S 0 0 / 2 8 0 9 5 ; P C T / U S 9 8 / 1 0 4 0 7、P C T / U S 9 8 / 1 0 5 1 4、P C T / U S 9 9 / 0 3 2 6 5、P C T / U S 9 9 / 0 3 2 6 8、P C T / U S 9 9 / 0 7 7 1 7、W O 9 7 / 0 9 4 2 8 および W O 9 7 / 0 9 4 2 9、W O 9 8 / 1 6 6 4 5、W O 9 8 / 1 6 6 4 6 に開示され、それぞれは、本明細書中において参考として援用さ

40

50

れる。

【0055】

本明細書中に記載される抗原は、多型改変体および保存的改変体、ならびに系統間および種間 *Mycobacterium* ホモログを包含する。さらに、本明細書中に記載される抗原は、下位配列 (subsequence) または短縮配列を包含する。融合タンパク質はまた、さらなるポリペプチド (必要に応じて、*Mycobacterium* または他の供給源由来の異種ペプチド) を含み得る。これらの抗原は、例えば、下に記載されるようなリンカーペプチド配列を加えることによって、改変され得る。これらのリンカーペプチドは、融合タンパク質のそれぞれを構成する一つ以上のポリペプチドの間に挿入され得る。

10

【0056】

(定義)

「融合ポリペプチド」または「融合タンパク質」は、直接またはアミノ酸リンカーを介してのいずれかで共有結合された、少なくとも2つの異種 *Mycobacterium* sp. ポリペプチドを有するタンパク質をいう。融合タンパク質を形成するポリペプチドは、典型的には、N末端にC末端を連結するが、これらはまた、C末端にC末端を、N末端にN末端を、またはC末端にN末端を連結し得る。融合タンパク質のポリペプチドは、任意の順序であり得る。この用語はまた、融合タンパク質を構成する抗原の保存的に改変された改変体、多型改変体、対立遺伝子、変異体、下位配列、種間免疫原性フラグメントおよび種間ホモログをいう。 *Mycobacterium tuberculosis* 抗原は、Cole et al., *Nature* 393:537 (1998) に記載され、これは、*Mycobacterium tuberculosis* ゲノム全体を開示する。*Mycobacterium tuberculosis* の完全な配列はまた、<http://www.sanger.ac.uk> および [http://www.pasteur.fr/mycdb/\(MycDB\)](http://www.pasteur.fr/mycdb/(MycDB)) に見出され得る。 *M. tuberculosis* 抗原に対応する他の *Mycobacterium* 種由来の抗原は、例えば、本明細書中に開示される、配列比較アルゴリズムか、または当業者に公知の他の方法 (例えば、ハイブリダイゼーションアッセイおよび抗体結合アッセイ) を使用して同定され得る。本発明の融合タンパク質はまた、成分抗原またはそれらの免疫原性フラグメントのさらなるコピーを含み得る。

20

30

【0057】

本発明の融合タンパク質を含むポリヌクレオチド配列は、MTB39またはそれらの免疫原性フラグメントおよびMTB32Aまたはそれらの免疫原性フラグメントからなる群から選択される抗原ポリペプチドを各々コードする、少なくとも2個のヌクレオチド配列に対して、ストリンジентな条件下でハイブリダイズする。従って、この融合ポリペプチドの個々の抗原をコードするポリヌクレオチド配列としては、MTB39およびMTB32Aの保存的に改変された変異体、多型変異体、対立遺伝子、変異体、下位配列、免疫原性フラグメントおよび種間ホモログを有する融合タンパク質が挙げられる。融合タンパク質の個々のポリペプチドをコードするポリヌクレオチド配列は、任意の順序であり得る。

40

【0058】

いくつかの実施形態において、この融合タンパク質の個々のポリペプチドは、大きいほうから小さいほうへの順番 (N末端~C末端) である。大きな抗原は、約30~150kDのサイズであり、中くらいの抗原は、約10~30kDのサイズであり、そして小さな抗原は、約10kD未満のサイズである。例えば、個々のポリペプチドをコードする配列は、例えば、免疫原性フラグメント (例えば、約8~9個のアミノ酸をコードする個々のCTLエピトープ、あるいはHTLまたはB細胞エピトープ) 程度の小ささである。このフラグメントはまた、多数のエピトープを含む。この免疫原性フラグメントはまた、大部分の抗原配列 (例えば、約50%以上のMTB39およびMTB32A) を示し得る (例えば、MTB32AのN-末端部分およびC-末端部分)。好ましいMTB32Aの免疫原性フラグメントとしては、Ra12、Ra35およびRa35 MutSAが挙げられる

50

。

## 【0059】

本発明の融合ポリペプチドは、少なくとも2個の抗原ポリペプチドに対して生じる抗体に特異的に結合し、ここで、核抗原ポリペプチドは、MTB39またはそれらの免疫原性部分もしくはフラグメントおよびMTB32Aまたはそれらの免疫原性部分からなる群から選択される。この抗体は、ポリクローナルまたはモノクローナルであり得る。必要に応じて、この融合ポリペプチドは、抗原の融合結合部に対して生じる抗体に特異的に結合し、この抗体は、個々に抗原に結合しない(すなわち、抗体が融合タンパク質の一部でない場合)。必要に応じて、この融合ポリペプチドは、さらなるポリペプチド(例えば、3、4、5、6個以上のポリペプチドから約25個のポリペプチド、必要に応じて、異種ポリペプチドまたは繰返し相同ポリペプチド)を含み、少なくとも2個の異質抗原に融合される。この融合タンパク質のさらなるポリペプチドは、必要に応じて、Mycobacteriumおよび他の供給源(例えば、他の細菌、ウイルス、無脊椎動物、脊椎動物または哺乳動物供給源)から誘導される。この融合タンパク質の個々のポリペプチドは、任意の順番であり得る。本明細書中に記載されるが、この融合タンパク質はまた、他の分子(さらなるポリペプチドを含む)に連結され得る。本発明の組成物はまた、本発明の融合タンパク質に連結しないさらなるポリペプチドを含み得る。これらのさらなるポリペプチドは、異種ポリペプチドまたは同種ポリペプチドであり得る。

10

## 【0060】

用語「融合(された)」は、融合タンパク質における2つのポリペプチド間の共有結合をいう。これらのポリペプチドは、互いに直接か、またはアミノ酸リンカーを介してのいずれかで、代表的にはペプチド結合を介して結合される。必要に応じて、ペプチドは、当業者に公知の非ペプチド共有結合を介して結合され得る。

20

## 【0061】

「FL」は、全長(すなわち、野生型ポリペプチドと同じ長さであるポリペプチド)をいう。

## 【0062】

用語「それらの免疫原性フラグメント」は、細胞傷害性Tリンパ球、ヘルパーTリンパ球またはB細胞により認識されるエピトープを含むポリペプチドをいう。例えば、MTB32Aの好ましい免疫原性フラグメントは、RA35、Ra35MutSA、またはRa12である。

30

## 【0063】

用語「結核菌群のMycobacterium種」は、疾患(結核)を引き起こすと伝統的に考えられる種、ならびに免疫簡易感染性患者(例えば、AIDSを有する患者)において結核および肺疾患を引き起こすMycobacterium環境種および日和見種(例えば、M.tuberculosis、M.bovis、またはM.africanum、BCG、M.avium、M.intracellulare、M.celatum、M.genavense、M.haemophilum、M.kansasii、M.simiae、M.vaccae、M.fortuitum、およびM.scrofulaceum(例えば、Harrison's Principles of Internal Medicine、第1巻、1004~1014頁および1019~1023頁(第14版、Fauciら編、1998)を参照のこと))を含む。

40

## 【0064】

アジュバントは、抗原に対する特定の免疫応答を増強させるワクチンまたは治療組成物における成分をいう(例えば、Edelman, AIDS Res. Hum. Retroviruses 8:1409~1411(1992)を参照のこと)。アジュバントは、Th1型応答およびTh2型応答の免疫応答を誘発する。Th1型サイトカイン(例えば、IFN-、IL-2およびIL-12)は、投与された抗原に対する細胞媒介性免疫応答の誘導を支持する傾向があり、一方、Th-2型サイトカイン(例えば、IL-4、IL-5、IL-6、IL-10およびTNF-)は、体液性免疫応答の誘導を支持す

50

る傾向がある。

【0065】

「核酸」は、デオキシリボヌクレオチドまたはリボヌクレオチドおよび一本鎖または二本鎖のいずれかの形態のそれらのポリマーをいう。この用語は、公知のヌクレオチドアナログまたは改変骨格残基または連結を含む核酸を包含する。これらは、合成核酸、天然に存在する核酸、および天然に存在しない核酸である。これは、参照核酸と類似の結合特性を有し、そしてこれは、参照ヌクレオチドと類似の様式で代謝される。このようなアナログの例としては、限定なしで、ホスホロチオエート、ホスホラミデート、メチルホスホネート、キラル-メチルホスホネート、2-O-メチルリボヌクレオチド、ペプチド-核酸(PNA)が挙げられる。

10

【0066】

他に示されない限り、特定の核酸配列はまた、これらの保存的に改変された改変体(例えば、変性コドン置換体)および相補配列ならびに明確に示された配列を暗に包含する。特に、変性コドン置換体は、1以上の選択された(または全ての)コドンの3番目の位置が混合塩基および/またはデオキシイノシン残基で置換された配列を生ずることにより達成され得る(Batzerら, *Nucleic Acid Res.* 19:5081(1991); Ohtsukaら, *J. Biol. Chem.* 260:2605~2608(1985); Rossoliniら, *Mol. Cell. Probes* 8:91~98(1994))。用語核酸は、遺伝子、cDNA、mRNA、オリゴヌクレオチド、およびポリヌクレオチドと互換可能に使用される。

20

【0067】

用語「ポリペプチド」、「ペプチド」および「タンパク質」とは、アミノ酸残基のポリマーをいうために本明細書中にて交換可能に使用される。この用語は、1以上のアミノ酸残基が、対応する天然に存在するアミノ酸の人工的な化学的模倣物であるアミノ酸ポリマー、ならびに天然に存在するアミノ酸ポリマー、および天然に存在しないアミノ酸ポリマーに適用される。

【0068】

用語「アミノ酸」とは、天然に存在するアミノ酸および合成アミノ酸、ならびに天然に存在するアミノ酸と同様な様式で機能するアミノ酸アナログおよびアミノ酸模倣物をいう。天然に存在するアミノ酸は、遺伝コードによりコードされるアミノ酸、ならびに後に改変されるアミノ酸(例えば、ヒドロキシプロリン、 $\alpha$ -カルボキシグルタメート、およびO-ホスホセリン)である。アミノ酸アナログとは、天然に存在するアミノ酸と同じ化学的基本骨格(すなわち、水素、カルボキシル基、アミノ基およびR基に結合する炭素)を有する化合物(例えば、ホモセリン、ノルロイシン、メチオニンスルホキシド、メチオニンメチルスルホニウム)をいう。このようなアナログは、改変されたR基(例えば、ノルロイシン)または改変されたペプチド骨格を有するが、天然に存在するアミノ酸と同じ化学的基本骨格を保持している。アミノ酸模倣物とは、アミノ酸の一般的化学構造とは異なる構造を有するが、天然に存在するアミノ酸と同様の様式で機能する化学的化合物をいう。

30

【0069】

アミノ酸は、IUPAC-IUB生化学命名委員会により推奨される、それらの一般的に知られた3文字表記または1文字表記のいずれかにより本明細書中で言及され得る。ヌクレオチドなどは、それらの一般的に受け入れられた1文字表記により言及され得る。

40

【0070】

「保存的に改変された改変体」は、アミノ酸配列およびヌクレオチド配列の両方に適用される。特定の核酸配列に関して、保存的に改変された改変体とは、同一または本質的に同一のアミノ酸配列をコードするそれらの核酸をいうか、または核酸がアミノ酸配列をコードしない場合、本質的に同一の配列をいう。遺伝コードの縮重により、多数の機能的に同一の核酸が任意の所定のタンパク質をコードする。例えば、コドンGCA、GCC、GCGおよびGCUはすべて、アミノ酸アラニンをコードする。従って、アラニンがコドンに

50

より特定されるすべての位置で、このコドンは、コードされたポリペプチドを改変することなく、示される対応するコドンのいずれかに改変され得る。このような核酸改変体は、「サイレントなバリエーション」である。このサイレントなバリエーションは、保存的に改変されたバリエーションの1種である。ポリペプチドをコードする本明細書中のすべての核酸配列はまた、その核酸のすべてのあり得るサイレントバリエーションを記載し得る。当業者は、核酸の各コドン（そもそもメチオニンの唯一のコドンであるAUG、およびそもそもトリプトファンの唯一のコドンであるTGGを除く）は、機能的に同一の分子を生じるように改変され得ることを認識する。従って、ポリペプチドをコードする核酸の各サイレントなバリエーションは、各記載される配列に含まれる。

#### 【0071】

アミノ酸配列に関して、核酸、ペプチド、ポリペプチド、またはタンパク質の配列に対する個々の置換、欠失、付加（コードされた配列中の1つのアミノ酸または小さな割合のアミノ酸を改変、付加、または欠失する）が「保存的に改変された改変体」であり、ここでこの変更は、化学的に類似のアミノ酸とのアミノ酸置換を生じることを、当業者は認識する。機能的に類似のアミノ酸を提供する保存的置換表は、当該分野で周知である。このような保存的に改変された改変体は、さらに、そして本発明の多型改変対、種内ホモログ、および対立遺伝子を排除しない。

#### 【0072】

以下の8つの群各々は、互いに保存的置換であるアミノ酸を含む：

- 1) アラニン (A)、グリシン (G)；
- 2) アスパラギン酸 (D)、グルタミン酸 (E)；
- 3) アスパラギン (N)、グルタミン (Q)；
- 4) アルギニン (R)、リジン (K)；
- 5) イソロイシン (I)、ロイシン (L)、メチオニン (M)、バリン (V)；
- 6) フェニルアラニン (F)、チロシン (Y)、トリプトファン (W)；
- 7) セリン (S)、スレオニン (T)；および
- 8) システイン (C)、メチオニン (M)。

(例えば、Creighton, Proteins (1984)を参照のこと)。

#### 【0073】

核酸の部分を参照して用いる場合、用語「異種」は、核酸が互いに天然で同じ関係性が見出されない2以上の部分配列を含むことを示す。例えば、この核酸は、代表的には、組換え生成され、新たな機能的核酸を生じるように配置された、関連しない遺伝子に由来する2以上の配列（例えば、1つの供給源に由来するプロモーターおよび別の供給源に由来するコード領域）を有する。同様に、異種タンパク質は、このタンパク質が、互いに天然で同じ関係性が見出されない2以上の部分配列を含むことを示す（例えば、融合タンパク質）。

#### 【0074】

句「～に選択的に（または特異的に）ハイブリダイズする」とは、配列が複雑な混合物（例えば、総細胞DNAまたはRNA、あるいはライブラリーDNAまたはRNA）中に存在する場合に、ストリンジェントなハイブリダイゼーション条件下で特定のヌクレオチド配列にのみ、分子が結合、二重鎖形成、またはハイブリダイズすることをいう。

#### 【0075】

句「ストリンジェントなハイブリダイゼーション条件」とは、その条件下で、代表的には、核酸の複雑な混合物中でプローブがその標的部分配列にハイブリダイズするが、他の配列にはハイブリダイズしない条件をいう。ストリンジェントな条件は、配列依存性であり、種々の環境で異なる。より長い配列は、より高い温度で特異的にハイブリダイズする。核酸のハイブリダイゼーションに対する広範な手引きは、Tijssen, Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Probes, 「Overview of principles of hybridization a

10

20

30

40

50

nd the strategy of nucleic acid assays」(1993)に見出される。一般に、ストリンジェントな条件は、規定されたイオン強度 pH で、特定の配列についての熱融解点 ( $T_m$ ) より約 5 ~ 10 低いように選択される。 $T_m$  は、標的に相補的なプローブの 50% が平衡で (規定されたイオン強度、pH、および核酸濃度下で) 標的配列にハイブリダイズする温度である (この標的配列が過剰に存在するので、 $T_m$  でプローブの 50% が平衡で占められる)。ストリンジェントな条件は、塩濃度が約 1.0 M 未満のナトリウムイオン濃度、代表的には、pH 7.0 ~ 8.3 にて約 0.01 M ~ 1.0 M のナトリウムイオン濃度 (または他の塩) であり、温度は、短いプローブ (例えば、10 ~ 50ヌクレオチド) については、少なくとも約 30、長いプローブ (例えば、50ヌクレオチドより多い) については、少なくとも約 60 である。ストリンジェントな条件はまた、ホルムアルデヒドのような変性剤の添加により達成され得る。選択的または特異的ハイブリダイゼーションについて、ポジティブなシグナルは、少なくともバックグラウンドの 2 倍であり、必要に応じてバックグラウンドハイブリダイゼーションの 10 倍である。例示的なストリンジェントなハイブリダイゼーション条件は、以下のとおりであり得る：50% のホルムアミド、5 x SSC、および 1% の SDS、42 でのインキュベーション、または 5 x SSC、1% の SDS、65 でのインキュベーション、65 での 0.2 x SSC および 0.1% SDS での洗浄。

#### 【0076】

ストリンジェントな条件下で互いにハイブリダイズしない核酸は、これらの核酸がコードするポリペプチドが実質的に同一である場合に、なお実質的に同一である。これは、例えば、核酸のコピーが遺伝コードにより許容される最大のコドン縮重を用いて作製される場合に生じる。このような場合、この核酸は、代表的には、中程度にストリンジェントなハイブリダイゼーション条件下でハイブリダイズする。例示的な「中程度にストリンジェントなハイブリダイゼーション条件」としては、37 にて 40% のホルムアミド、1 M の NaCl、1% の SDS の緩衝液中でのハイブリダイゼーション、および 45 での 1 x SSC 中での洗浄が挙げられる。ポジティブなハイブリダイゼーションは、バックグラウンドの少なくとも 2 倍である。当業者は、代替的なハイブリダイゼーションおよび洗浄条件が類似のストリンジェンシーの条件を提供するために利用され得ることを容易に認識する。

#### 【0077】

「抗体」とは、抗原を特異的に結合し、認識する免疫グロブリン遺伝子に由来するフレーム領域を含むポリペプチド、またはそのフラグメントをいう。認識された免疫グロブリン遺伝子としては、 $\kappa$ 、 $\lambda$ 、 $\mu$ 、 $\delta$ 、および  $\mu$  の定常領域遺伝子、ならびに無数の免疫グロブリン可変領域遺伝子が挙げられる。軽鎖は、 $\kappa$ 、 $\lambda$ 、または  $\mu$  のいずれかとして分類される。重鎖は、 $\mu$ 、 $\delta$ 、 $\gamma$ 、 $\alpha$ 、または  $\epsilon$  として分類され、これは、続いて、免疫グロブリンのクラスである IgG、IgM、IgA、IgD、および IgE をそれぞれ規定する。

#### 【0078】

例示的な免疫グロブリン (抗体) 構造単位は、テトラマーを含む。各テトラマーは、2 つの同じ対のポリペプチド鎖 (各々の対は、1 つの「軽」鎖 (約 25 kDa) および 1 つの「重」鎖 (約 50 ~ 70 kDa) を有する) から構成される。各鎖の N 末端は、抗原認識を主に担う約 100 ~ 110 以上のアミノ酸の可変領域を規定する。用語、可変軽鎖 ( $V_L$ ) および可変重鎖 ( $V_H$ ) は、それぞれ、これらの軽鎖および重鎖をいう。

#### 【0079】

抗体は、例えば、インタクトな免疫グロブリンとして、または種々のペプチダーゼを用いた消化により生成された多くの十分に特徴付けられたフラグメントとして存在する。従って、例えば、ペプシンは、以下のヒンジ領域のジスルフィド結合を消化して、 $F(ab)'_2$  (それ自体、ジスルフィド結合により  $V_H - C_H1$  に結合した軽鎖である  $Fab$  のダイマーである) を生じる。 $F(ab)'_2$  は、温和な条件下で還元されて、ヒンジ領域のジスルフィド連結が壊され得る。 $Fab'$  モノマーは、本質的にはヒンジ領域の一部を有

10

20

30

40

50

するFabである(Fundamental Immunology (Paul編、第3版、1993)を参照のこと)。種々の抗体フラグメントがインタクトな抗体の消化によって規定されるが、当業者は、このようなフラグメントが、化学的または組換えDNA方法論を使用するかのいずれかにより新規合成され得ることを理解する。従って、本明細書中で用いられる場合、用語、抗体はまた、完全な抗体の改変により生成された抗体フラグメント、または組換えDNA方法論を用いて新規合成された抗体フラグメント(例えば、単鎖Fv)もしくはファージディスプレイライブラリー(例えば、McCaffertyら、Nature 348:552-554(1990)を参照のこと)を用いて同定された抗体フラグメントのいずれも含む。

【0080】

モノクローナル抗体もしくはポリクローナル抗体の調製については、当該分野で公知の任意の技術が用いられ得る(例えば、Kohler & Milstein, Nature 256:495-497(1975); Kozborら、Immunology Today 4:72(1983); Coleら、pp77-96、Monoclonal Antibodies and Cancer Therapy(1985)を参照のこと)。単鎖抗体の生成についての技術(米国特許第4,946,778号)は、本発明のポリペプチドに対する抗体を生成するために適合され得る。また、トランスジェニックマウス、または他の哺乳動物のような他の生物は、ヒト化抗体を発現するために用いられ得る。あるいは、ファージディスプレイ技術は、選択された抗原に特異的に結合する抗体およびヘテロマーFabフラグメントを同定するために用いられ得る(例えば、McCaffertyら、Nature 348:552-554(1990); Marksら、Biotechnology 10:779-783(1992)を参照のこと)。

【0081】

タンパク質またはペプチドを参照する場合に、句、抗体に「特異的に(または選択的に)結合」するまたは「特異的に(または選択的に)免疫反応性」であるとは、タンパク質または他の生物学的物質(biologic)の不均一な集団中にこのタンパク質が存在することを決定する結合反応をいう。従って、指定された免疫アッセイ条件下で、特定された抗体は、特定のタンパク質にバックグラウンドの少なくとも2倍で結合し、サンプル中に存在する他のタンパク質に有意な量で実質的に結合しない。このような条件下での抗体に対する特異的結合は、特定のタンパク質に対するその特異性に関して選択された抗体を必要とし得る。例えば、融合タンパク質に対して惹起されたポリクローナル抗体は、この融合タンパク質と特異的に免疫反応性であり、かつ融合タンパク質の個々の成分とは免疫反応性でないポリクローナル抗体のみを得るために選択され得る。この選択は、個々の抗原と交叉反応する抗体を差し引くことにより達成され得る。種々の免疫アッセイ形式が用いられて、特定のタンパク質と特異的に免疫反応性の抗体が選択され得る。例えば、固相ELISA免疫アッセイが慣用的に用いられて、タンパク質と特異的に免疫反応性の抗体が選択される(特異的免疫反応性を決定するために用いられ得る免疫アッセイ形式および条件の記載については、例えば、Harlow & Lane、Antibodies, A Laboratory Manual(1988)を参照のこと)。代表的には、特異的反応または選択的反応は、バックグラウンドシグナルまたはノイズの少なくとも2倍であり、より代表的には、バックグラウンドの10~100倍を超える。

【0082】

ポリヌクレオチドは、ネイティブな配列(すなわち、個々の抗原またはその一部をコードする内因性の配列)を含んでいてもよいし、このような配列の改変体を含んでいてもよい。ポリヌクレオチド改変対は、1以上の置換、付加、欠失および/または挿入を含み得、その結果、コードされた融合ポリペプチドの生物学的活性は、ネイティブな抗原を含む融合ポリペプチドと比較して減少されない。改変体は、ネイティブなポリペプチドまたはその一部をコードするポリヌクレオチド配列に対して、好ましくは、少なくとも約70%の同一性、より好ましくは、少なくとも約80%の同一性、もっとも好ましくは、約90%の同一性を示す。

10

20

30

40

50



## 【0083】

2以上の核酸配列またはポリペプチド配列の状況で、用語「同一」または「同一性」%とは、以下の配列比較アルゴリズムのうちの1つを用いて、または手動の整列および視覚検査により測定される、比較ウィンドウまたは指定された領域にわたり最大の一致について比較および整列した場合に、同じかあるいは同じアミノ酸残基もしくはヌクレオチドの特定の%を有する(すなわち、特定された領域にわたり70%同一、必要に応じて75%、80%、85%、90%、または95%同一)2以上の配列または部分配列をいう。次いで、このような配列は、「実質的に同一」といわれる。この規定はまた、試験配列のコンプライメント(compliant)という。必要に応じて、この同一性は、長さ少なくとも約25~約50アミノ酸または長さ少なくとも約25~約50ヌクレオチドの領域にわたり存在し、あるいは必要に応じて、長さ75~100アミノ酸または長さ約75~100ヌクレオチドの領域にわたり存在する。

10

## 【0084】

配列比較に関して、代表的には、1つの配列が参照配列として機能し、この参照配列に対して試験配列が比較される。配列比較アルゴリズムを用いた場合、試験配列および参照配列は、コンピューターに入力され、必要であれば部分配列座標が指定され、配列アルゴリズムプログラムパラメーターが指定される。デフォルトプログラムパラメーターが用いられるか、あるいは代替のパラメーターが指定され得る。次いで、配列比較アルゴリズムは、プログラムパラメーターに基づいて、参照配列に対する試験配列についての配列同一性%を計算する。

20

## 【0085】

本明細書中で用いられる場合、「比較ウィンドウ」は、25~500、通常は約50~約200、より通常は、約100~約150からなる群より選択される連続位置の数のいずれか1つのセグメントに対する参照を含む。ここで配列は、2つの配列が最適に整列されたあとに、連続位置の同じ数の参照配列に対して比較され得る。比較のための配列の整列方法は、当該分野で周知である。比較のための配列の最適な整列は、例えば、Smith & Waterman, Adv. Appl. Math. 2: 482 (1981)の局所的相同性アルゴリズムにより、Needleman & Wunsch, J. Mol. Biol. 48: 443 (1970)の相同性整列アルゴリズムにより、Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85: 2444 (1988)の類似性方法の検索により、これらのアルゴリズムのコンピューター化された実行(Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr. Madison, WI)のGAP、BESETFIT、BLASTFIT、FASTAおよびTFASTA)により、または手動の整列および視覚検査により(例えば、Current Protocols in Molecular Biology (Ausubelら編、1995、補遺)を参照のこと)行われ得る。

30

## 【0086】

有用なアルゴリズムの一例は、PILEUPである。PILEUPは、連続的な対をなす整列を用いて関連する配列の群から複数の配列整列を作製して、関連性および配列同一性%を示す。これはまた、整列を作製するために用いられるクラスター化関連性を示すツリーまたは系統樹(dendrogram)をプロットする。PILEUPは、Feng & Doolittle, J. Mol. Evol. 35: 351-360 (1987)の連続的整列方法の単純化したものを用いる。用いられる方法は、Higgins & Sharp, CABIOS 5: 151-153 (1989)により記載される方法に類似する。このプログラムは、300までの配列を整列し得、各々は、5,000のヌクレオチドまたは5,000アミノ酸の最大長である。複数の整列手順は、2つの最も類似する配列のペアをなす整列で始まり、2つの整列された配列のクラスターを生成する。次いで、このクラスターは、次に最も関連する配列または整列された配列のクラスターに整列される。配列の2つのクラスターは、2つの個々の配列のペアをなす整列を単純に拡大することにより

40

50

整列される。最終的な整列は、一連の連続的な対をなす整列により達成される。このプログラムは、配列比較の領域のための、特定の配列およびそれらのアミノ酸座標またはヌクレオチド座標を指定することにより、そしてプログラムパラメーターを指定することにより実行される。PILEUPを用いて、参照配列は、以下のパラメーターを用いて配列同一性%の関係を決定するために、他の試験配列に対して比較される：デフォルトギャップ重み付け(3.00)、デフォルトギャップ長重み付け(0.10)および重み付けエンドギャップ。PILEUPは、GCG配列分析ソフトウェアパッケージ(例えば、バージョン7.0)(Devereauxら、Nuc. Acids Res. 12:387-395(1984))から入手され得る。

#### 【0087】

パーセント配列同一性および配列類似性を決定するのに適しているアルゴリズムの別の例は、BLASTアルゴリズムおよびBLAST 2.0アルゴリズムであり、これらは、Altschulら、Nuc. Acids Res. 25:3389-3402(1977)およびAltschulらJ. Mol. Biol. 215:403-410(1990)にそれぞれ記載される。BLAST分析を実行するためのソフトウェアは、National Center for Biotechnology Information(<http://www.ncbi.nlm.nih.gov/>)を通じて公に利用可能である。このアルゴリズムは、問い合わせ配列中の長さWの短いワードを同定することによって高スコア付け配列対(HSP)を最初に同定することを含み、これは、データベース配列中で同じ長さのワードと整列した場合に、あるポジティブに値付けられた閾値スコアTに、一致するかまたは満たすかのいずれかである。Tは、隣接ワードスコア閾値として参照される(Altschulら、前出)。これらの最初の隣接ワードヒットは、これらを含むより長いHSPを見出すための検索を開始するための種として作用する。ワードヒットは、累積的整列スコアが増加し得る限り、各配列に沿って両方向に伸長される。累積スコアは、ヌクレオチド配列に関しては、パラメーターM(一致する残基対に関するリワードスコア(reward score); 常に>0)およびパラメーターN(ミスマッチ残基に関するペナルティスコア; 常に<0)を用いて算出される。アミノ酸配列に関しては、スコア付けマトリックスは、累積スコアを算出するために使用される。各方向におけるワードヒットの伸長は、以下の場合に停止される：累積整列スコアがその最大に達成された値から量X落ちる場合；累積スコアが、1つ以上の負のスコア付けの残基整列の蓄積に起因して、0またはそれより下になる場合；またはいずれかの配列の末端に達する場合。BLASTアルゴリズムパラメーターW、T、およびXは、整列の感度および速度を決定する。BLASTNプログラム(ヌクレオチド配列に関して)は、11のワード長さ(W)、10の期待値(E)、M=5、N=-4および両鎖の比較をデフォルトとして使用する。アミノ酸配列に関しては、BLASTPプログラムは、3のワード長さ、および10の期待値(E)、およびBLOSUM62スコア付けマトリックス(HenikoffおよびHenikoff, Proc. Natl Acad. Sci. U.S.A. 89:10915(1989)を参照のこと)、50の整列(B)、10の期待値(E)、M=5、N=-4、および両鎖の比較をデフォルトとして使用する。

#### 【0088】

BLASTアルゴリズムはまた、2つの配列間の統計的な類似性分析を実行する(例えば、KarlinおよびAltschul Proc. Nat'l Acad. Sci. U.S.A. 90:5873-5787(1993)を参照のこと)。BLASTアルゴリズムによって提供される類似性の1つの測定は、最小の和可能性(P(N))であり、これは、2つのヌクレオチド配列間または2つのアミノ酸配列間的一致が偶然生じる可能性の指標を提供する。例えば、参照核酸に対する試験核酸の比較におけるこの最小の和可能性が、約0.2よりも小さい場合、より好ましくは約0.01よりも小さい場合、最も好ましくは約0.001よりも小さい場合、核酸は、参照配列に類似すると考えられる。

#### 【0089】

(ポリヌクレオチド組成物)

10

20

30

40

50

本明細書中に使用される場合、用語「DNAセグメント」および「ポリヌクレオチド」は、特定種の総ゲノムDNAを含まない単離されたDNA分子をいう。従って、ポリペプチドをコードするDNAセグメントは、1つ以上のコード配列を含むが、DNAセグメントが得られる種の総ゲノムDNAから実質的に単離されているかもしくはこの総ゲノムDNAを含まずに精製されているDNAセグメントをいう。DNAセグメントおよびこのようなセグメントのより小さいフラグメント、および組換えベクター（例えば、プラスミド、コスミド、ファージミド、ファージ、ウイルスなどを含む）もまた、この用語「DNAセグメント」および「ポリヌクレオチド」内に含まれる。

【0090】

当業者に理解されるように、本発明のDNAセグメントは、ゲノム配列、ゲノム外かつプラスミドにコードされる配列、およびタンパク質、ポリペプチド、ペプチドなどを発現するかまたは発現するように適応され得るより小さな操作された遺伝子セグメントを含み得る。このようなセグメントは、天然で単離されていてもよいし、人の手によって人工的に改変されていてもよい。

【0091】

従って、用語「単離された」、「精製された」または「生物学的に純粋な」は、その天然の状態で見出される、天然に付随する成分を実質的にまたは本質的に含まない物質をいう。当然、これは、元来単離され、人の手によってこの組成物に後で付加される、他の単離されたタンパク質、遺伝子、またはコード領域を除外しないようなDNAセグメントをいう。純度および相同性は、代表的に、ポリアクリルアミドゲル電気泳動または高速液体クロマトグラフィーのような分析化学技術を使用して決定される。調製物に存在する優勢な種であるタンパク質は、実質的に精製される。単離された核酸は、その遺伝子に隣接し、その遺伝子以外のタンパク質をコードする他のオープンリーディングフレームから分離される。

【0092】

当業者によって認識されるように、ポリヌクレオチドは、一本鎖（コード鎖またはアンチセンス鎖）または二本鎖であり得、そしてDNA分子（ゲノム、cDNAまたは合成）またはRNA分子であり得る。RNA分子は、HnRNA分子（これは、イントロンを含みそして1対1の様式でDNA分子に対応する）およびmRNA分子（これは、イントロンを含まない）を含む。さらなるコード配列または非コード配列が、（必要ではないが）本発明のポリヌクレオチド内に存在し得、そしてポリヌクレオチドは、（必要ではないが）他の分子および/または支持材料に連結され得る。

【0093】

ポリヌクレオチドは、ネイティブな配列（すなわち、Mycobacterium抗原またはその部分をコードする内因性配列）を含み得るか、または改変体またはこのような配列の生物学的機能等価物もしくは抗原性機能等価物を含み得る。ポリヌクレオチド改変体は、以下にさらに記載されるように、1つ以上の置換、付加、欠失および/または挿入を含み得、その結果好ましくはコードされるポリペプチドの免疫原性は、ネイティブな腫瘍タンパク質に関して減少されていない。コードされるポリペプチドの免疫原性に対する効果は、一般的に、本明細書中に記載されるように評価され得る。用語「改変体」はまた、異種起源の相同性遺伝子を含む。

【0094】

さらなる実施形態において、本発明は、本明細書中に開示される配列の1つ以上と同一または相補的な配列の種々の長さの連続したストレッチを含む単離されたポリヌクレオチドおよびポリペプチドを提供する。例えば、本明細書中で開示される配列のうち1つ以上の、少なくとも約15、20、30、40、50、75、100、150、200、300、400、500または1000以上の連続したヌクレオチド、ならびにその間の全ての中間の長さを含むポリヌクレオチドが、本発明によって提供される。この文脈において、「中間の長さ」は、引用された値の間の任意の長さ（例えば、16、17、18、19など；21、22、23など；30、31、32など；50、51、52、53など；10

10

20

30

40

50

0、101、102、103など；150、151、152、153など；)(200～500；500～1,000などの間中の全ての整数を含む)を意味することが容易に理解される。

【0095】

本発明のポリヌクレオチド、またはそのフラグメントは、そのコード配列の長さに関係なく、他のDNA配列(例えば、プロモーター、ポリアデニル化シグナル、さらなる制限酵素部位、マルチプルクロニング部位、他のコードセグメントなど)と結合され得、その結果、その全体長さは、かなり変化し得る。従って、ほとんどの任意の長さの核酸フラグメントが利用され得、全長は、好ましくは調製の容易さおよび意図される組換えDNAプロトコルにおける使用によって制限されることが意図される。例えば、約10,000、約5000、約3000、約2,000、約1,000、約500、約200、約100、約50塩基対長など(全ての中間の長さを含む)の全体の長さの例示的なDNAセグメントは、本発明の多くの実行において有用であることが意図される。

10

【0096】

さらに、遺伝子コードの縮重の結果として、本明細書中に記載されるようなポリペプチドをコードする多くのヌクレオチド配列が存在するということが当業者に理解される。これらのポリヌクレオチドのいくつかは、任意のネイティブな遺伝子のヌクレオチド配列に対して最少の相同性を有する。それにもかかわらず、コドンの使用の差異に起因して変化するポリヌクレオチド(例えば、ヒトおよび/または霊長類のコドンの選択のために最適化されたポリヌクレオチド)が、特に本発明により意図される。さらに、本明細書中に提供されるポリヌクレオチド配列を含む遺伝子の対立遺伝子は、本発明の範囲内である。対立遺伝子は、1つ以上の変異(例えば、ヌクレオチドの欠失、付加および/または置換)の結果として変更される内因性遺伝子である。生じたmRNAおよびタンパク質は、変更された構造または機能を有し得るが、必須ではない。対立遺伝子は、標準的な技術(例えば、ハイブリダイゼーション、増幅および/またはデータベース配列比較)を使用して同定され得る。

20

【0097】

(ポリヌクレオチドの同定および特徴付け)

ポリヌクレオチドは、種々の十分に確立された技術のいずれかを使用して同定、調製および/または作製され得る。例えば、ポリヌクレオチドは、以下により詳細に記載されるように、腫瘍関連発現(すなわち、本明細書中で提供される代表的なアッセイを使用して決定された、正常な組織における発現よりも少なくとも2倍多い、腫瘍における発現)についてのcDNAのマイクロアレイのスクリーニングによって、同定され得る。このようなスクリーニングは、例えば、Synteniマイクロアレイ(Palo Alto, CA)を使用し、製造業者の説明書に従って、(そして、本質的に、Schenara, Proc. Natl. Acad. Sci. USA 93:10614-10619(1996)およびHeller, Proc. Natl. Acad. Sci. USA 94:2150-2155(1997)に記載されるように)実施され得る。あるいは、ポリヌクレオチドを、本明細書中に記載されるタンパク質を発現する細胞(例えば、M.tuberculosis細胞)から調製されるcDNAから増幅し得る。このようなポリヌクレオチドは、ポリメラーゼ連鎖反応(PCR)を介して増幅され得る。このアプローチのために、配列特異的プライマーは、本明細書中に提供される配列に基づいて設計され得、そして購入され得るか、または合成され得る。

30

40

【0098】

本発明のポリペプチドの増幅された部分は、周知技術を使用して適切なライブラリー(例えば、M.tuberculosis cDNAライブラリー)から、全長遺伝子を単離するために使用され得る。このような技術において、ライブラリー(cDNAまたはゲノム)は、増幅に適切な1以上のポリヌクレオチドプローブまたはポリヌクレオチドプライマーを使用してスクリーニングされる。好ましくは、ライブラリーは、より大きい分子を含むようにサイズ選択される。ランダムプライムライブラリーもまた、遺伝子の5'領域

50

および上流領域を同定するために好ましくあり得る。ゲノムライブラリーは、イントロンおよび伸長5'配列を得るために好ましい。

【0099】

ハイブリダイゼーション技術のために、部分配列が、周知技術を使用して標識され得る（例えば、<sup>32</sup>Pを用いるニックトランスレーションまたは末端標識によって）。次いで、細菌ライブラリーまたはバクテリオファージライブラリーは、一般に、変性させた細菌コロニーを含むフィルター（またはファージブランクを含む菌叢）にその標識プローブをハイブリダイズさせることによってスクリーニングされる（例えば、Sambrookら、Molecular Cloning: A Laboratory Manual (1989)を参照のこと）。ハイブリダイズしたコロニーまたはブランクは、選択および増殖され、そのDNAをさらなる分析のために単離する。cDNAクローンは、例えば、その部分配列由来のプライマーおよびベクター由来のプライマーを使用するPCRによって分析され、付加配列の量が決定され得る。制限地図および部分配列を作製し、1以上の重複クローンが同定され得る。次いで、完全配列は、標準的技術（これは、一連の欠失クローンの作製を含み得る）を使用して決定され得る。次いで、得られた重複配列を、単一の連続する配列にアセンブルする。全長cDNA分子を、周知技術を使用して、適切なフラグメントを連結することによって作製し得る。

10

【0100】

あるいは、部分cDNA配列から全長コード配列を得るための、多くの増幅技術が存在する。このような技術において、増幅は、一般に、PCRを介して行われる。任意の種々の市販のキットを使用して、この増幅工程を行い得る。プライマーは、例えば、当該分野で周知のソフトウェアを使用して設計され得る。プライマーは、好ましくは、22~30ヌクレオチド長であり、少なくとも50%のGC含量を有し、そして約68~72の温度で標的配列にアニールする。この増幅された領域を、上記のように配列決定し得、そして重複配列を、連続する配列にアセンブルし得る。

20

【0101】

1つのこのような増幅技術は、逆PCRである（Trigliaら、Nucl. Acids Res. 16: 8186 (1988)を参照のこと）。逆PCRは、制限酵素を使用して、遺伝子の既知の領域にフラグメントを作製する。次いで、このフラグメントを、分子内連結によって環状化し、そしてこのフラグメントを、その既知領域由来の異なるプライマーを用いるPCRのテンプレートとして使用する。代替的アプローチにおいて、部分配列に隣接する配列を、リンカー配列に対するプライマーおよび既知領域に特異的なプライマーを用いる増幅によって、回収し得る。この増幅された配列を、代表的には、同じリンカープライマーおよび既知領域に特異的な第2のプライマーを用いる2回目の増幅に供する。この手順の変形型（これは、その既知配列から反対方向への伸長を開始する2つのプライマーを使用する）が、WO96/38591に記載される。別のこのような技術は、「cDNA末端の迅速増幅（rapid amplification of cDNA ends）」すなわちRACEとして公知である。この技術は、ポリA領域またはベクター配列とハイブリダイズする、内部プライマーおよび外部プライマーの使用を包含し、既知配列の5'側および3'側の配列を同定する。さらなる技術としては、捕捉PCR（Lagerstromら、PCR Methods Applic. 1: 111-19 (1991)）およびウォーキングPCR（Parkerら、Nucl. Acids Res. 19: 3055-60 (1991)）が挙げられる。増幅を使用する他の方法もまた、全長cDNA配列を得るために使用され得る。

30

40

【0102】

特定の例において、発現配列タグ（EST）データベース（例えば、Genbankから利用可能なデータベース）に提供された配列の分析によって全長cDNA配列を得ることが可能である。重複ESTについての検索は、一般に、周知のプログラム（例えば、NCBI BLAST検索）を使用して実施され得、そしてこのようなESTを使用して、連続した全長配列を製作し得る。全長DNA配列はまた、ゲノムフラグメントの分析によ

50

て得られ得る。

【0103】

(宿主細胞におけるポリヌクレオチドの発現)

本発明の他の実施形態において、ポリヌクレオチド配列またはそのフラグメント(これは、本発明のポリペプチドをコードする)、あるいはその融合タンパク質または機能的均等物は、適切な宿主細胞におけるポリペプチドの発現を指向するために、組換えDNA分子中で使用され得る。遺伝コードの固有の縮重に起因して、実質的に同じかまたは機能的に等価なアミノ酸配列をコードする他のDNA配列が産生され得、そしてこれらの配列は、所定のポリペプチドをクローン化および発現するために使用され得る。

【0104】

当業者に理解されるように、いくつかの場合において、天然に存在しないコドン保有するポリペプチドをコードするヌクレオチド配列を産生することが有利であり得る。例えば、特定の原核生物または真核生物の宿主に好ましいコドンは、タンパク質発現の速度を増加するために選択され得るか、または所望される特性(例えば、天然に存在する配列から産生される転写物よりも長い半減期)を有する組換えRNA転写物を産生するために選択され得る。

【0105】

さらに、本発明のポリヌクレオチド配列は、種々の理由のためにポリペプチドコード配列を変化(クローニング、プロセッシング、および/または遺伝子産物の発現を改変する変化が挙げられるがこれらに限定されない)するために、当該分野で一般に公知の方法を使用して操作され得る。例えば、ランダム断片化によるDNAシャッフリング、ならびに遺伝子フラグメントおよび合成オリゴヌクレオチドのPCR再構築は、ヌクレオチド配列を操作するために使用され得る。さらに、部位特異的変異誘発は、新しい制限部位を挿入するため、グリコシル化パターンを変化するため、コドンの優先度を変化するため、スプライス改変体を生成するため、または変異を導入するためなどに使用され得る。

【0106】

本発明の別の実施形態において、天然、改変、または組換え核酸配列は、融合タンパク質をコードするために異種配列に連結され得る。例えば、ポリペプチド活性のインヒビターについてペプチドライブラリーをスクリーニングするために、市販の抗体によって認識され得るキメラタンパク質をコードすることが有用であり得る。融合タンパク質はまた、ポリペプチドコード配列と異種タンパク質配列との間に位置する切断部位を含むように操作され得、その結果、ポリペプチドは切断され得、そして異種部分から精製され得る。

【0107】

所望のポリペプチドをコードする配列は、全体的または部分的に、当該分野で周知の化学的方法を使用して合成され得る(Caruthers, M. H.ら, Nucl. Acid. Res. Symp. Ser. 215-223頁(1980), Hornら, Nucl. Acids Res. Symp. Ser. 225-232頁(1980)を参照のこと)。あるいは、このタンパク質自身は、ポリペプチドのアミノ酸配列、またはその部分を合成するために、化学的方法を使用して生成され得る。例えば、ペプチド合成は、種々の固相技術を使用して行われ得(Robergerら, Science 269:202-204(1995))、そして自動合成は、例えば、ABI 431Aペプチド合成機(Perkin Elmer, Palo Alto, CA)を使用して達成され得る。

【0108】

新しく合成されるペプチドは、分離用高速液体クロマトグラフィー(例えば、Creighton, Proteins, Structures and Molecular Principles(1983))または当該分野で利用可能な他の匹敵する技術によって実質的に精製され得る。この合成ペプチドの組成は、アミノ酸分析または配列決定(例えば、エドマン分解手順)によって確認され得る。さらに、ポリペプチドのアミノ酸配列またはその任意の部分は、直接合成の間に変化し得、そして/または、他のタンパク質に由来する配列もしくはその任意の部分を用いる化学的方法を使用して結合して、改変体ボ

10

20

30

40

50

リペプチドを生成し得る。

【0109】

所望のポリペプチドを発現するために、このポリペプチドをコードするヌクレオチド配列、またはその機能的均等物は、適切な発現ベクター（すなわち、挿入されるコード配列の転写および翻訳に必要なエレメントを含むベクター）に挿入され得る。当業者に周知の方法は、目的のポリペプチドをコードする配列、および適切な転写制御エレメントおよび翻訳制御エレメントを含む発現ベクターを構築するために使用され得る。これらの方法としては、インビトロ組換えDNA技術、合成技術、およびインビボ遺伝子組換えが挙げられる。このような技術は、Sambrookら、*Molecular Cloning, A Laboratory Manual* (1989)、およびAusubelら、*Current Protocols in Molecular Biology* (1989)に記載される。

10

【0110】

種々の発現ベクター/宿主系が、ポリヌクレオチド配列を含み、そして発現するために利用され得る。これらとしては、微生物（例えば、組換えバクテリオファージ、プラスミド、またはコスミドDNA発現ベクターで形質転換された細菌）；酵母発現ベクターで形質転換された酵母；ウイルス発現ベクター（例えば、バキュロウイルス）に感染した昆虫細胞系；ウイルス発現ベクター（例えば、カリフラワーモザイクウイルス、CaMV；タバコモザイクウイルス、TMV）または細菌発現ベクター（例えば、TiまたはpBR322プラスミド）で形質転換され植物細胞系；あるいは動物細胞系が挙げられるがこれらに限定されない。

20

【0111】

発現ベクター中に存在する「制御エレメント」または「調節配列」は、ベクターの非翻訳領域 - エンハンサー、プロモーター、5'および3'非翻訳領域 - であり、これは宿主細胞タンパク質と相互作用して、転写および翻訳を行う。このようなエレメントは、その長さおよび特異性が変化し得る。用いられるベクター系および宿主に依存して、任意数の適切な転写および翻訳エレメント（構造的プロモーターおよび誘導性プロモーターを含む）が使用され得る。例えば、細菌系におけるクローニングの場合、誘導性プロモーター（例えば、PBLUESCRIPTファージミド (Stratagene, La Jolla, Calif.) またはPSPORT1プラスミド (Gibco BRL, Gaithersburg, MD) のハイブリッドlacZプロモーターなど）が使用され得る。哺乳動物細胞系において、哺乳動物遺伝子由来または哺乳動物ウイルス由来のプロモーターが一般に好ましい。ポリペプチドをコードする配列の複数のコピーを含む細胞株の産生が必要な場合、SV40またはEBVベースのベクターは、適切な選択マーカーを用いて有利に使用され得る。

30

【0112】

細菌系において、多くの発現ベクターが、発現されるポリペプチドについて意図される使用に依存して選択され得る。例えば、大きな量が必要とされる場合（例えば、抗体の誘導のために）、容易に精製される融合タンパク質の高レベルの発現を指向するベクターが使用され得る。このようなベクターとしては、以下が挙げられるがこれらに限定されない：多機能のE. coliクローニングベクターおよび発現ベクター（例えば、BLUESCRIPT (Stratagene)）ここで、目的のポリペプチドをコードする配列が、  
- ガラクトシダーゼのアミノ末端Metおよび続く7残基についての配列とインフレームでベクターに連結され得、その結果、ハイブリッドタンパク質が産生される；pINベクター (Van Heeke & Schuster, J. Biol. Chem. 264: 5503-5509 (1989)) など。pGEXベクター (Promega, Madison, Wis.) もまた、グルタチオンS-トランスフェラーゼ (GST) との融合タンパク質として外来ポリペプチドを発現するために使用され得る。一般に、このような融合タンパク質は可溶性であり、グルタチオン-アガロースビーズへの吸着、続いて遊離グルタチオンの存在下での溶出によって溶解細胞から容易に精製され得る。このような系

40

50

で作製されたタンパク質は、ヘパリン、トロンピン、または第 X A 因子プロテアーゼ切断部位を含むように設計され得、その結果、目的のクローニングされたポリペプチドが、G S T 部分から自由に放出され得る。

【0113】

酵母 *Saccharomyces cerevisiae* において、構成的プロモーターまたは誘導性プロモーター（例えば、因子、アルコールオキシダーゼおよび P G H など）を含む多くのベクターが、使用され得る。総説については、A u s u b e l ら（前出）および G r a n t ら, *Method Enzymol.* 153:516-544 (1987) を参照のこと。

【0114】

植物発現ベクターが使用される場合において、ポリペプチドをコードする配列の発現は、多くのプロモーターのいずれかによって駆動され得る。例えば、ウイルスプロモーター（例えば、C a M V の 3 5 S および 1 9 S プロモーター）が、単独であるいは T M V 由来のリーダー配列と組合せて使用され得る（T a k a m a t s u, *EMBO J.* 6:307-311 (1987)）。あるいは、植物プロモーター（例えば、R U B I S C O の小分子サブユニット）または熱ショックプロモーターが使用され得る（C o r u z z i ら, *EMBO J.* 3:1671-1680 (1984); B r o g l i e ら, *Science* 224:838-843 (1984); および W i n t e r ら, *Results Probl. Cell Differ.* 17:85-105 (1991)）。これらの構築物は、直接的な D N A 形質転換または病原体媒介トランスフェクションによって植物細胞に導入され得る。このような技術は、多くの一般的に入手可能な総説に記載されている（例えば、H o b b s, *McGraw Hill Yearbook of Science and Technology* 191-196 (1992) を参照のこと）。

【0115】

昆虫系もまた、目的のポリペプチドを発現するために使用され得る。例えば、1つのこのような系において、*Autographa californica* 核多角体病ウイルス（A c N P V）が、*Spodoptera frugiperda* 細胞または *Trichoplusia larvae* において外来遺伝子を発現するためのベクターとして使用される。このポリペプチドをコードする配列は、ウイルス（ポリヘドリン遺伝子）の必須でない領域にクローニングされ得、そしてポリヘドリンプロモーターの制御下に置かれる。ポリペプチドコード配列の首尾よい挿入は、ポリヘドリン遺伝子を不活性にし、そしてコートタンパク質を欠く組み換えウイルスを産生する。これらの組み換えウイルスは、例えば、目的のポリペプチドが発現され得る *S. frugiperda* 細胞または *Trichoplusia larvae* を感染するために使用され得る（E n g e l h a r d ら, *Proc. Natl. Acad. Sci. U.S.A.* 91:3224-3227 (1994)）。

【0116】

哺乳動物宿主細胞において、多くのウイルスベースの発現系が、一般に利用可能である。例えば、アデノウイルスが発現ベクターとして使用される場合において、目的のポリペプチドをコードする配列は、アデノウイルス転写/翻訳複合体に連結され得、この複合体は、後期プロモーターおよび3つの部分からなるリーダー配列からなる。ウイルスゲノムの必須でない E 1 または E 3 領域における挿入は、感染された宿主細胞においてこのポリペプチドを発現し得る生存可能なウイルスを得るために使用され得る（L o g a n & S h e n k, *Proc. Natl. Acad. Sci. U.S.A.* 81:3655-3659 (1984)）。さらに、転写エンハンサー（例えば、ラウス肉腫ウイルス（R S V）エンハンサー）が、哺乳動物宿主細胞における発現を増加させるために使用され得る。

【0117】

特定の開始シグナルもまた、目的のポリペプチドをコードする配列のより効率的な翻訳を達成するために使用され得る。このようなシグナルは、A T G 開始コドンおよび隣接する

10

20

30

40

50



配列を含む。このポリペプチドをコードする配列、その開始コドン、および上流配列が、適切な発現ベクターに挿入される場合において、さらなる転写または翻訳制御シグナルが、必要とされることはないであろう。しかし、コード配列のみ、またはその一部が挿入される場合、ATG翻訳開始コドンを含む外因性翻訳制御シグナルが提供されるべきである。さらに、この開始コドンは、全挿入物の翻訳を確実にするために、正確なリーディングフレーム内にあるべきである。外因性翻訳エレメントおよび開始コドンは、種々の起源（天然および合成の両方）であり得る。発現の効率は、使用される特定の細胞系に適切なエンハンサーの包含によって増強され得、特定の細胞系は、文献（Scharfら, Results Probl. Cell Differ. 20: 125-162 (1994)）に記載されている。

10

## 【0118】

さらに、宿主細胞株は、挿入された配列の発現を調節するその能力または所望の様式で発現タンパク質を操作するその能力について、選択され得る。ポリペプチドのこのような変化は、以下を含むがこれらに限定されない：アセチル化、カルボキシル化、グリコシル化、ホスホリル化、脂質化（lipidation）およびアシル化。翻訳後プロセッシング（これは、タンパク質の「プレプロ」形態を切断する）はまた、正確な挿入、折り畳みおよび/または働きを容易にするために、使用され得る。異なる宿主細胞（例えば、CHO、HeLa、MDCK、HEK293およびWI38）（これらは、このような翻訳後活性について特異的な細胞性の機械的機構および特徴的な機構を有する）は、外来タンパク質の正確な改変およびプロセッシングを確実にするために選択され得る。

20

## 【0119】

組換えタンパク質の長期にわたる高収量産生のために、安定な発現が一般に好ましい。例えば、目的のポリヌクレオチドを安定に発現する細胞株は、発現ベクターを使用して形質転換され得、この発現ベクターは、ウイルス複製起点および/または内因性発現エレメントおよび同じまたは異なるベクター上に選択マーカー遺伝子を含み得る。ベクターの導入に続いて、細胞は、それらが選択培地にスイッチされる前に、富化培地で1~2日増殖され得る。選択マーカーの目的は、選択に対する耐性を与えることであり、そしてその存在は、導入された配列を首尾よく発現する細胞の増殖および回収を可能にする。安定に形質転換された細胞の耐性クローンは、その細胞型に適切な組織培養技術を使用して増殖され得る。

30

## 【0120】

かなりの数の選択系が、形質転換された細胞株を回収するために使用され得る。これらは、以下を含むが、これらに限定されない：単純疱疹ウイルスチミジンキナーゼ（Wiglerら, Cell 11: 223-32 (1977)）およびアデニンホスホリボシルトランスフェラーゼ（Lowyら, Cell 22: 817-23 (1990)）遺伝子（これらは、それぞれ、tk.sup.-またはaprt.sup.-細胞において使用され得る）。また、代謝拮抗物質、抗生物質または除草剤耐性が、選択の基準として使用され得る；例えば、dhfr（これは、メトトレキセートに対する耐性を与える）（Wiglerら, Proc. Natl. Acad. Sci. U.S.A. 77: 3567-70 (1980)）；npt（これは、アミノグリコシド、ネオマイシンおよびG-418に対する耐性を与える）（Colbere-Garapinら, J. Mol. Biol. 150: 1-14 (1981)）；およびalsまたはpat（これらは、それぞれ、クロスルフロンおよびホスフィノトリシンアセチルトランスフェラーゼに対する耐性を与える）（Murry, 前出）。例えば、以下のようなさらなる選択遺伝子が記述されている：trpB（これは、細胞が、トリプトファンの代わりにインドールを使用することを可能にする）またはhisD（これは、細胞がヒスチジンの代わりにヒスチノールを利用することを可能にする）（Hartman & Mulligan, Proc. Natl. Acad. Sci. U.S.A. 85: 8047-51 (1988)）。最近、可視マーカーの使用は、一般的であり、アントシアニン、-グルクロニダーゼおよびその基質GUS、ならびにルシフェラーゼおよびその基質ルシフェリンのようなマーカーが、形質転換

40

50

体を同定するのみでなく、特定のベクター系に寄与可能な一過性または安定なタンパク質発現の量を定量するためにも広く使用される (Rhodesら, *Methods Mol Biol*, 55: 121-131 (1995))。

【0121】

マーカー遺伝子発現の存在/非存在は、目的の遺伝子がまた存在することを示唆するが、その存在および発現は、確認されることを必要とし得る。例えば、ポリペプチドをコードする配列が、マーカー遺伝子配列内に挿入される場合、配列を含む組換え細胞が、マーカー遺伝子機能の非存在によって同定され得る。あるいは、マーカー遺伝子は、単一プロモーターの制御下でポリペプチドをコードする配列とタンデムで配置され得る。誘導または選択への応答におけるマーカー遺伝子の発現は、同様に、通常、タンデム遺伝子の発現を示す。

10

【0122】

あるいは、所望のポリヌクレオチド配列を含みそして発現する宿主細胞は、当業者に公知の種々の手順によって同定され得る。これらの手順は、以下を含むが、これらに限定されない: DNA-DNAまたはDNA-RNAハイブリダイゼーションおよびタンパク質バイオアッセイまたは免疫アッセイ技術(これらは、核酸またはタンパク質の検出および/または定量化のための膜、溶液またはチップに基づく技術を含む)。

【0123】

産物に特異的なポリクローナル抗体またはモノクローナル抗体のいずれかを使用する、ポリヌクレオチドをコードした産物の発現を検出および測定するための種々のプロトコルが、当該分野で公知である。例としては、酵素結合免疫ソルベントアッセイ(ELISA)、ラジオ免疫アッセイ(RIA)、および蛍光細胞分析分離法(FACS)が挙げられる。所定のポリペプチド上の2つの非干渉エピトープに反応性であるモノクローナル抗体を利用する2つの部位のモノクローナルベースの免疫アッセイがいくつかの適用に好ましくあり得るが、競合結合アッセイもまた使用され得る。これらおよび他のアッセイは、とりわけ、Hamptonら、*Serological Methods, a Laboratory Manual* (1990)およびMaddoxら、*J. Exp. Med.* 158: 11211-1216 (1983)に記載される。

20

【0124】

広範な種々の標識および結合技術が、当業者に公知であり、そして種々の核酸およびアミノ酸アッセイにおいて使用され得る。ポリヌクレオチドに関連した配列を検出するための標識されたハイブリダイゼーションプローブまたはPCRプローブを生成するための手段は、標識されたヌクレオチドを使用するオリゴ標識化、ニックトランスレーション、末端標識またはPCR増幅を含む。あるいは、配列またはその任意の部分は、mRNAプローブの産生のためにベクターにクローニングされ得る。このようなベクターは、当該分野で公知であり、市販され、そして適切なRNAポリメラーゼ(例えば、T7、T3またはSP6)および標識されたヌクレオチドの添加によって、インビトロでRNAプローブを合成するために使用され得る。これらの手順は、種々の市販のキットを使用して実施され得る。使用され得る適切なレポーター分子または標識は、放射性核種、酵素、蛍光剤、化学発光剤、または色素剤ならびに基質、共因子、インヒビター磁気粒子などを含む。

30

40

【0125】

目的のポリヌクレオチド配列で形質転換された宿主細胞は、タンパク質の発現および細胞培養物からのタンパク質の回収のために適した条件下で培養され得る。組換え細胞によって産生されたタンパク質は、使用される配列および/またはベクターに依存して、細胞内で分泌され得るか、または得られ得る。当業者に理解されるように、本発明のポリヌクレオチドを含む発現ベクターは、原核生物細胞膜または真核生物細胞膜を介するコードされたポリペプチドの分泌を指向するシグナル配列を含むように設計され得る。他の組換え構築物は、目的のポリペプチドをコードする配列を、可溶性タンパク質の精製を容易にするポリペプチドドメインをコードするヌクレオチド配列に連結するために使用され得る。このような精製が容易なドメインは、金属キレートペプチド(例えば、固定化された金属上

50

での精製を可能にするヒスチジン - トリプトファンモジュール)、プロテインAドメイン(これは、固定化された免疫グロブリン上での精製を可能にする)、およびFLAG伸長/アフィニティー精製系(ImmuneX Corp., Seattle, Wash.)において使用されるドメインを含むが、これらに限定されない。精製ドメインとコードされたポリペプチドとの間のFactor第XA因子またはエンテロキナーゼ(Invitrogen, San Diego, Calif.)に特異的な配列のような切断可能なリンカー配列の包含が、精製を容易にするために使用され得る。1つのこのような発現ベクターは、目的のポリペプチドおよびチオレドキシンまたはエンテロキナーゼ切断部位に先行する6個のヒスチジン残基をコードする核酸を含む融合タンパク質の発現を提供する。Porathら, Prot. Exp. Purif. 3: 263 - 281 (1992)に記載されるように、これらのヒスチジン残基は、IMAC上での精製(固定化された金属イオンアフィニティークロマトグラフィー)を容易にするが、エンテロキナーゼ切断部位は、融合タンパク質から所望のポリペプチドを精製するための手段を提供する。融合タンパク質を含むベクターの議論は、Krollら, DNA Cell Biol. 12: 441 - 453 (1993)に提供される。

#### 【0126】

組換え産生方法に加えて、本発明のポリペプチド、およびそのフラグメントは、固相技術を使用する直接的なペプチド合成によって産生され得る(Merrifield, J. Am. Chem. Soc. 85: 2149 - 2154 (1963))。タンパク質合成は、手動技術を使用してまたは自動化によって実施され得る。自動化合成は、例えば、Applied Biosystems 431A Peptide Synthesizer(Perkin Elmer)を使用して達成され得る。あるいは、種々のフラグメントが、個別に化学合成され、そして化学的な方法を使用して合されて、全長分子を生成し得る。

#### 【0127】

(インビボにおけるポリヌクレオチド送達技術)

さらなる実施形態において、1つ以上の本発明のポリヌクレオチドを含む遺伝子構築物が、インビボにおいて細胞に導入される。これは、様々なアプローチまたは周知のアプローチ(これらのいくつかは、例示の目的で以下に示される)のいずれかを使用して達成され得る。

#### 【0128】

(1. アデノウイルス)

1つ以上の核酸配列のインビボ送達のための好ましい方法の一つは、アデノウイルス発現ベクターの使用を包含する。「アデノウイルス発現ベクター」は、(a)この構築物のパッケージングを支持し、そして(b)センス配向またはアンチセンス配向でそこにクローン化されたポリヌクレオチドを発現するのに十分なアデノウイルス配列を含む構築物を含むことを意味する。もちろん、アンチセンス構築物の場合において、発現は、遺伝子産物が合成されることを必要としない。

#### 【0129】

発現ベクターは、遺伝子操作した形態のアデノウイルスを含む。アデノウイルス(36 kb、線状の二本鎖DNAウイルス)の遺伝子構築の知識により、7 kbまでの異種配列を有する大きな断片のアデノウイルスのDNA置換が可能となる(Grunhaus & Horwitz, 1992)。レトロウイルスとは対照的に、宿主細胞のアデノウイルス感染は、染色体組込みを生じない。なぜなら、アデノウイルスDNAは、強力な遺伝子毒性を伴わないエピソーム様式で複製され得るからである。また、アデノウイルスは、構造的に安定であり、そしてゲノム再配列は広範な複製後には検出されなかった。アデノウイルスは、それらの細胞周期段階に関係なく、実質的に全ての上皮細胞に感染し得る。これまでは、アデノウイルス感染は、ヒトにおける急性呼吸器疾患のような軽度の疾患のみに関連するようである。

#### 【0130】

アデノウイルスは、その中間のサイズのゲノム、操作の容易さ、高い力価、広い標的細胞範囲、および高い感染力のため、遺伝子伝達ベクターとして使用するために特に適切である。ウイルスゲノムの両端は、100～200塩基対の逆方向反復（ITR）を含み、これはウイルスDNAの複製およびパッケージングのために必要なシスエレメントである。ゲノムの早期（E）および後期（L）領域は、ウイルスDNA複製の開始によって分裂される異なる転写単位を含む。E1領域（E1AおよびE1B）は、ウイルスゲノムおよびわずかな細胞遺伝子の転写の制御の原因であるタンパク質をコードする。E2領域（E2AおよびE2B）の発現は、ウイルスDNA複製のためのタンパク質の合成を導く。これらのタンパク質は、DNA複製、後期遺伝子発現、および宿主遮断に参与する（Renan, 1990）。後期遺伝子の産物（大多数のウイルスカプシドタンパク質を含む）は、  
10  
主な後期プロモーター（MLP）により起こる単一の一次転写物の有意なプロセシングの後だけに発現される。MLP（16.8 m.u.に位置する）は、感染の後期の間に特に効率的であり、そしてこのプロモーターから生じる全てのmRNAは、それらを翻訳のために好ましいmRNAにする5'-三分裂リーダー（5'-tripartite leader）（TPL）配列を有する。

#### 【0131】

現存のシステムにおいて、組換えアデノウイルスは、シャトルベクターとプロウイルスベクターとの間の相同性組換えによって生成される。2つのプロウイルスベクターの間の可能な組換えに起因して、野生型アデノウイルスはこのプロセスから生成され得る。従って、  
20  
ウイルスの単一のクローンを、個々のプラークから単離し、そしてそのゲノム構造を試験することが重要である。

#### 【0132】

現存のアデノウイルスベクター（これは、複製欠失性である）の生成および増殖は、特有のヘルパー細胞株（これは、293と命名され、Ad5 DNAフラグメントによってヒト胚腎細胞から形質転換され、そしてE1タンパク質を構成的に発現する）に依存する（Grahamら、1977）。E3領域は、アデノウイルスゲノムには必要ではないため（JonesおよびShenk, 1978）、現存のアデノウイルスベクターは、293細胞の助けをかりて、異種DNAをE1、D3、またはその両方の領域のいずれかに運搬する（GrahamらおよびPrevec, 1991）。天然では、アデノウイルスは、  
30  
野生型ゲノムの約105%をパッケージングし（Ghosh-Choudhuryら、1987）、DNAの約2kBを越えるDNAの容量を提供し得る。E1およびE3領域において複製可能な約5.5kBのDNAと組み合わせると、現存のアデノウイルスベクターの最大容量は、7.5kBより下であるか、またはベクターの全長の約15%である。80%を越えるアデノウイルスゲノムは、ベクター骨格中にとどまり、そしてベクターにより運ばれた細胞傷害性の供給源である。また、E1欠失ウイルスの複製欠失は不完全である。例えば、ウイルス遺伝子発現の漏出は、高い感染効率（MOI）で、現在利用可能なベクターを用いて観察された（Mulligan, 1993）。

#### 【0133】

ヘルパー細胞株は、ヒト細胞（例えば、ヒト肺腎細胞、筋肉細胞、造血細胞または他のヒト胚間葉細胞もしくは上皮細胞）由来であり得る。あるいは、ヘルパー細胞は、ヒトアデ  
40  
ノウイルスに対して許容性の他の哺乳動物種の細胞由来であり得る。このような細胞としては、Vero細胞、または他のサル胚間葉細胞もしくは上皮細胞が挙げられる。上記のように、現存の好ましいヘルパー細胞株は293である。

#### 【0134】

最近、Racherら（1995）は、293細胞を培養しそしてアデノウイルスを増殖するための改良された方法を開示した。1つの様式において、天然の細胞凝集物を、100～200mlの培養液を含む1リットルのシリコン処理したスピナーフラスコ（Tehne, Cambridge, UK）に個々の細胞を播種することによって増殖した。40rpmで撹拌した後、細胞生存度を、トリパンブルーを用いて推定した。別の様式において、Fibra-Cellマイクロキャリア（Bibby Sterlin, Stone  
50

、UK) (5 g / l) を以下のように使用した。5 ml の培養液に再懸濁した細胞播種物を、250 ml のエルレンマイヤーフラスコ中のキャリア (50 ml) に添加し、そして時折攪拌しながら、1 ~ 4 時間静置した。この培養物を、次いで、50 ml の新鮮培地に入れ、そして攪拌を開始した。ウイルス産生のために、細胞を約 80 % の集密度まで増殖し、その後この培地を (最終容量の 25 % または約まで) 置換し、そしてアデノウイルスを 0.05 の MOI で添加した。培地を一晩静止し、その後この容量を 100 % まで増やし、そしてさらに 72 時間の振盪を開始した。

#### 【0135】

アデノウイルスベクターが複製欠失であるか、または少なくとも条件付きで欠失であるという要件以外、アデノウイルスベクターの特徴は本発明の首尾良い実施に対して重要ではないと考えられる。アデノウイルスは、42 個の異なる既知の血清型またはサブグループ A ~ F のいずれかを有し得る。サブグループ C の 5 型アデノウイルスは、本発明において使用するための条件的複製欠失アデノウイルスベクターを得るための好ましい出発物質である。なぜなら、5 型アデノウイルスは、多くの生化学的情報および遺伝的情報が知られているヒトアデノウイルスであり、そしてこれはアデノウイルスをベクターとして使用するほとんどの構築のために歴史的に使用されているからである。

10

#### 【0136】

上記のように、本発明に従う典型的なベクターは、複製欠失であり、そしてアデノウイルス E1 領域を有さない。従って、E1 コード配列が除去される位置において目的の遺伝子をコードするポリヌクレオチドを導入することが最も簡便である。しかし、アデノウイルス配列内のこの構築物の挿入位置は、本発明に重要ではない。目的の遺伝子をコードするポリヌクレオチドはまた、Karlssonら (1986) に記載されるように、E3 置換ベクター中の欠失 E3 領域の変わりに挿入され得るか、またはヘルパー細胞株またはヘルパーウイルスが E4 欠失を補完する E4 領域に挿入される。

20

#### 【0137】

アデノウイルスは、増殖および操作し易く、そしてインビトロおよびインビボにおける広範な宿主範囲を示す。この群のウイルスは、高力価 (例えば、1 ml あたり  $10^9 \sim 10^{11}$  プラーク形成単位) で得られ得、そしてこれらは高い感染力である。アデノウイルスのライフサイクルは、宿主細胞ゲノムへの組み込みを必要としない。アデノウイルスベクターによって送達される異種遺伝子はエピソーム性であり、従って、宿主細胞に対して低い遺伝子毒性を有する。副作用は野生型アデノウイルスを用いるワクチン接種の研究において報告されず (Couchら、1963; Topら、1971)、それらの安全性およびインビボ遺伝子伝達ベクターとしての治療的可能性を示す。

30

#### 【0138】

アデノウイルスベクターは、真核生物遺伝子発現 (Levreroら、1991; Gomeix-Foixら、1992)、およびワクチン開発 (GrunhausおよびHorwitz、1992; GrahamおよびPrevec、1992) において使用されてきた。最近、動物研究により、組換えアデノウイルスは、遺伝子治療のために使用され得ることが示唆された (G Stratford - PerricaudetおよびPerricaudet、1991; Stratford - Perricaudetら、1990; Richら、1993)。組換えアデノウイルスを異なる組織に投与する研究は、気管注入 (Rosenfeldら、1991; Rosenfeldら、1992)、筋肉注射 (Ragotら、1993)、末梢静脈注射 (HerzおよびGerard、1993)、および脳への定位接種 (Le Gal La Salleら、1993) を含む。

40

#### 【0139】

##### (2. レトロウイルス)

レトロウイルスは、逆転写のプロセスによって、感染細胞においてそれらの RNA を二本鎖 DNA に変換する能力によって特徴付けられる群の二本鎖 RNA ウイルスである (Coffin、1990)。次いで、得られた DNA は、細胞の染色体にプロウイルスとして安定に組み込み、そしてウイルスタンパク質の合成を指向する。この組み込みは、レシピエン

50

トの細胞およびその子孫におけるウイルス遺伝子配列の保持を生じる。レトロウイルスゲノムは、3つの遺伝子 (*gag*、*pol*、および *env*) を含み、これらはそれぞれ、カプシドタンパク質、ポリメラーゼ酵素およびエンベロープ成分をコードする。*gag* 遺伝子から上流に見出される配列は、ゲノムのウイルスへのパッケージングのためのシグナルを含む。2つの長末端反復 (LTR) 配列は、ウイルスゲノムの5' および3' 末端に存在する。これらは、強力なプロモーターおよびエンハンサー配列を含み、そしてまた、宿主細胞ゲノムへの組み込みのために必要とされる (Coffin、1990)。

#### 【0140】

レトロウイルスベクターを構築するために、本発明の1つ以上のオリゴヌクレオチドまたはポリヌクレオチド配列をコードする核酸が、複製欠失のウイルスを産生するために、特定のウイルス配列のかわりにウイルスゲノムに挿入される。ピリオンを産生するために、*gag*、*pol* および *env* を含むが LTR およびパッケージング成分を含まないパッケージング細胞株が構築される (Mannら、1983)。cDNA を含む組換えプラスミドが、レトロウイルス LTR およびパッケージング配列と共にこの細胞株に導入される場合 (例えば、リン酸カルシウム沈降によって)、パッケージング配列により、組換えプラスミドの RNA 転写物はウイルス粒子にパッケージングされ得、これは次いで、培地に分泌される (Nicolas および Rubenstein、1988; Temin、1986; Marunら、1983)。次いで、組換えレトロウイルスを含む培地を収集し、必要に応じて濃縮し、そして遺伝子伝達のために使用する。レトロウイルスベクターは、広範な様々な細胞型を感染し得る。しかし、組み込みおよび安定な発現は、宿主細胞の分裂を必要とする (Paskindら、1975)。

#### 【0141】

レトロウイルスベクターの特異的ターゲティングを可能にするように設計された新規のアプローチは、最近、ウイルスエンベロープにラクトース残基を化学的に付加することによるレトロウイルスの化学修飾に基づいて開発された。この修飾は、シアログリコタンパク質レセプターを介する肝細胞の特異的感染を可能にし得る。

#### 【0142】

組換えレトロウイルスのターゲティングのための異なるアプローチが設計され、ここでレトロウイルスエンベロープタンパク質および特異的細胞レセプターに対するビオチン化抗体が使用された。これらの抗体は、ストレプトアビジンを使用することによってビオチン成分を介して連結された (Rouxら、1989)。主な組織適合性結合体クラス I およびクラス II 抗体に対する抗体を使用して、これらはインビトロにおいてエクトロピックウイルスで表面抗原に穴を開ける様々なヒト細胞の感染を示した (Rouxら、1989)。

#### 【0143】

##### (3. アデノ随伴ウイルス)

AAV (Ridgeway、1988; Hermonat および Muzycska、1984) は、アデノウイルス株の混入物として発見されたパルボウイルスである。これはいずれの疾患にも関連しない偏在性ウイルス (抗体は、US ヒト集団の85%に存在する) である。これはまた、デPENDウイルスとして分類される。なぜなら、この複製物は、アデノウイルスのようなヘルパーウイルスの存在に依存するからである。5個の血清型が単離され、このうち AAV-2 は最もよく特徴付けされる。AAV は、カプシドタンパク質である VP1、VP2 および VP3 にカプシド形成された一本差の線状 DNA を有し、20~24 nm の直径の正二十面体ピリオンを形成する (Muzyczka および McLaughlin、1988)。

#### 【0144】

AAV DNA は、約4.7キロ塩基長である。これは2つのオープンリーディングフレームを含み、そして2つの ITR により隣接される。2つの主な遺伝子 (*rep* および *cap*) が AAV ゲノム中に存在する。*rep* 遺伝子は、ウイルス複製の原因となるタンパク質についてコードし、一方 *cap* は、カプシドタンパク質 VP1~3 についてコードす

10

20

30

40

50

る。各 I T R は、T 型ヘアピン構造を形成する。これらの末端反復は、染色体組込みのための A A V の必須シス成分のみである。従って、A A V は、送達のための遺伝子のカセットによって除去および置換される全てのウイルスコード配列を有するベクターとして使用され得る。3つのウイルスプロモーターが同定され、そしてそれらのマップ位置に従って、p 5、p 19 および p 40 と命名された。p 5 および p 19 からの転写は、r e p タンパク質の産生を生じ、そして p 40 からの転写は、カプシドタンパク質を産生する (H e r m o n a t および M u z y c z k a、1984)。

#### 【0145】

発現ベクターとして r A A V を使用する可能性を研究するように研究者を促すいくつかの因子が存在する。1つは、遺伝子を送達して宿主染色体に組込むための要件が驚くほど少ないことである。これは、145bpの I T R (これは、A A V ゲノムのたった6%である)を有することが必要である。これは、ベクターにおいて4.5kbの D N A 挿入物を構築する余地がある。この運搬能力は、A A V が大きな遺伝子を送達することを妨げ得るが、これは本発明のアンチセンス構築物を送達するために十分適切である。

10

#### 【0146】

A A V はまた、その安全性に起因して送達ベシクルの良好な選択肢である。比較的複雑な救出メカニズムが存在する。野生型アデノウイルスだけでなく A A V 遺伝子が r A A V を動員するために必要とされる。同様に、A A V は病原性ではなく、いずれの疾患にも関連しない。ウイルスコード配列の除去は、ウイルス遺伝子発現に対する免疫反応を最小にし、従って r A A V は、炎症性応答を誘起しない。

20

#### 【0147】

##### (4. 発現構築物としての他のウイルスベクター)

本発明において、他のウイルスベクターは、宿主細胞へのオリゴヌクレオチドまたはポリヌクレオチド配列の送達のための発現構築物として利用され得る。痘疹ウイルス (R i d g e w a y、1988; C o u p a r ら、1988)、レンチウイルス、ポリオウイルスおよびヘルペスウイルスのようなウイルスに由来するベクターが利用され得る。これらは、種々の哺乳動物細胞に対していくつかの魅力的な特性を提供する (F r i e d m a n n、1989; R i d g e w a y、1988; C o u p a r ら、1988; H o r w i c h ら、1990)。

#### 【0148】

欠損 B 型肝炎ウイルスの最近の認識によって、異なるウイルス配列の構造と機能の関係に対して新しい見識が得られた。インビトロ研究は、ウイルスは、そのゲノムの80%に至る欠失にもかかわらず、ヘルパー依存性 (h e l p e r - d e p e n d e n t) パッケージングおよび逆方向転写のための能力を保持し得ることを示した (H a o r w i c h ら、1990)。これは、大部分のゲノムは、外来の遺伝的材料によって置換され得ることを示唆した。肝臓向性 (h e p a t o t r o p i s m) および持続性 (組込み) は、肝臓に方向付けられた遺伝子転移にとって特に魅力的な特性である。C h a n g ら (1991) は、クロラムフェニコールアセチルトランスフェラーゼ (C A T) 遺伝子を、ポリメラーゼ、表面、および前表面 (p r e - s u r f a c e) をコードする配列のかわりに、カモの B 型肝炎ウイルスゲノムへ導入した。この遺伝子は、野生型ウイルスと共に、鳥類のヘパトーム細胞株へと同時トランスフェクトされた。高力価の組換えウイルスを含む細胞培地を使用して、子ガモの初代肝細胞を感染させた。安定な C A T 遺伝子発現は、トランスフェクション後少なくとも24日の間検出された (C h a n g ら、1991)。

30

40

#### 【0149】

##### (5. ウイルスを含まないベクター)

本発明のオリゴヌクレオチドまたはポリヌクレオチド配列を発現させるために、発現構築物は細胞へ送達されねばならない。この送達は、インビトロで、細胞株を形質転換するための研究室の手順におけるように、あるいは、インビボまたはエキソビボで、ある疾患の状態の処置におけるように、達成され得る。上に記載されるように、1つの送達のための好ましい機構は、発現構築物が感染ウイルス粒子の中に封入されたウイルス感染を介する

50

。

## 【0150】

一旦他の発現構築物が細胞内に送達された場合、所望のオリゴヌクレオチドまたはポリヌクレオチド配列をコードする核酸は、異なる部位に配置され、そして発現し得る。特定の実施形態において、この構築物をコードする核酸は、細胞のゲノムの中へ安定に組み込まれ得る。この組み込みは、相同の組換えによって特定の配置および配向にあり得る（遺伝子置換）か、またはこれは、無作為な、特定されない位置に組み込まれ得る（遺伝子増強）。なおさらなる実施形態において、核酸は、DNAの別のエピソーム部分として、細胞内に安定に保持され得る。このような核酸部分または「エピソーム」は、宿主の細胞周期から独立または同調して、保持および複製を可能とするのに十分な配列をコードする。発現構築物がどのように細胞に送達されるか、および細胞のどこに核酸が留まるかは、利用される発現構築物の型に依存する。

10

## 【0151】

本発明の特定の実施形態において、1以上のオリゴヌクレオチドまたはポリヌクレオチド配列を含む発現構築物は、単に、裸の組換えDNAまたはプラスミドからなり得る。この構築物の転移は、細胞膜を物理的または化学的に透過させる、上記の任意の方法によって行われ得る。これは、特にインビトロでの転移に対して利用可能であるが、これはまたインビボでの使用にも適用され得る。Dubenskyら(1984)は、成体および新生児のマウスの肝臓および脾臓へリン酸カルシウム沈降物の形態で、ポリオーマウイルスDNAを首尾よく注入し、このマウスは、活発なウイルス複製および急性感染を示した。BenvenistyおよびReshef(1986)らはまた、サルチル酸カルシウムを沈降させたプラスミドの直接の腹腔内注射は、トランスフェクトした遺伝子の発現を引き起こすことを示した。目的の遺伝子をコードするDNAはまた、インビトロで同様の形式で転移され、そして遺伝子産物を発現することが想像される。

20

## 【0152】

裸のDNA発現構築物を細胞へ転移するための本発明の別の実施形態は、粒子のボンバードメントを含み得る。この方法は、DNAでコートされた微小発射体が細胞膜を貫通し、そして細胞を殺傷することなく細胞内へ入ることを可能とする高い速度まで、微小発射体を加速する能力に依存する(Kleinら、1987)。小さな粒子を加速するためのいくつかのデバイスが開発されている。このようなデバイスの1つは、電流を発生させるための高い電圧の放電に依存し、この電流は、代わりに、輸送力を提供する(Yangら、1990)。使用された微小発射体は、タングステンまたは金のビーズのような生物学的に不活性な物質から構成された。

30

## 【0153】

ラットおよびマウスの肝臓、皮膚ならびに筋肉組織を含む選択された器官は、インビボでボンバードされた(Yangら、1990; Zeleninら、1991)。これは、銃と標的器官との間のあらゆる介在組織を排除するために、組織または細胞の外科的露出を必要とし得る(すなわち、エキソビボ処置)。再び、特定の遺伝子をコードするDNAは、この方法によって送達され得、そして依然として本発明によって組み込まれ得る。

40

## 【0154】

(ポリペプチド組成物)

本発明は、他の局面において、ポリペプチド組成物を提供する。一般に、本発明ポリペプチドは、哺乳動物種に由来する単離されたポリペプチド(あるいは、エピトープ、変異体またはそれらの活性なフラグメント)である。好ましくは、ポリペプチドは、本明細書中で開示されるポリヌクレオチド配列、または中程度にストリンジェントな条件下で、本明細書中で開示されるポリヌクレオチド配列にハイブリダイズする配列によってコードされる。あるいは、ポリペプチドは、本明細書中で開示されるアミノ酸配列由来の連続したアミノ酸配列を含むか、または本明細書中で開示されるアミノ酸配列全体を含むポリペプチドとして規定され得る。

## 【0155】

50



免疫原性部分は、一般に、周知の技術（例えば、Paul, Fundamental Immunology, 第3版、243-247(1993)およびその中で引用される参考文献に概説される技術）を用いて同定され得る。このような技術としては、抗原特異的抗体、抗血清および/またはT細胞株もしくはクローンと反応する能力についてのポリペプチドのスクリーニングが挙げられる。本明細書中で使用される場合、抗血清および抗体は、抗原に特異的に結合する（すなわち、これらが、ELISAまたは他の免疫アッセイにおいてタンパク質と反応し、そして関連しないタンパク質と検出可能に反応しない）場合、「抗原特異的」である。このような抗血清および抗体は、本明細書中に記載されるように、そして周知の技術を用いて調製され得る。Mycobacterium sp. タンパク質の免疫原性部分は、全長ポリペプチドの反応性より実質的に勝るとも劣らないレベルで、（例えば、ELISAおよび/またはT細胞反応アッセイにおいて）このような抗血清および/またはT細胞と反応する部分である。このような免疫原性部分は、全長のポリペプチドの反応性と類似の、またはそれより高いレベルで、このようなアッセイ内で反応し得る。このようなスクリーニングは、一般に、当業者に周知の方法（例えば、HarlowおよびLane, Antibodies: A Laboratory Manual (1988)に記載される方法）を用いて実施され得る。例えば、ポリペプチドは、固体支持体に固定され、そして患者の血清と接触され、固定されたポリペプチドへの血清内での抗体の結合を可能とし得る。次いで、結合しない血清は、除去され、そして結合した抗体は、例えば、<sup>125</sup>I 標識されたプロテインAを用いて検出され得る。

10

## 【0156】

20

ポリペプチドは、任意の種々の周知の技術を用いて調製され得る。上記のようにDNA配列によってコードされる組換えポリペプチドは、当業者に公知の任意の種々の発現ベクターを用いて、DNA配列から容易に調整され得る。発現は、組換えポリペプチドをコードするDNA分子を含む発現ベクターで形質転換されるかまたはトランスフェクトされた任意の適切な宿主細胞において達成され得る。適切な宿主細胞としては、原核生物、酵母および高等真核細胞（例えば、哺乳動物細胞または植物細胞）が挙げられる。好ましくは、使用される宿主細胞は、E. coli、酵母あるいは哺乳動物細胞株（例えば、COSまたはCHO）である。培養培地に組換えタンパク質またはポリペプチドを分泌する適切な宿主/ベクター系由来の上清は、市販のフィルターを用いてまず濃縮され得る。濃縮後、この濃縮物は、アフィニティマトリックスまたはイオン交換樹脂のような適切な精製マトリックスに適用され得る。最後に、1以上の逆相HPLC工程が使用され、組換えポリペプチドをさらに精製し得る。

30

## 【0157】

本発明のポリペプチド、それらの免疫原性フラグメント、ならびに約100アミノ酸未満、そして一般に約50アミノ酸未満を有する他の改変体はまた、当業者に周知の技術を用いて、合成手段によって生成され得る。例えば、このようなポリペプチドは、任意の市販の固相技術（例えば、Merrifieldの固相合成方法（この技術では、アミノ酸が、成長するアミノ酸鎖に連続的に付加される））を用いて合成され得る。Merrifield, J. Am. Chem. Soc. 85: 2149-2146 (1963)を参照のこと。ポリペプチドの自動合成のための装置は、Perkin Elmer/Applied Biosystems Division (Foster City, CA)のような供給者から市販され、そして製造者の指示に従って操作され得る。

40

## 【0158】

ある特定の実施形態において、ポリペプチドは、本明細書中に記載されるような複数のポリペプチドを含むか、または本発明に記載されるような少なくとも1つのポリペプチドおよび無関係の配列（例えば、公知の腫瘍タンパク質）を含む融合タンパク質であり得る。融合パートナーは、例えば、Tヘルパーエピトープ、好ましくはヒトによって認識されるTヘルパーエピトープを提供するのを補助し得るか（免疫学的融合パートナー）、またはネイティブな組換えタンパク質より高い収率でタンパク質を発現するのを補助し得る（発現エンハンサー）。特定の好ましい融合パートナーは、免疫学的融合パートナーと発現増

50

強融合パートナーとの両方である。他の融合パートナーは、タンパク質の溶解性を増大させるため、またはタンパク質が所望の細胞内画分に対して標的化されることを可能とするために選択され得る。なおさらなる融合パートナーとしては、タンパク質の精製を容易にするアフィニティータグが挙げられる。

【0159】

融合タンパク質は、一般に、化学結合体化を含む標準的技術を用いて調製され得る。好ましくは、融合タンパク質は、組換えタンパク質として発現され、発現系において、非融合タンパク質と比較して、増大したレベルの産生を可能とする。簡潔にいうと、ポリペプチド成分をコードするDNA配列は、別々に組み立てられ得、そして適切な発現ベクターへ連結され得る。1つのポリペプチド成分をコードするDNA配列の3'末端は、ペプチドリinkerありまたはなしで、第2のポリペプチド成分をコードするDNA配列の5'末端に連結し、その結果、配列のリーディングフレームは、一致する。これは、両方の成分ポリペプチドの生物学的活性を保持する単一の融合タンパク質への翻訳を可能とする。

10

【0160】

ペプチドリinker配列を使用して、各ポリペプチドが、その二次および三次構造へと折り畳まれるのを確かめるのに十分な距離で第1および第2のポリペプチド成分を分離し得る。このようなペプチドリinker配列は、当該分野において周知の標準的な技術を用いて、融合タンパク質へと組み込まれる。適切なペプチドリinker配列は、以下の因子に基づいて選択され得る：(1)可撓性の伸長されたコンフォメーションをとる能力；(2)第1および第2のポリペプチド上の機能的エピトープと相互作用し得る二次構造をとることができないこと；および(3)ポリペプチドの機能的エピトープと反応し得る疎水性残基または荷電残基の欠如。好ましいペプチドリinker配列は、Gly、AsnおよびSer残基を含む。他のほぼ中性のアミノ酸（例えば、ThrおよびAla）もまた、リンカー配列において使用され得る。リンカーとして有用に使用され得るアミノ酸配列としては、Marateaら、Gene 40:39-46(1985)；Murphyら、Proc. Natl. Acad. Sci. USA 83:8258-8262(1986)；米国特許第4,935,233号および米国特許第4,751,180号に開示されるアミノ酸配列が挙げられる。リンカー配列は、一般に、1~約50アミノ酸長であり得る。リンカー配列は、第1および第2のポリペプチドが必須でないN末端アミノ酸領域（これは、機能的ドメインを分離しそして位置的干渉を妨げるために使用され得る）を有する場合、必要とされない。

20

30

【0161】

連結されたDNA配列は、適切な転写または翻訳調節エレメントに作動可能に結合される。DNAの発現を担うこの調節エレメントは、第1のポリペプチドをコードするDNA配列に対して5'側のみ位置する。同様に、翻訳終らせるために必要とされる終止コドンおよび転写終結シグナルは、第2のポリペプチドをコードするDNA配列に対して3'側のみ存在する。

【0162】

融合タンパク質もまた提供される。このようなタンパク質は、無関係の免疫原性タンパク質と共に、本明細書中に記載されるようなポリペプチドを含む。好ましくは、この免疫原性タンパク質は、リコール応答を誘発し得る。このようなタンパク質の例としては、破傷風、結核および肝炎タンパク質が挙げられる（例えば、Stouteら、New Engl. J. Med. 336:86-91(1997)を参照のこと）。

40

【0163】

好ましい実施形態において、免疫原性融合パートナーは、プロテインD（グラム陰性細菌Haemophilus influenza Bの表面タンパク質）に由来する（WO 91/18926）。好ましくは、プロテインD誘導体は、このタンパク質のほぼ最初の3分の1（例えば、最初のN末端の100~110個のアミノ酸）を含み、そしてプロテインD誘導体は、脂質化（lipidate）され得る。特定の好ましい実施形態において、リポプロテインD融合パートナーの最初の109個の残基は、N末端に含まれ、ポリ

50

ペプチドにさらなる外因性T細胞エピトープを提供し、そしてE. coliにおける発現レベルを増大させる(従って、発現エンハンサーとして機能する)。脂質の尾部は、抗原提示細胞に対する、最適な抗原の提示を保証する。他の融合パートナーとしては、インフルエンザウイルス由来の非構造的タンパク質NS1(血球凝集素)が挙げられる。典型的には、N末端の81アミノ酸が使用されるが、Tヘルパーエピトープを含む異なるフラグメントが使用され得る。

#### 【0164】

別の実施形態において、免疫学的融合パートナーは、LYTAとして公知のタンパク質、またはその一部分(好ましくはC末端部分)である。LYTAは、アミダーゼLYTA(LytA遺伝子によってコードされる; Gene 43: 265-292(1986))として公知のN-アセチル-L-アラニンアミダーゼを合成するStreptococcus pneumoniae由来である。LYTAは、ペプチドグリカン骨格において特定の結合を特異的に分解する自己溶解素である。LYTAタンパク質のC末端ドメインは、コリンまたはいくつかのコリンアナログ(例えば、DEAE)に対する親和性を担う。この特性は、融合タンパク質の発現に有用であるE. coli C-LYTA発現プラスミドの開発のために活用された。アミノ末端でC-LYTAフラグメントを含むハイブリッドタンパク質の精製は、記載されている(Biotechnology 10: 795-798(1992))を参照のこと)。好ましい実施形態において、LYTAの反復部分は、融合タンパク質の中に取り込まれ得る。反復部分は、178残基で始まるC末端領域において見出される。特に好ましい反復部分は、188残基~305残基を取り込む。

#### 【0165】

一般的に、本明細書中に記載されるポリペプチド(融合タンパク質を含む)およびポリヌクレオチドは、単離される。「単離された」ポリペプチドまたはポリヌクレオチドは、その本来の環境から取り出されたポリペプチドまたはポリヌクレオチドである。例えば、天然系において共存する物質のいくつかまたは全てから分離される場合、天然に存在するタンパク質は、単離される。好ましくは、このようなポリペプチドは、少なくとも約90%純粋、より好ましくは少なくとも約95%純粋および最も好ましくは少なくとも約99%純粋である。ポリヌクレオチドは、例えば、天然な環境の一部でないベクター中にクローンされる場合、単離されるとみなされる。

#### 【0166】

(T細胞)

免疫治療組成物はまた、あるいは免疫治療組成物は、Mycobacterium抗原に特異的なT細胞を含み得る。このような細胞は、一般的に、標準的な手順を使用してインビトロまたはエキソビボで調製され得る。例えば、T細胞は、市販されている細胞分離システム(例えば、Nexell Therapeutics, Inc. (Irvine, CA; 米国特許第5,240,856号; 米国特許第5,215,926号; WO89/06280; WO91/16116およびWO92/07243をまた参照のこと)から入手可能なIsoplex<sup>TM</sup> System)を使用して、骨髄、末梢血、または患者の骨髄もしくは末梢血の画分から単離され得る。あるいは、T細胞は、関連もしくは非関連のヒト、非ヒト哺乳動物、細胞株または培養物由来であり得る。

#### 【0167】

T細胞は、本発明のポリペプチド、このようなポリペプチドをコードするポリヌクレオチド、および/またはこのようなポリペプチドを発現する抗原提示細胞(APC)で刺激され得る。このような刺激は、ポリペプチドに特異的なT細胞の生成が可能な十分な条件下および時間で実行される。好ましくは、ポリペプチドまたはポリヌクレオチドは、送達ビヒクル(例えば、ミクロスフェア)中に存在して、特異的なT細胞の生成を容易にする。

#### 【0168】

T細胞が、特異的に増殖するか、サイトカインを分泌するか、またはポリペプチドで被覆されたか、もしくはポリペプチドをコードする遺伝子を発現する標的細胞を殺す場合、T細胞は、本発明のポリペプチドに特異的であるとみなされる。T細胞特異性は、任意の種

10

20

30

40

50

々の標準的な技術を使用して評価され得る。例えば、クロム放出アッセイまたは増殖アッセイにおいて、ネガティブコントロールと比較して溶解および/または増殖において2倍より大きい増加の刺激指数は、T細胞特異性を示す。このようなアッセイは、例えば、Chenら、Cancer Res. 54:1065-1070(1994)に記載されるように実行され得る。あるいは、T細胞の増殖の検出は、種々の公知の技術によって達成され得る。例えば、T細胞増殖は、DNA合成の増加した速度を測定することによって(例えば、トリチウム化チミジンを用いるT細胞のパルス標識培養物およびDNA中に取り込まれたトリチウム化チミジンの量の測定によって)検出され得る。3~7日間の本発明のポリペプチドとの接触(100ng/ml~100μg/ml、好ましくは200ng/ml~25μg/ml)は、T細胞の増殖において少なくとも2倍の増加を生じるはずである。2~3時間の上記の接触は、サイトカイン(例えば、TNFまたはIFN-)放出のレベルにおける2倍の増加が、T細胞活性化の指標である標準的サイトカインアッセイを使用して測定されるような、T細胞の活性化を生じるはずである。(Coliganら、Current Protocols in Immunology、第一巻(1998)を参照のこと)。ポリペプチド、ポリヌクレオチドまたはポリペプチド発現APCに应答して活性化されたT細胞は、CD4<sup>+</sup>および/またはCD8<sup>+</sup>であり得る。タンパク質特異的T細胞は、標準的技術を使用して増殖され得る。好ましい実施形態において、T細胞は、患者、関連ドナー、または非関連ドナー由来であり、そして刺激および増殖の後に患者に投与される。

10

## 【0169】

20

治療的目的のために、ポリペプチド、ポリヌクレオチドまたはAPCに应答して増殖するCD4<sup>+</sup>T細胞またはCD8<sup>+</sup>T細胞は、インビトロまたはインビボで数の上で増殖され得る。インビトロでのこのようなT細胞の増殖は、種々の方法において達成され得る。例えば、T細胞は、T細胞増殖因子(例えば、インターロイキン-2)および/もしくはポリペプチドを合成する刺激細胞の添加を伴ってかまたは伴わずに、ポリペプチド、またはこのようなポリペプチドの免疫原性部分に対応する短いペプチドに対して再曝露され得る。あるいは、タンパク質の存在下で増殖する1つ以上のT細胞は、クローニングによって数の上で増殖され得る。細胞をクローニングする方法は、当該分野で周知であり、そして限定希釈を含む。

## 【0170】

30

(薬学的組成物)

さらなる実施形態において、本発明は、細胞もしくは動物への投与のための、薬学的に受容可能な溶液または生理学的に受容可能な溶液中の、単独でかあるいは1つ以上の他の治療様式と組み合わせるのいずれかでの、本明細書中で開示される1つ以上のポリヌクレオチド、ポリペプチド、T細胞および/または抗体組成物の処方物に関する。このような組成物はまた、診断的使用のために有用である。

## 【0171】

所望の場合、本明細書中に開示されるポリペプチドを発現する核酸セグメント、RNA、DNAまたはPNA組成物が、他の薬剤および、例えば他のタンパク質またはポリペプチドあるいは種々の薬学的に活性な薬剤と組み合わせる投与され得ることがまた、理解される。実際に、さらなる薬剤が、標的細胞または宿主組織との接触において有意な悪影響を生じないことを考慮すると、また含まれ得る他の成分に対して実質的には限定は存在しない。従って、この組成物は、特定の例において必要とされる種々の他の薬剤と共に送達され得る。このような組成物は、宿主細胞または他の生物学的供給源から精製され得るか、あるいは、本明細書中に記載されるように化学的に合成され得る。同様に、このような組成物はさらに、置換されたまたは誘導体化されたRNAまたはDNA組成物を含む。

40

## 【0172】

薬学的に受容可能な賦形剤およびキャリア溶液の処方、種々の処置レジメン(例えば、経口、非経口、静脈内、鼻腔内、および筋肉内投与および処方を含む)で本明細書中に記載される特定の組成物を使用する、適切な用量および処置レジメンの開発と同様、当業者に

50

周知である。

【0173】

(1. 経口送達)

特定の適用において、本明細書中に開示される薬学的組成物は、動物への経口投与を介して送達される。このように、これらの組成物は、不活性の賦形薬と共にもしくは同化可能な食用キャリアと共に処方され得るか、またはこれらの組成物は、硬殻ゼラチンカプセルもしくは軟殻ゼラチンカプセルに封入され得るか、またはこれらの組成物は、錠剤中に圧縮され得るか、またはこれらの組成物は、治療食の食物と共に直接取り込まれ得る。

【0174】

活性化化合物でさえ、賦形剤と共に取り込まれ、そして摂取錠剤、バツカル錠、トローチ、カプセル剤、エリキシル剤、懸濁剤、シロップ剤、カシエ剤などの形態で使用される (Mathiowitzら、1997; Hwangら、1998; 米国特許第5,641,515号; 米国特許第5,580,579号および米国特許第5,792,451号、それぞれは、その全体が参考として明確に本明細書中に援用される)。錠剤、トローチ、丸剤、カプセル剤などもまた、以下: 結合剤 (ガムトラガント、アカシア、コーンスターチ、またはゼラチンなど); 賦形剤 (リン酸二カリウムなど); 崩壊剤 (コーンスターチ、ポテトスターチ、アルギン酸など); 滑沢剤 (ステアリン酸マグネシウム); および甘味剤 (添加され得るスクロース、ラクトースもしくはサッカリンなど) または香味剤 (ペパーミント、ウィンターグリーンオイル、またはチェリー香味など) を含み得る。投薬単位形態が、カプセル剤である場合、上記の型の材料に加えて、液体キャリアを含み得る。コーティング剤としてか、または他に投薬単位の物理的形態を改変するために、種々の他の物質が、存在し得る。例えば、錠剤、丸剤またはカプセル剤は、シェラック、砂糖、または両方でコートされ得る。エリキシルのシロップ剤は、活性化化合物、甘味剤としてスクロース、防腐剤としてメチルおよびプロピルパラベン、色素ならびに香味剤 (チェリーまたはオレンジ香味など) を含み得る。もちろん、任意の投薬単位形態を調製する際に使用する任意の材料は、薬学的にきれいで、そして使用される量において実質的に無毒性でなければならない。さらに、活性化化合物は、持続放出調製物および処方物中に組み込まれ得る。

【0175】

代表的に、これらの処方物は、少なくとも約0.1%以上の活性化化合物を含み得るが、活性成分のパーセンテージは、もちろん変化し得、そして都合よく総処方物の重量または容量の約1%もしくは2%と約60%もしくは70%以上との間である。自然に、各治療的に有用な組成物中の活性化化合物の量は、適切な投薬が、化合物の任意の所定の単位用量において得られるような方法で調製され得る。溶解度、生物学的利用能、生物学的半減期、投与経路、生成物貯蔵寿命、および他の薬理学的考慮などの因子は、このような薬学的処方物を調製する当業者によって意図され、そのため、種々の投薬および処置のレジメンが、所望され得る。

【0176】

あるいは、経口投与について、本発明の組成物は、うがい薬、歯磨き剤、バツカル錠、経口スプレー、または舌下経口的投与処方物の形態で1つ以上の賦形剤と共に組み込まれ得る。例えば、うがい薬は、適切な溶媒 (ホウ酸ナトリウム溶液 (Dobell's 溶液) など) 中に必要とされる量で活性成分を組み込んで調製され得る。あるいは、活性成分は、経口溶液中 (ホウ酸ナトリウム、グリセリンおよび炭酸水素カリウムを含む溶液など) に組み込まれるか、または歯磨き剤に分散されるか、または水、結合剤、研磨剤、香味剤、発泡剤、および湿潤剤を含み得る組成物に治療的に有効な量で加えられ得る。あるいは、組成物は、舌下に置かれ得るかまたは別に口の中で溶解される錠剤または溶液型に形成され得る。

【0177】

(2. 注入可能送達)

特定の状況において、米国特許第5,543,158号; 米国特許第5,641,515

10

20

30

40

50

号および米国特許 5, 399, 363号(各々は、その全体が参考として明確に本明細書中に援用される)に記載されるように本明細書中に開示される薬学的組成物を非経口に、静脈内に、筋肉内にまたは腹腔内でさえ送達することが所望される。遊離塩基または薬学的に受容可能な塩としての活性化化合物の溶液は、界面活性剤(ヒドロキシプロピルセルロースなど)と適度に混合された水において調製され得る。分散はまた、グリセロール、液体ポリエチレングリコール、およびそれらの混合物中で、ならびに油中で調製され得る。貯蔵および使用の通常の条件下で、これらの調製物は、微生物の増殖を阻止するため防腐剤を含む。

#### 【0178】

注入可能な使用に適切な薬学的形態は、滅菌注入可能な溶液または分散剤の即時調製のための滅菌水溶液または分散剤および滅菌粉末を含む(米国特許第5, 466, 468号、その全体が参考として明確に本明細書中に援用される)。全ての場合において、形態は、滅菌されるべきであり、簡単な注入能力がある程度まで流動性であるべきである。形態は、製造および貯蔵条件下で安定であるべきであり、そして微生物(例えば、細菌および真菌)の汚染行為に対して保護されるべきである。キャリアは、溶媒または分散培地(例えば、水、エタノール、ポリオール(例えば、グリセロール、プロピレングリコール、および液体ポリエチレングリコールなど)、適切なそれらの混合物、および/または植物性油を含む)であり得る。適切な流動率は、例えば、コーティング剤(例えば、レシチン)の使用によって、分散の場合には必要とされる粒子サイズの維持によって、および界面活性剤の使用によって維持され得る。微生物の作用の予防は、種々の抗細菌および抗真菌薬剤(例えば、パラベン、クロロブタノール、フェノール、ソルビン酸、チメローサルなど)によって促進され得る。多くの場合において、等張剤(例えば、砂糖または塩化ナトリウム)を含むことが好ましい。注入可能な組成物の持続性の吸収は、吸収遅延剤の組成物(例えばモノステアリン酸アルミニウムおよびゼラチン)における使用によって引き起こされ得る。

#### 【0179】

例えば、水溶液の非経口投与のために、必要であれば溶液は適切に緩衝化され、そして液体の希釈液は、最初に十分な生理食塩水またはグルコースで等張性を与えられるべきである。これらの特定の水溶液は、特に、静脈内投与、筋肉内投与、皮下投与、および腹腔内投与に適する。これに関連して、使用され得る滅菌した水溶性の媒体が本開示から考慮して当業者に公知である。例えば、1回の投与量が、1mlの等張性のNaCl溶液に溶解され得、そして1000mlの皮下注入流体へ添加され得るかまたは提案された注入部位へ注入され得るかのいずれかである(例えば、Remington's Pharmaceutical Sciences、第15版、pp. 1035~1038および1570~1580を参照のこと)。処置される被験体の状態に依存して投与量のいくつかのバリエーションが必然的に生じる。いずれにしても、投与責任がある人が個々の被験体に対する適切な用量を決定する。さらに、ヒトへの投与のために、調製物は、FDA生物製剤局の基準により要求される無菌性基準、パイロジェン性基準(pyrogenicity)、ならびに一般的な安全性基準および純度基準を満たすべきである。

#### 【0180】

滅菌した注射可能な溶液は、上記に列挙された種々の他の成分を有する適切な溶媒中に必要な量で活性化化合物を組み込むことにより調製され、必要に応じてその後濾過滅菌される。分散物は、一般的に、種々の滅菌した活性成分を基本的な分散媒体および上記に列挙されたこれらに由来する必要な他の成分を含む滅菌ビヒクルに組み込むことにより調製される。滅菌した注射可能な溶液の調製物のための滅菌粉末の場合、好ましい調製方法は滅菌乾燥技術および凍結乾燥技術であり、これらは、活性成分の粉末およびあらかじめ濾過滅菌したこれらの溶液由来のさらに所望の任意の成分の粉末を得る技術である。

#### 【0181】

本明細書中に開示された組成物は、中性または塩の形態で処方され得る。薬学的に受容可能な塩として、酸付加塩(タンパク質の遊離アミノ基で形成された)が挙げられ、これら

10

20

30

40

50

は無機酸（例えば、塩酸またはリン酸など）、または有機酸（酢酸、シュウ酸、酒石酸、マンデル酸など）で形成される。遊離カルボキシル基で形成された塩もまた、無機塩基（例えば、ナトリウム、カリウム、アンモニウム、カルシウム、または水酸化第二鉄など）および有機塩基（イソプロピルアミン、トリメチルアミン、ヒスチジン、プロカインなど）から誘導され得る。処方に関して、溶液が、投与処方物と適合する様式で投与され、このような量で治療的に有効である。注射可能な溶液、薬物放出カプセルなどのような種々の投与形態で、処方物が容易に投与される。

#### 【0182】

本明細書中で使用される場合、「キャリア」として、任意のそして全ての溶媒、分散媒体、ビヒクル、コーティング剤、希釈液、抗菌剤ならびに抗真菌剤、等張剤ならびに吸収遅延剤、緩衝液、キャリア溶液、懸濁液、コロイドなどが挙げられる。薬学的活性物質のためのこのような媒体および薬剤の使用は当業者に周知である。従来の任意の媒体または薬剤が活性成分と不適合である範囲を除き、治療的組成物におけるその使用が企図される。補助的な活性成分がまた、組成物に組み込まれ得る。

10

#### 【0183】

句「薬学的に受容可能」とは、ヒトへ投与される場合に、アレルギー反応または類似の有害反応を生じない分子実体および組成物を言う。活性成分としてタンパク質を含む水性組成物の調製物が、当業者に周知である。代表的には、このような組成物は注射可能に、液体溶液または懸濁液のいずれかとして調製され；注射の前に液体中への溶解または懸濁に適した固形形態もまた調製され得る。調製物はまた、乳化され得る。

20

#### 【0184】

##### （3．経鼻送達）

特定の実施形態において、薬物学的組成物は、鼻腔スプレー、吸入、および/または他のエアロゾルの送達ビヒクルにより送達され得る。経鼻エアロゾルスプレーを介して直接肺に遺伝子、核酸、およびペプチドの組成物を送達する方法が記載されてきた（例えば、米国特許第5,756,353号および米国特許第5,804,212号（各々その全体が参考として本明細書中で詳細に援用される））。同様に、鼻腔微粒子樹脂（Takenagaら、1998）およびリゾホスファチジル-グリセロール化合物（米国特許第5,725,871号、その全体が参考として本明細書中で詳細に援用される）を使用する薬物送達はまた、薬学分野で周知である。同様に、ポリテトラフルオロエチレン支持マトリックス形態の粘膜輸送による薬物送達が米国特許第5,780,045号（その全体が参考として詳細に援用される）に記載される。

30

#### 【0185】

##### （4．リポソーム-媒介送達、ナノカプセル-媒介送達、およびマイクロ粒子-媒介送達）

特定の実施形態において、発明者らは、本発明の組成物の適切な宿主細胞への導入のためにリポソーム、ナノカプセル、マイクロ粒子、ミクロスフェア、脂質粒子、ビヒクルなどの使用を企図する。特に、本発明の組成物は、脂質粒子、リポソーム、小胞、ナノスフェアまたはナノ粒子などのいずれかにカプセル化された送達のために処方され得る。

#### 【0186】

このような処方物は本明細書中に開示された核酸または構築物の薬学的に受容可能な処方物の導入のために好まれ得る。リポソームの形成および使用は一般的に当業者に公知である（例えば、Couvreurら、1977；Couvreur、1988；Lasica、1998を参照のこと；これらは、細胞内細菌感染および疾患に対する標的化された抗生物質治療におけるリポソームおよびナノカプセルの使用を記載する）。最近、血清の安定性および循環半減期が改善されたりポソームが開発された（GabizonおよびPapahadjopoulos、1988；AllenおよびChoun、1987；米国特許第5,741,516号、その全体が本明細書中で参考として詳細に援用される）。さらに、潜在的な薬物キャリアとしてリポソーム様調製物およびリポソーム様調製物の種々の方法が総説されてきた（Takakura、1998；Chandranら、199

40

50

7; Margalit, 1995; 米国特許第5,567,434号; 米国特許第5,552,157号; 米国特許第5,565,213号; 米国特許第5,738,868号および米国特許第5,795,587号、各々その全体が本明細書中で参考として詳細に援用される)。

#### 【0187】

他の手順によるトランスフェクションに対して通常は、耐性の多数の細胞型(T細胞の懸濁液、初代肝細胞培養物およびPC12細胞を含む)でリポソームがうまく使用された(Renneisenら、1990; Mullerら、1990)。さらに、リポソームは、ウイルスに基いた送達系の典型であるDNA長の制限がない。遺伝子、薬物(HeathおよびMartin、1986; Heathら、1986; Balazsovitsら、1989; FrestaおよびPuglisi、1996)、放射線治療剤(Pikulら、1987)、酵素(Imaizumiら、1990a; Imaizumiら、1990b)、ウイルス(FallerおよびBaltimore、1984)、転写因子およびアロステリックエフェクター(NicolauおよびGersonde、1979)を種々の培養された細胞株および動物に効果的に導入するためにリポソームが使用されてきた。さらに、リポソームに媒介された薬物送達の有効性を審査するいくつかの成功した臨床試験が終了している(Lopez-Beresteinら、1985a; 1985b; Coune、1988; Sculierら、1988)。さらに、いくつかの研究によって、リポソームの使用が全身送達後の自己免疫応答、毒性、または性腺の局在化と関連しないことが示唆される(MoriおよびFukatsu、1992)。

#### 【0188】

リポソームは水性の媒体に分散されるリン脂質から形成され、そして多重層の同心状二重層の小胞(多重層の小胞(MLV)とも呼ばれる)を自然に形成する。一般的にMLVは25nm~4μmの直径を有する。MLVの超音波処理は、200~500の範囲の直径を有する小さな単層の小胞(SUV)(核に水溶液を含む)の形成を生じる。

#### 【0189】

リポソームは細胞膜の類似物を保有し、そして本発明との組み合わせにおいてペプチド組成物のためのキャリアとしての使用を企図される。これらは、水に可溶性の物質および脂質に可溶性の物質の両者が捕捉され得る場合、(すなわち、それぞれ水性の空間中、および二重膜自体の内部にある場合に)広く適している。薬物を保有するリポソームは、リポソームの処方を選択的に改変することにより活性薬物の部位特異的送達のために、さらに使用され得ることが可能である。

#### 【0190】

Couvreurら(1977; 1988)の教示に加えて、以下の情報が、リポソームの処方物の作製に利用され得る。リン脂質は、脂質対水のモル比に依存して、水中に分散した場合、リポソーム以外の種々の構築物を形成し得る。低い比で、リポソームは好ましい構築物である。リポソームの物理的特徴はpH、イオン強度および二価のカチオンの存在に依存する。リポソームは、イオン性物質および極性物質に対して低い透過性を示し得るが、上昇した温度で相転移を起こし、リポソームの透過性が顕著に変化する。相転移は、ゲル状態として公知のきっちり詰まった整列した構築物から、流動状態として公知のゆるく詰まったより整列していない構築物への変化を含む。これは特徴的な相転移温度で起こり、そしてイオン、糖および薬物に対する透過性の増加を生じる。

#### 【0191】

温度に加えて、タンパク質への曝露がリポソームの透過性を変化し得る。特定の可溶性のタンパク質(チトクロムcなど)は結合し、変形し、そして二重膜を透過することにより透過性の変化を生じる。コレステロールは、リン脂質をよりきつく詰めることにより明らかにタンパク質の浸透を阻害する。抗生物質の送達およびインヒビターの送達のための最も有用なリポソーム処方物は、コレステロールを含むことが企図される。

#### 【0192】

溶質を捕捉する能力は、異なるリポソームの型の間で変化する。例えば、MLVは溶質の



捕捉において中程度の効率であるが、SUVは非常に非効率的である。SUVは大きさの分布において均一性および再現性の利点を提供する。しかし、大きさと捕捉する効率の間の妥協点は大きな単層の小胞(LUV)により提供される。これらはエーテルエバポレーションにより調製され、そして溶質の捕捉においてMLVよりも3から4倍効率的である。

#### 【0193】

リポソームの特徴に加えて、化合物を捕捉することにおける重要な決定因子は、化合物自身の物理科学的性質である。極性化合物は水性の空間に捕捉され、そして非極性の化合物は小胞の脂質二重層に結合する。極性化合物は、透過により、または二重層が破壊された場合に放出されるが、非極性化合物は温度または脂質タンパク質への曝露により二重層が破壊されない限り二重層に付着したままである。両方の型は、相転移温度で最大流出速度を示す。

10

#### 【0194】

4つの異なるメカニズムを介してリポソームと細胞が相互作用する：細網内皮系の食細胞(マクロファージおよび好中球など)によるエンドサイトーシス；非特異的な弱い疎水性の力もしくは弱い静電的な力、または細胞表面の成分との特定の相互作用のいずれかによる細胞表面への吸着；原形質膜へのリポソームの脂質二重膜の挿入による原形質細胞膜との融合、続いて細胞質へのリポソームの内容物の同時放出；そして、リポソームの内容物の結合のないリポソーム脂質の細胞膜または細胞成分膜(subcellular membrane)への転移による(逆も同様)。どのメカニズムが作用するか決定すること、および1より多くを同時に操作し得ることは、しばしば難しい。

20

#### 【0195】

静脈内に注射されたリポソームの運命および性質は、それらの物理的性質(大きさ、流動性、および表面の電荷など)に依存する。これらはその組成物に依存して、時間または日単位で組織中で持続し得、そして血中半減期は分から数時間の範囲である。MLVおよびLUVなどの大きいリポソームは、細網内皮系の食細胞により迅速に取りこまれるが、生理学的な循環系は、ほとんどの部位でこのような大きな種の流出を抑制する。これらは、毛質状の内皮(肝臓または脾臓の洞様毛細血管など)に存在する大きな開口または大きな孔のある場所にしか流出できない。従って、これらの器官は取りこみの主な部位である。一方、SUVは広範な組織分布を示すが、なお肝臓および脾臓に大きく離されている。一般的に、インビボでのこの挙動は、リポソームの標的化の能力を、これらの大きいサイズが利用し易いこれらの器官および組織のみに限定する。これらとして、血液、肝臓、脾臓、骨髄、およびリンパ様器官が挙げられる。

30

#### 【0196】

標的化は、一般的に、本発明において限定されない。しかし、特定の標的化が所望される場合、このことを達成するための方法が利用可能である。抗体をリポソーム表面と結合させ、方向付けるために抗体が使用され得、そして、特定の抗原レセプターに対する組成物の薬物の内容物は、特定の細胞の型の表面上に位置する。糖質の決定因子(細胞-細胞の認識、相互作用および接着において役割を果たす糖タンパク質または糖脂質の細胞表面成分)もまた、特定の細胞型へリポソームを方向付ける能力を有する認識部位として使用され得る。リポソームの調製物はほとんど、静脈内注射が使用されることが企図されるが、他の投与経路もまた考えられる。

40

#### 【0197】

あるいは、本発明は、本発明の組成物の薬学的に受容可能なナノカプセル処方物を提供する。ナノカプセルは、一般的に安定かつ再生可能な様式で化合物を捕らえる(Henry-Michellandら、1987；Quintanar-Guerreroら、1998；Douglasら、1987)。細胞内の重合体の過剰負荷に起因する副作用を回避するために、このような超微細粒子(約0.1 $\mu$ mのサイズ)は、インビボで分解可能なポリマーを用いて設計されるべきである。これらの要求に合う生分解性ポリアルキルシリアノアクリレートナノ粒子が、本発明における使用に意図される。このような粒子は、記

50

載される (Couvreurら、1980; 1988; zur Muhlenら、1998; Zambauxら、1998; Pinto-Alphandryら、1995、および米国特許第5,145,684号(これらは、その全体において本明細書中に詳細に参考として援用される)) ように、容易に作製され得る。

【0198】

(ワクチン)

本発明の特定の好ましい実施形態においては、ワクチンが提供される。このワクチンは、一般的に、免疫賦活剤と組み合わせて、上記で考察されるような1つ以上の薬学的組成物を含む。免疫賦活剤は、外来性の抗原に対する免疫応答(抗体媒介性および/または細胞媒介性)を向上または増強させる任意の物質であり得る。免疫賦活剤の例としては、アジュバント、生分解性マイクロスフェア(例えば、ポリ乳酸ガラクトド)およびリポソーム(この中に化合物が組み込まれる; 例えば、Fullerton、米国特許第4,235,877号を参照のこと)が挙げられる。ワクチン調製物は、一般的に、例えば、PowellおよびNewman(編)、Vaccine Design (the subunit and adjuvant approach) (1995)に記載される。本発明の範囲内の薬学的組成物およびワクチンはまた、他の化合物に含まれ得、この化合物は、生物学的に活性であっても不活性であってもよい。例えば、他の腫瘍抗原の1つ以上の免疫原性部分が、融合ポリペプチドに組み込まれてか、または別個の化合物としてのいずれかで、この組成物またはワクチン中に存在し得る。

【0199】

例示的なワクチンは、上記のようなポリペプチドのうちの1つ以上をコードするDNAを含み得、その結果、このポリペプチドは、インサイチュで産生される。上記のように、このDNAは、当業者に公知の種々の送達系(核酸発現系、細菌およびウイルス発現系を含む)のいずれかにおいて存在し得る。多数の遺伝子送達技術が当該分野において周知であり、例えば、Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198 (1998)およびそこに引用される参考文献に記載される技術がある。適切な核酸発現系は、患者における発現のために必要なDNA配列(例えば、適切なプロモーターおよび終止シグナル)を含む。細菌送達系は、その細胞表面上でこのポリペプチドの免疫原性部分を発現するか、またはこのようなエピソードを分泌する細菌(例えば、Bacillus-Calmette-Guerrin)の投与を含む。好ましい実施形態においては、DNAは、ウイルス発現系(例えば、ワクシニアウイルスまたは他のポックスウイルス、レトロウイルス、またはアデノウイルス)を用いて導入され得、これは非病原性(欠損)の複製コンピテントなウイルスの使用を含み得る。適切な系は、例えば、以下に開示される: Fisher-Hochら、Proc. Natl. Acad. Sci. USA 86:317-321 (1989); Flexnerら、Ann. N.Y. Acad. Sci. 569:86-103 (1989); Flexnerら、Vaccine 8:17-21 (1990); 米国特許第4,603,112号、同第4,769,330号、および同第5,017,487号; WO89/01973; 米国特許第4,777,127号; GB2,200,651; EP0,345,242; WO91/02805; Berkner, Biotechniques 6:616-627 (1988); Rosenfeldら、Science 252:431-434 (1991); Kollsら、Proc. Natl. Acad. Sci. USA 91:215-219 (1994); Kass-Eislerら、Proc. Natl. Acad. Sci. USA 90:11498-11502 (1993); Guzmanら、Circulation 88:2838-2848 (1993); および Guzmanら、Cir. Res. 73:1202-1207 (1993)。DNAをこのような発現系に組み込むための技術は、当業者に周知である。DNAはまた、例えば、Ulmerら、Science 259:1745-1749 (1993)に記載され、そしてCohen, Science 259:1691-1692 (1993)により概説されるように、「裸(naked)」であり得る。裸のDNAの取込みは、このDNAを生

10

20

30

40

50

分解性ビーズにコーティングすることによって増加され得、これは、細胞中に効率的に輸送される。ワクチンは、ポリヌクレオチド成分およびポリペプチド成分の両方を含み得ることは明らかである。このようなワクチンは、増強された免疫応答を提供し得る。

【0200】

ワクチンは、本明細書中で提供されるポリヌクレオチドおよびポリペプチドの薬学的に受容可能な塩を含み得ることが明らかである。このような塩は、薬学的に受容可能な非毒性の塩基（有機塩基（例えば、一級アミン、二級アミンおよび三級アミンならびに塩基性アミノ酸の塩））および無機塩基（例えば、ナトリウム塩、カリウム塩、リチウム塩、アンモニウム塩、カルシウム塩およびマグネシウム塩）を含む）から調製され得る。

【0201】

当業者に公知の任意の適切なキャリアは、本発明のワクチン組成物において使用され得るが、キャリアの型は、投与の様式に依存して変化する。本発明の組成物は、任意の適切な投与の様式（例えば、局所投与、経口投与、鼻腔内投与、静脈内投与、頭蓋内投与、腹腔内投与、皮下投与または筋内投与を含む）のために処方され得る。非経口投与（例えば、皮下注射）のためには、キャリアは、好ましくは、水、生理食塩水、アルコール、脂肪、ロウまたは緩衝液を含む。経口投与のためには、上記のキャリアまたは固体キャリア（例えば、マンニトール、ラクトース、スターチ、ステアリン酸マグネシウム、サッカリンナトリウム、滑石粉、セルロース、グルコース、スクロースおよび炭酸マグネシウム）が使用され得る。生分解性マイクロスフェア（例えば、ポリ乳酸ポリグリコレート）はまた、本発明の薬学的組成物のためのキャリアとして使用され得る。適切な生分解性マイクロスフェアは、例えば、米国特許第4,897,268号；同第5,075,109号；同第5,928,647号；同第5,811,128号；同第5,820,883号；同第8,853,763号；同第5,814,344号；および同第5,942,252号に開示される。米国特許第5,928,647号に記載される粒子状タンパク質複合体を含むキャリアもまた使用され得、これは、宿主においてクラスI-制限細胞障害性Tリンパ球応答を誘導し得る。

【0202】

このような組成物はまた、緩衝液（例えば、中性緩衝化生理食塩水またはリン酸緩衝化生理食塩水）、炭水化物（例えば、グルコース、マンノース、スクロースまたはデキストラン）、マンニトール、タンパク質、ポリペプチドまたはアミノ酸（例えば、グリシン）、抗酸化剤、静菌剤、キレート剤（例えば、EDTAまたはグルタチオン）、アジュバント（例えば、水酸化アルミニウム）、この処方物を、レシピエントの血液と等張、低浸透圧または弱高浸透圧にする溶質、懸濁剤、濃化剤および/または保存剤を含み得る。あるいは、本発明の組成物は、凍結乾燥剤として処方され得る。化合物はまた、周知の技術を用いてリポソームにカプセル化され得る。

【0203】

任意の種々の免疫賦活剤が、本発明のワクチンに使用され得る。例えば、アジュバントが含まれ得る。ほとんどのアジュバントは、抗原を急速な異化作用から保護するように設計された物質（例えば、水酸化アルミニウムまたは鉱油）、および免疫応答の刺激因子（例えば、リポドA、*Bordetella pertussis*または*Mycobacterium*属または*Mycobacterium*誘導タンパク質）を含む。例えば、脱脂された、脱糖脂質化されたM.vaccinae（「pVac」）が使用され得る。別の実施形態においては、BCGがアジュバントとして使用され得る。さらに、このワクチンは、予めBCGに暴露された被験体に投与され得る。適切なアジュバントは、例えば、以下として市販されている：フロイント不完全アジュバントおよびフロイント完全アジュバント（Difco Laboratories, Detroit, MI）；Merck Adjuvant 65（Merck and Company, Inc., Rahway, NJ）；AS-2およびその誘導体（SmithKline Beecham, Philadelphia, PA）；CWS、TDM、Leif、水酸化アルミニウムゲル（ミョウバン）またはリン酸アルミニウムのようなアルミニウム塩；カルシウム、鉄、または亜鉛

10

20

30

40

50

の塩；アシル化チロシンの不溶性懸濁液；アシル化された糖類；カチオン性またはアニオン性に誘導体化された多糖類；ポリホスファゼン；生分解性マイクロスフェア；モノホスホリルリピドAおよびクイルA。サイトカイン（例えば、GM-CSFまたはインターロイキン-2、インターロイキン-7もしくはインターロイキン-12もまた、アジュバントとして使用され得る。

#### 【0204】

本明細書中で提供されるワクチンにおいて、アジュバント組成物は、好ましくは、主にTh1型の免疫応答を誘導するように設計される。高レベルのTh1型サイトカイン（例えば、IFN-、TNF、IL-2およびIL-12）は、投与された抗原に対する細胞媒介性免疫応答の誘導を助ける傾向がある。対照的に、高レベルのTh2型サイトカイン（例えば、IL-4、IL-5、IL-6およびIL-10）は、体液性免疫応答の誘導を助ける傾向がある。本明細書中で提供されるようなワクチンの適用後に、患者は、Th1型およびTh2型応答を含む免疫応答を支持する。応答が主にTh1型である好ましい実施形態においては、Th1型サイトカインのレベルは、Th2型サイトカインのレベルを大いに超えるまで増加する。これらのサイトカインのレベルは、標準的なアッセイを用いて容易に評価され得る。サイトカインのファミリーの概説については、MosmannおよびCoffman, Ann. Rev. Immunol. 7: 145-173 (1989)を参照のこと。

10

#### 【0205】

Th1型応答を主に誘発する際の使用のための好ましいアジュバントとしては、例えば、モノホスホリルリピドA、好ましくは3-脱-O-アシル化モノホスホリルリピドA（3D-MPL）とアルミニウム塩との組合せが挙げられる。MPLアジュバントは、Corixa Corporation (Seattle, WA；米国特許第4,436,727号；同第4,877,611号；同第4,866,034号；および同第4,912,094号を参照のこと)より入手可能である。CpG含有オリゴヌクレオチド（ここで、CpGジヌクレオチドはメチル化されていない）はまた、主にTh1型応答を誘導する。このようなオリゴヌクレオチドは周知であり、例えば、WO96/02555、WO99/33488ならびに米国特許第6,008,200号および同第5,856,462号に記載される。免疫刺激性DNA配列はまた、例えば、Satoら、Science 273: 352 (1996)により記載される。別の好ましいアジュバントは、サポニン（例えば、Quil Aまたはその誘導体（QS21およびQS7 (Aquila Biopharmaceuticals Inc., Framingham, MA)を含む）；エスシン (Escin)；ジギトニン；あるいはGypsophilaまたはChenopodium quinoaサポニン）を含む。他の好ましい処方物は、本発明のアジュバントの組合せにおける1つよりも多いサポニン（例えば、QS21、QS7、Quil A、エスシンまたはジギトニンからなる群の少なくとも2つの組合せ）を含む。

20

30

#### 【0206】

あるいは、サポニン処方物は、キトサンまたは他のポリカチオンポリマー、ポリ乳酸およびポリ乳酸-co-グリコシド粒子、ポリ-N-アセチルグルコサミンベースのポリマーマトリックス、多糖類または化学的に改変された多糖類から構成される粒子、リポソームおよび脂質ベースの粒子、グリセロールモノエステルにより構成される粒子などにより構成されるワクチンピヒクルと組合せられ得る。サポニンはまた、コレステロールの存在下で処方されて、リポソームまたはISCOMのような粒子構造を形成し得る。さらに、サポニンは、非粒子状溶液もしくは懸濁液中、または小層状 (paucilamellar) リポソームまたはISCOMのような粒子構造中のいずれかで、ポリオキシエチレンエーテルまたはポリオキシエチレンエステルとともに処方され得る。サポニンはまた、Carbopol (登録商標)のような賦形剤とともに処方されて粘度を増加させ得るか、またはラクトースのような粉末賦形剤とともに乾燥粉末形態で処方され得る。

40

#### 【0207】

1つの好ましい実施形態においては、アジュバント系は、モノホスホリルリピドAとサポ

50

ニン誘導体との組合せ（例えば、WO94/00153に記載されるようなQS21と3D-MPL（登録商標）アジュバント）との組合せ、またはWO96/33739に記載されるような、QS21がコレステロールでクエンチされたより反応発生性（reactogenic）の低い組成物）を含む。他の好ましい処方物は、水中油エマルジョンおよびトコフェロールを含む。水中油エマルジョン中のQS21、3D-MPL（登録商標）アジュバントおよびトコフェロールを使用する、別の特に好ましいアジュバント処方物は、WO95/17210に記載される。

【0208】

別の増強されたアジュバント系は、CpG含有オリゴヌクレオチドとサポニン誘導体との組合せ（特に、WO00/09159に記載されるようなCpGとQS21との組合せ）を含む。好ましくは、この処方物はさらに、水中油エマルジョンおよびトコフェロールを含む。

10

【0209】

他の好ましいアジュバントとしては、以下が挙げられる：Montanide ISA 720（Seppic, France）、SAF（Chiron, California, United States）、ISCOMS（CSL）、MF-59（Chiron）、アジュバントのSBASシリーズ（例えば、SBAS-2、AS2'、AS2''、SBAS-4またはSBAS6（SmithKline Beecham, Rixensart, Belgiumより入手可能））、Detox（Corixa, Hamilton, MT）、RC-529（Corixa, Hamilton, MT）、および係属中の米国特許出願第08/853,826号および同第09/074,720号（これらの開示は、本明細書中にその全体が参考として援用される）に記載されるような、他のアミノアルキルグルコサミニド4-ホスフェート（AGP）、およびWO99/52549 A1に記載されるような、ポリオキシエチレンエーテルアジュバント。

20

【0210】

他の好ましいアジュバントとしては、一般式（I）： $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{-A-R}$ のアジュバント分子が挙げられ、ここでnは、1~50であり、Aは、結合または-C(O)-であり、Rは、 $\text{C}_{1-50}$ アルキルまたはフェニル $\text{C}_{1-50}$ アルキルである。

【0211】

本発明の1つの実施形態は、nが、1と50との間、好ましくは4~24、最も好ましくは9であり；R成分が、 $\text{C}_{1-50}$ 、好ましくは $\text{C}_4\text{-C}_{20}$ アルキル、最も好ましくは $\text{C}_{12}$ アルキルであり、そしてAが結合である、一般式（I）のポリオキシエチレンエーテルを含むワクチン処方物からなる。ポリオキシエチレンエーテルの濃度は、0.1~20%の範囲、好ましくは0.1~10%、最も好ましくは0.1~1%の範囲であるべきである。好ましいポリオキシエチレンエーテルは、以下の群より選択される：ポリオキシエチレン-9-ラウリルエーテル、ポリオキシエチレン-9-ステオリルエーテル、ポリオキシエチレン-8-ステオリルエーテル、ポリオキシエチレン-4-ラウリルエーテル、ポリオキシエチレン-35-ラウリルエーテル、およびポリオキシエチレン-23-ラウリルエーテル。ポリオキシエチレンラウリルエーテルのようなポリオキシエチレンエーテルは、Merck index（第12版：エントリー7717）に記載される。これらのアジュバント分子は、WO99/52549に記載される。

30

40

【0212】

上記の一般式（I）に従うポリオキシエチレンエーテルは、所望の場合、別のアジュバントと組合せ得る。例えば、好ましいアジュバントの組合せは、好ましくは、係属中のUK特許出願GB9820956.2に記載されるようなCpGとの組合せである。

【0213】

本明細書中で提供される任意のワクチンは、抗原、免疫応答エンハンサーおよび適切なキャリアまたは賦形剤の組合せを生じる周知の方法を使用して調製され得る。本明細書中に記載される組成物は、持続放出处方物（すなわち、投与後に化合物の緩徐な放出を行うカプセル、スポンジまたはゲル（例えば、多糖類で構成される）のような処方物）の一部と

50

して投与され得る。このような処方物は、一般的に、周知の技術（例えば、C o o m b e sら、V a c c i n e 14:1429-1438(1996)を参照のこと）を用いて調製され得、そして例えば、経口、直腸もしくは皮下移植により、または所望の標的部位への移植により投与され得る。持続放出处方物は、キャリアマトリックス中に分散され、そして/または速度制御膜に囲まれたレザバに含まれた、ポリペプチド、ポリヌクレオチドまたは抗体を含み得る。

#### 【0214】

このような処方物における使用のためのキャリアは生体適合性であり、そしてまた生分解性であり得；好ましくは、この処方物は、比較的一定のレベルの活性成分の放出を提供する。このようなキャリアとしては、ポリ(ラクチド-*c o*-グリコリド)、ポリアクリレート、ラテックス、スターチ、セルロース、デキストランなどの微粒子が挙げられる。他の遅延放出キャリアとしては、超分子バイオベクターが挙げられ、これは、非液体親水性コア（例えば、架橋多糖またはオリゴ糖）、および必要に応じて、両親媒性化合物（例えば、リン脂質）を含む外部層を含む（例えば、米国特許第5,151,254号およびPCT出願WO94/20078、WO94/23701およびWO96/06638を参照のこと）。持続放出处方物に含まれる活性化合物の量は、移植の部位、放出の速度および予測される持続時間ならびに処置されるかまたは予防されるべき状態の性質に依存する。

10

#### 【0215】

任意の種々の送達ビヒクルが、薬学的組成物およびワクチンにおいて使用され、腫瘍細胞を標的化する抗原特異的免疫応答の生成を容易にし得る。送達ビヒクルとしては、抗原提示細胞（APC）（例えば、樹状細胞、マクロファージ、B細胞、単球および有効なAPCであるように操作され得る他の細胞）が挙げられる。このような細胞は、必須ではないが、抗原を提示する能力を増加させるため、T細胞応答の活性化および/もしくは維持を向上させるため、それ自体で抗腫瘍効果を有するように、ならびに/またはレシーバーと免疫学的に適合性である（すなわち、一致したHLAハプロタイプ）であるように、遺伝的に改変され得る。APCは、一般的に、任意の種々の生物学的流体および器官（腫瘍および腫瘍周囲の組織を含む）から単離され得、そして自己細胞、同種異系細胞、同系細胞または外因性細胞であり得る。

20

#### 【0216】

本発明の特定の好ましい実施形態は、抗原提示細胞として樹状細胞またはその前駆細胞を用いる。樹状細胞は、非常に強力なAPCであり（BanchereauおよびSteinman, Nature 392:245~251(1998)）、そして予防的または治療的な抗腫瘍免疫を惹起するための生理学的なアジュバントとして有効であることが示されている（TimmermanおよびLevy, Ann. Rev. Med. 50:507~529(1999)を参照のこと）。一般には、樹状細胞は、その典型的な形状（インサイチュで星状（インピトロで可視の顕著な細胞質突起（樹状突起）を有する））、その取り込み能、高い効率のプロセッシングおよび抗原提示、ならびにナイーブなT細胞応答を活性化する能力に基いて同定され得る。当然ながら、樹状細胞は、インピボまたはエキソピボで樹状細胞に通常見出されない特定の細胞表面レセプターまたはリガンドを発現するように操作され得、そしてこのような改変された樹状細胞が、本発明によって意図される。樹状細胞の代替として、分泌型小胞抗原負荷樹状細胞（secreted vesicles antigen-loaded dendritic cells）（エキソソーム（exosome）と呼ばれる）をワクチン内で用い得る（Zitvogelら、Nature Med. 4:594~600(1998)）。

30

40

#### 【0217】

樹状細胞および前駆体は、末梢血、骨髄、腫瘍浸潤細胞、腫瘍周囲組織浸潤細胞、リンパ節、脾臓、皮膚、臍帯血、または任意の他の適切な組織もしくは体液から得ることができ。例えば、樹状細胞は、GM-CSF、IL-4、IL-13および/またはTNFのようなサイトカインの組合せを、末梢血から収集した単球の培養物に添加することによ

50

り、エキソピボで分化し得る。あるいは、末梢血、臍帯血、または骨髄から収集したCD34陽性細胞は、GM-CSF、IL-3、TNF、CD40リガンド、LPS、flt3リガンド、および/または他の化合物(樹状細胞の分化、成熟、および増殖を誘導する)の組み合わせを、培養培地に添加することにより、樹状細胞に分化し得る。

#### 【0218】

樹状細胞は、2つの十分特徴付けされた表現型の間の識別を簡単に行うため、「未成熟」細胞、および「成熟」細胞と、簡便に分類される。しかし、この命名法は、全ての可能性のある中間段階の分化を除くと解釈されるべきではない。未成熟の樹状細胞は、高い能力の抗原取り込みおよびプロセッシングを有するAPCとして特徴付けられる。この細胞は、Fcレセプターおよびマンノースレセプターの高い発現に関連する。成熟表現型は、代表的にこれらのマーカーの発現がより低い、T細胞活性化を担う細胞表面分子(例えば、クラスIおよびクラスII MHC、接着分子(例えば、CD54およびCD11)ならびに同時刺激分子(例えば、CD40、CD80、CD86および4-1BB))の発現が高いことにより特徴付けられる。

#### 【0219】

APCは、一般にタンパク質(またはその部分もしくは他の改変体)をコードするポリヌクレオチドでトランスフェクトされ得、その結果、この細胞の表面上でこのポリペプチドまたはその免疫原性部分を発現する。このようなトランスフェクションは、エキソピボで生じ得、次いで、このようなトランスフェクトされた細胞を含む組成物またはワクチンは、本明細書に記載のように、治療目的に用いられ得る。あるいは、樹状細胞または他の抗原提示細胞を標的化する遺伝子送達ビヒクルが、患者に投与され得、インピボで生じるトランスフェクションを生じる。例えば、インピトロおよびエキソピボでの樹状細胞のトランスフェクションは、一般的に、当該分野で公知の任意の方法(例えば、WO97/24447に記載のような方法)、またはMahviら、*Immunology and Cell Biology* 75:456~460(1997)に記載の遺伝子銃アプローチを用いて実施され得る。樹状細胞の抗原ローディングは、ポリペプチド、DNA(裸のDNAまたはプラスミドベクター中のDNA)、またはRNA;あるいは抗原を発現する組換え細菌もしくはウイルス(例えば、ワクシニア、鶏痘、アデノウイルスまたはレンチウイルスベクター)とともに樹状細胞または前駆細胞をインキュベートすることによって達成され得る。ローディングの前、このポリペプチドは、T細胞の助けを提供する免疫学的パートナー(例えば、キャリア分子)と共有結合的に結合体化され得る。あるいは、樹状細胞は、結合体化されない免疫学的パートナーで、別々にまたはポリペプチドの存在下で、パルスされ得る。

#### 【0220】

ワクチンおよび薬学的組成物は、密閉アンプルまたはバイアルのような、単位用量または複数用量の容器で提供され得る。このような容器は、好ましくは、密閉してシールされ、使用時までその処方物の滅菌性を保持する。一般に、処方物は、懸濁液、溶液、または油性もしくは水性ビヒクル中のエマルジョンとして貯蔵され得る。あるいは、ワクチンまたは薬学的組成物は、使用の直前に滅菌液体キャリアの添加しか要さない凍結乾燥状態で貯蔵され得る。

#### 【0221】

(診断キット)

本発明は上記の任意の診断方法における使用のためのキットをさらに提供する。このようなキットは、代表的に、診断アッセイを実施するのに必要な2つ以上の成分を含む。成分は、化合物、試薬、容器、および/または備品であり得る。例えば、キット内の1つの容器は、タンパク質に特異的に結合する、モノクローナル抗体またはそのフラグメントを含み得る。このような抗体またはフラグメントは、上記のような、支持材料に接着され得る。1つ以上のさらなる容器が、このアッセイで用いられるべき、試薬または緩衝液のような、要素を含み得る。このようなキットはまた、あるいは、上記のような検出試薬を含み、この試薬は、抗体結合の直接検出または間接検出に適切なレポーター基を含む。

## 【 0 2 2 2 】

あるいは、キットは、生物学的サンプル中のタンパク質をコードする mRNA のレベルを検出するように設計され得る。このようなキットは一般に、上記のように、タンパク質をコードするポリヌクレオチドにハイブリダイズする、少なくとも 1 つのオリゴヌクレオチドプローブまたはプライマーを含む。このようなオリゴヌクレオチドは、例えば、PCR またはハイブリダイゼーションアッセイにおいて用いられ得る。このようなキット中に存在し得るさらなる成分としては、本発明のタンパク質をコードするポリヌクレオチドの検出を容易にするための第二オリゴヌクレオチド、および/または診断試薬または容器が挙げられる。

## 【 0 2 2 3 】

本明細書に引用される全ての刊行物および特許出願は、各々の個々の刊行物または特許出願が、参考として援用されることが特別にかつ個々に示されているように、本明細書において参考として援用される。

## 【 0 2 2 4 】

前述の発明は、ある程度詳細に、例示のためにそして理解の明確化のために記載してきたが、本発明の教示に照らして、ある変化および改変が、添付の特許請求の範囲の精神および範囲から逸脱することなく本明細書においてなされ得ることが当業者に容易に明白である。

## 【 0 2 2 5 】

(実施例)

以下の実施例は、限定の目的ではなく例示の目的のみで示される。当業者は、本質的に類似の結果を得るために、変更または改変され得る様々な重要ではないパラメーターを容易に認識する。

## 【 0 2 2 6 】

(実施例 1 : M T B 7 2 F 融合タンパク質および個々の抗原を含む組成物を用いるモルモットのワクチン接種)

モルモットを、アジュバントのみ ( S B A S 1、S B A S 2 または A S A S 7 および A 1 ( O H ) 3 )、アジュバント中の M T B 7 2 F 融合タンパク質、または T b H 9 および R a 3 6 抗原組成物で免疫した。

## 【 0 2 2 7 】

方法 :

群 : 1 ) S B A S 1

2 ) S B A S 2

3 ) S B A S 7 + A 1 ( O H ) 3

4 ) T b H 9 + R a 3 5 + S B A S 1

5 ) T b H 9 + R a 3 5 + S B A S 2

6 ) T b H 9 + R a 3 5 + S B A S 7 ( A 1 ( O H ) 3 )

7 ) M T B 7 2 F ( S B A S 1 中 )

8 ) M T B 7 2 F ( S B A S 2 中 )

9 ) M T B 7 2 F ( S B A S 7 + A 1 ( O H ) 3 中 )

1 0 ) P B S

1 1 ) B C G。

## 【 0 2 2 8 】

投薬量 : それぞれ 4 μ g の T b H 9 および R a 3 5

8 μ g の M T B 7 2 F

プロトコル : 一回目の免疫、約 3 週間後の二回目の免疫、約 2 週間半後の三回目の免疫。

## 【 0 2 2 9 】

プレチャレンジ : D T H ( 遅延型過敏性、抗原性を測定するために使用 ; 1 0 μ g の抗原 )

チャレンジ : 約 3 0 c f u の E r d m a n 株を含むエアロゾル

10

20

30

40

50



チャレンジ後のモニタリング：体重の減少  
死亡（チャレンジ後約6ヶ月）。

結果：

1. D T H

免疫抗原に対するポジティブな反応。個々の抗原または融合タンパク質に対する反応は、比較可能であった。P P Dに対する皮膚試験反応性は、B C Gで免疫した群を用いた場合にのみ観察された。

【0230】

2. 保護：M T B 7 2 F 融合タンパク質でワクチン接種したモルモットは、抗原の混合物で免疫したモルモットと比較して、保護を与えた（図1を参照のこと）。

10

【0231】

（実施例2：M T B 7 2 F 融合タンパク質および個々の抗原を含む組成物を用いるマウスのワクチン接種）

上記のように、マウスを、アジュバントのみ（S B A S 2、S B A S 2'、S B A S 2'' またはA S A S 6）、アジュバント中のM T B 7 2 F 融合タンパク質、M T B 7 2 F D N A、アジュバント中のM T B 5 9 F 融合タンパク質、またはT b H 9、R a 3 5、およびR a 1 2 抗原組成物で免疫した。

【0232】

方法：

群：1) M T B 7 2 F + S B A S 2

20

2) M T B 7 2 F + S B A S 2'

3) M T B 7 2 F + S B A S 2''

4) M T B 7 2 F + S B A S 6

5) R a 1 2 + T b H 9 + R a 3 5 ( S B A S 2 中 )

6) M T B 5 9 F ( S B A S 2 中 )

7) S B A S 2

8) M T B 7 2 F + 脱脂、脱糖脂質化した M . v a c c a e

9) M T B 7 2 F D N A

10) M T B 7 2 F + I F A

11) M T B 7 2 F + B C G

30

12) 脱脂、脱糖脂質化した M . v a c c a e

13) B C G

14) 生理食塩水

15) M T B 7 2 F + S B A S 2 ( 実験室内処方 )。

【0233】

1群あたり8匹の動物。

【0234】

免疫スケジュール：一回目の免疫、約3週間後の二回目の免疫、約3週間後の三回目の免疫。

【0235】

40

一回目の投薬の約3ヶ月後にエアロゾルチャレンジ。

【0236】

脾臓細胞または肺細胞を単離し、そして培養した；プレーティングの約3週間後に培養物中のC F Uをカウントした。

【0237】

用量：8 μ g の M T B 7 2 F、6 . 5 6 μ g の M T B 5 9 F、またはそれぞれ、1 . 5 2、4 . 3 および 2 . 2 4 μ g の R a 1 2、T b H 9 および R a 3 5 ( 混合 )

結果：

A S アジュバントのうち、A S 2'' + M B 7 2 F は、このセットの実験における脾臓および肺の両方において最も良好な保護を示した（図2 A および 2 B を参照のこと）。M T

50

B 7 2 F は、このセットの実験における脾臓および肺の両方において、M T B 5 9 F よりも約 1 l o g 良好な保護を示した。このことは、R a 1 2 がさらなる利点を提供することを示す。1 2 / H 9 / 3 5 + A S 2 の混合物は、この実験において M T B 7 2 F よりも良好な保護を示した。M T B 7 2 F D N A は、この実験において、特に脾臓において最も良好な保護を示した (> 2 l o g)。肺における保護は、この実験において、M T B 7 2 F タンパク質 + A S 2 ' ' を用いた場合に観察された保護に匹敵した。

【 0 2 3 8 】

(実施例 3 : M T B 7 2 F 融合タンパク質および個々の抗原を含む組成物を用いるモルモットのワクチン接種)

上記のように、モルモットを、アジュバントのみ ( S B A S 2、S B A S 2 '、S B A S 2 ' ' または A S A S 6 )、アジュバント中の M T B 7 2 F 融合タンパク質、M T B 7 2 F D N A、アジュバント中の M T B 5 9 F 融合タンパク質、または T b H 9、R a 3 5 および R a 1 2 抗原組成物で免疫した。

10

【 0 2 3 9 】

方法 :

群 : 1 ) M T B 7 2 F + S B A S 2

2 ) M T B 7 2 F + S B A S 2 '、

3 ) M T B 7 2 F + S B A S 2 ' '、

4 ) M T B 7 2 F + S B A S 6

5 ) R a 1 2 + T b H 9 + R a 3 5 ( S B A S 2 中)

20

6 ) M T B 5 9 F ( S B A S 2 中)

7 ) S B A S 2

8 ) M T B 7 2 F + p v a c

9 ) M T B 7 2 F D N A

1 0 ) M T B 7 2 F + I F A

1 1 ) M T B 7 2 F + B C G

1 2 ) B C G

1 3 ) 生理食塩水

1 4 ) 脱脂、脱糖脂質化した M . v a c c a e

抗原 : 抗原は、1 モル当量で処方した。

30

【 0 2 4 0 】

1 群あたり 5 匹の動物。

【 0 2 4 1 】

1 投薬あたりの注射容量は、2 5 0 μ l ( I M ) であり、以下を含む :

M T B 7 2 F 2 0 μ g

R a 1 2、T B H 9、R a 3 5 3 . 8、1 0 . 8 および 5 . 6 μ g

M T B 5 9 F 1 6 . 4 μ g

スケジュール : 一回目の免疫、約 3 週間後の二回目の免疫、約 3 週間後の三回目の免疫。

【 0 2 4 2 】

チャレンジ : 最初の免疫から約 1 ヶ月半後

40

結果 :

チャレンジの約 3 8 週間後

【 0 2 4 3 】

【 表 1 】

<u>群</u>	<u>生存</u>	<u>状態</u>	
G1. MTB72F + AS2	1/5	[体重が減少]	
G2. MTB72F + AS2'	2/5	[体重の増加なし]	
G3. MTB72F + AS2''	3/5	[問題なく見えるが、体重の増加なし]	
G4. MTB72F + AS6	2/5	[これら両方とも体重が増加]	
G5. MTBRa12+H9+Ra35 +AS2	4/5	[1匹は少し衰弱しているようであるが、2匹は体重が増加している]	
G6. MTB59F + AS2	2/5	[両方とも少し体重が減少]	
G7. AS2	2/5	[両方とも体重が減少]	10
G8. MTB72F + pVac	1/5	[あまり良好には見えない]	
G9. MTB72F DNA	3/5	[全て安定に維持している]	
G10. MTB72F + IFA	2/5	[問題なし]	
G11. MTB72F + BCG	5/5	[非常によく食べる]	
G12 BCG	4/5	[良好]	
G13 生理食塩水	全て死亡		
G14 pVac	2/5	[体重の増加なし]	20

チャレンジの50週間後までに、BCG + Mt b 7 2 Fで免疫したモルモットのうちの80% (5匹のうち4匹) がなお生存していたが、BCG単独で免疫したモルモットのうち20% (5匹のうち1匹) のみが生存していた。85週目には、BCG + Mt b 7 2 Fで免疫した5匹のモルモットのうち4匹が、なお生存し、かつ健康であった(図7を参照のこと)。

【0244】

(実施例4: 長期保護)

上記のように、モルモットを、アジュバント単独(AS2またはAS2'')、アジュバント中のMTB72F融合タンパク質、TbH9、Ra35およびRa12抗原組成物、またはアジュバント中の種々の個々の抗原で免疫した。

【0245】

(方法)

【0246】

【表2】

<u>群</u>	<u>抗原用量</u>	
1. AS2”+MTB39 (TbH9)	20ug/250ul (IM)	
2. AS2”+MTB8.4 (DPV)	20ug	
3. AS2”+MTB9.9 (MTI)	20ug	
4. AS2”+MTB41 (MTCC#2)	20ug	
5. AS2”+MTB40 (HTCC#1)	20ug	
6. AS2”+MTB9.8 (MSL)	20ug	10
7. AS2”+MTB72F	20ug	
8. AS2”+Ra12+TbH9+Ra35 (モル当量)	3.8 µg +10.8 µg +5.6 µg	
9. AS2”+MTB71F+MTB72F+HTCC#1	20 µg +20 µg +10 µg	
10. AS2”+Ra12	20 µg	
11. BCG		
12. AS2”		
13. AS2+MTB72F		20
14. AS2+Ra12+TbH9+Ra35		
15. AS2		

(実施例5：MTB72F融合タンパク質および個々の抗原を含む組成物でのサルのワクチン接種)

上記のように、サルを、S B A S 2 アジュバント中のMTB72F融合タンパク質、またはアジュバント中のMTB8.4抗原組成物、またはMTB72FおよびMTB8.4の混合物で免疫した。

【0247】

方法：

群

1. 生理食塩水
2. B C G
3. M T B 8 . 4 / A S 2
4. M T B 7 2 F / A S 2
5. M T B 7 2 F / A S 2 (一方の腕) + M T B 8 . 4 / A S 2 (他方の腕)

各抗原40 µg。

【0248】

結果：

チャレンジの8週間後に、BCGで免疫したサルは、感染の徴候を見せている。

【0249】

チャレンジの16週間後についての現在のデータによって、以下の傾向が明らかになっている：

MTB72Fで免疫した群(4および5)は、体重を維持し、群3(MTB8.4免疫)と比較して低いESR値を有する(表1および2)。

【0250】

【表3】

30

40

表1  
 チャレンジ20週間後の、AS2中に処方したMTB8.4および  
 MTB72Fでのカニクイザルにおける予防ワクチン研究

群	ID	正味の体重 変化(kg)	胸部X線(発症)	状態	
AS2	1398K	-24%	Pn, bil, prog (8週間)	生存	10
	4437B	-33%	Pn, bil, prog (4週間)	死亡	
	2959G	-8.30%	Pn, bil, prog (4週間)	生存	
	605AE	-14.00%	Pn, rt, 安定 (8週間)	生存	
BCG	3436A	-15.00%	Neg	生存	20
	3642G	+ 4.5%	Pn, rt, prog (8週間)	生存	
	1190H	0%	Neg	生存	
	1051I	-30%	Pn, rt, prog (8週間)	死亡	
MTB8.4	3665C	-25%	Pn, rt, prog (8週間)	死亡	
	2200F	-18.00%	Pn, rt, 安定 (8週間)	生存	
	1654J	-33.00%	Pn, bil, prog (4週間)	死亡	
	4141C	-33%	Pn, bil, prog (4週間)	死亡	
MTB72F	3061C*	ITチャレンジ後に死亡			30
	1228G	+ 3.6%	Bron, bil, 3mo安定 (8週間)	生存	
	3462E	-2.20%	Neg	生存	
MTB8.4	4254C	+ 1.21	Pn, rt, 3mo安定 (4週間)	生存	
	4496A	+ 7%	Pn, rt, 1mo安定 (8週間)	生存	
	4422C	-39.00%	Pn, bil, prog (4週間)	死亡	
MTB72F	4416A	+ 11%	Pn, rt, 2mo安定 (12週間)	生存	
	2734E	+ 12.5%	Susp infil rt, 3mo安定 (8週間)	生存	
【 0 2 5 1 】					40
【 表 4 】					

表2

AS2中に処方したMTB8.4およびMTB72Fを用いた  
カニクイザルにおける予防ワクチン研究

群	ID	チャレンジ後の経過週				16週間 胸部X線	
		4	8	12	16		
AS2	1398K	3	3	10	19	Pn, bil, progrsv	10
	4437B	10	20	3		死亡	
	2959G	6	3	3	0	Pn, rt, progrsv	
	605AE	1	4	7	3	Pn, rt, 安定	
BCG	3436A	0	8	7	15	Neg	20
	3642G	0	0	0	0	Pn, rt, progrsv	
	1190H	1	0	2	0	Neg	
	1051I	0	8	22	7	Pn, bil, w/furt pro, 死亡	
MTB8.4	3665C	12	30	19		死亡	30
	2200F	1	7	2	0	Pn, rt, progrsv	
	1654J	20	8	21	7	Pn, bil, w/fur progr	
	4141C	13	8	2	15	Pn, bil, w/fur progr	
MTB72F	3061C*	ITチャレンジ後に死亡					30
	1228G	0	1	20	0	現在安定	
	3462E	0	0	0	0	Neg	
MTB8.4/ MTB72F	4254C	13	0	0	0	Pn, 現在安定	40
	4496A	5	1	0	5	Pn, rt, w/furt prog	
	4422C	10	3	0		死亡	
	4416A	6	0	1	0	Pn, 現在安定	
	2734E	0	0	0	0	Susp infil, 現在安定	

(実施例6:サルにおけるBCGプライミング実験)

4つの群(5匹/群)をBCGで免疫し、休ませ、次いで上記のように免疫し、そしてチャレンジする。以下のプロトコルを使用する:

【0252】

【表5】

群	動物の数	免疫する抗原	抗原用量	
1. なし	5	AS2		
2. BCG	5	AS2		
3. BCG	5	MTB72F	40ug	
4. BCG	4	Ra12+TbH9+Ra35	MTB72F用量中 モル当量の抗原	
5. BCG	4	MTB72F + MTB71F + MTB40	40ug MTB72F 40ug MTB72F 20ug MTB40	10

全ての抗原を、AS2中に処方する。

群4および5は、各4匹の動物を有する。BCG免疫したサルのうち2匹が死亡した。

【0253】

【表6】

群	動物の数	免疫する抗原	T細胞増殖および サイトカイン産生アッセイ についての抗原	
1. なし	5	AS2	PHA, PPD, MTB72F, MTCC#2, Ra12, TbH9, Ra35, MSL, MTI	20
2. BCG	5	AS2	PHA, PPD, MTB72F, MTB71F, HTCC#1, DPV, MTCC#2, Ra12, TbH9, Ra35, MSL, MTI	30
3. BCG	5	MTB72F	PHA, PPD, MTB72F, Ra12, TbH9, Ra35	
4. BCG	4	Ra12+TbH9+Ra35	PHA, PPD, MTB72F, Ra12, TbH9, Ra35	
5. BCG	4	MTB72F + MTB71F + MTB40	PHA, PPD, MTB72F, MTB71F, HTCC#1, DPV, MTCC-2, Ra12, TbH9, Ra35, MSL, MTI	40

(実施例7: Ra35 Mut SAおよびMTB72F Mut SAの構築)

代表的に、Mtb72fの発現は、いくつかの破損した産物を生じする。さらに、E. coliにおけるRa35 (Mtb32A)の成熟または全長形態の全長配列の発現は、困難である。発現された産物は、低レベルのタンパク質発現を示すポリクローナルウサギ抗Ra35Abを用いた免疫プロット後にのみ、可視化できた。その際さえも、組換え抗原

の自己触媒破損（分解）を示す複数の特定の種（バンド）が検出された。これは、生物学的に活性な形態としての E . c o l i における R a 3 5 F L の発現に起因すると推定された。

【 0 2 5 4 】

N 末端の半分（ R a 3 5 N 末端、 R a 3 5 と称される ）および C 末端の半分（ R a 3 5 C 末端、 R a 1 2 と称される ）を含む 2 つの重複する半分のものとして R a 3 5 F L を発現することが可能であったことが、以前に示された。 R a 3 5 分子全体の発現を増強および安定化するために、単一点変異を、 3 つ組活性部位内の残基のうちの 1 つに導入した（ S e r から A l a の置換、図 6 を参照のこと）。そうするとこの変異誘発した形態の M t b 3 2 A を、安定形態で高レベルで簡単に発現させることができた。さらに、 M t b 7 2 F の発現を安定化させるために、単一ヌクレオチド置換（融合タンパク質の 7 1 0 位に S e r から A l a の変更を生じる、 T から G への置換）を、 M t b 7 2 F 配列中のヌクレオチド 2 1 2 8 位に導入した（図 5 を参照のこと）。

10

【 0 2 5 5 】

この安定化はまた、 R a 3 5 F L、 R a 3 5、もしくは M t b 7 2 F または R a 3 5 を含む他の融合タンパク質中の 3 つ組活性部位を含む 3 つの残基（ H i s、 A s p、または S e r ）のうちの任意の 1 つ、任意の 2 つ、または 3 つ全てを変異誘発することによって、容易に達成される。変異誘発は、当業者に公知の任意の技術を使用して実行され得る。

【 0 2 5 6 】

（実施例 8： R a 3 5 F L M u t S A - T b H 9 および M T B 7 2 F M u t S A を用いたマウスの免疫）

20

1 群につき 8 匹のマウスを、以下に列挙する組成物（アジュバント A S 2 A を含む）で免疫した。次いで、このマウスを M y c o b a c t e r i u m t u b e r c u l o s i s でチャレンジし、そしてマウスの生存を測定した。

【 0 2 5 7 】

【表 7】

<u>群</u>	<u>タンパク質またはDNAの濃度</u>
1. Mtb72f タンパク質	1.5 mg/ml
2. Mtb72f DNA	1.2 mg/ml
3. Mtb72f-85b タンパク質	0.6 mg/ml
4. Mtb72f-85b DNA	1.1 mg/ml
5. Mtb72f-MTI タンパク質	1.3 mg/ml
6. Mtb72f-MTI DNA	1.1 mg/ml
7. Mtb72f MutSA タンパク質	1.7 mg/ml
8. MTB3AMutSA-TbH9タンパク質	2.4 mg/ml
9. BCG	
10. AS2	
11. ベクター単独	1.5 mg/ml
12. 生理食塩水	

30

40

【図面の簡単な説明】

【図 1】 図 1 は、 M T B 7 2 F ポリタンパク質をワクチン接種したモルモットのパーセ

50



ント生存率を示す。

【図2】 図2は、MTB72F、MTB59F、MTB72F DNA、またはRa12抗原、TbH9抗原、およびRa35抗原を含む組成物による免疫後の脾臓細胞（図2A）および肺細胞からのCFUを示す。

【図3】 図3は、MTB72Fの概略図を示す。

【図4】 図4は、Ra35（MTB32AのN末端部分から195アミノ酸）のヌクレオチド配列およびアミノ酸配列を示す。

【図5】 図5は、MTB72Fおよび変異版MTB72F MutSAのアミノ酸配列のアラインメントを示す。

【図6】 図6は、成熟（全長）Ra35 / MTB32Aおよび変異版Ra35 FL MutSAのアミノ酸配列のアラインメントを示す。

10

【図7】 図7は、Mt b 72F処方物を用いてワクチン接種したモルモットの長期間の生存率を示す。

【配列表】

SEQUENCE LISTING

(2) INFORMATION FOR SEQ ID NO:1: MTB32A (Ra35 FL)

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1872 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

GACTACGTTG GTGTAGAAAA ATCCTGCCGC CCGGACCCTT AAGGCTGGGA CAATTTCTGA      60
TAGCTACCCC GACACAGGAG GTTACGGGAT GAGCAATTCG CGCCGCGGCT CACTCAGGTG      120
GTCATGGTTG CTGAGCGTGC TGGCTGCCGT CGGGCTGGGC CTGGCCACGG CGCCGGCCCA      180
GGCGGCCCCG CCGGCCTTGT CGCAGGACCG GTTCGCCGAC TTCGCCGCGC TGCCCTCGGA      240
CCCCTCCGCG ATGGTCCGCC AAGTGGCGCC ACAGGTGGTC AACATCAACA CCAAACCTGGG      300
CTACAACAAC GCCGTGGGCG CCGGGACCGG CATCGTCATC GATCCCAACG GTGTCGTGCT      360
GACCAACAAC CACGTGATCG CGGGCGCCAC CGACATCAAT GCGTTCAGCG TCGGCTCCGG      420
CCAAACCTAC GGCCTCGATG TGGTCCGGTA TGACCGCACC CAGGATGTCT CGGTGCTGCA      480
GCTCGCGGTT GCCCGTGGCC TGCCGTGGCC GGCATGGGTG GCGCGGCTCG CGGTGGGTGA      540
GCCCTCCGTC CCGATGGGCA ACAGCGTGGG GCAGGGGGGA ACGCCCGCTG CGGTGCTGCG      600
CAGGGTGGTC GCGCTCGGCC AAACCGTGCA GCGCTCGGAT TCGCTGACCG GTGCCGAAGA      660
GACATTGAAC GGGTTGATCC AGTTCGATGC CGCAATCCAG CCCGGTGATT CCGGCGGGCC      720
CGTCGTCAAC GGCCTAGGAC AGGTGGTCCG TATGAACACG GCCCGCTCCG ATAACTTCCA      780
GCTGTCCCAG GGTGGGCAGG GATTCCGCCAT TCCGATCCGG CAGGCGATGG CGATCGCGGG      840
CCAAATCCGA TCCGGTGGGG GGTCAACCAC CGTTCATATC GGGCCTACCG CCTTCCTCGG      900
CTTGGGTGTT TCCGACAACA ACGGCAACGG CGCACGAGTC CAACGCGTGG TCGGAAGCGC      960
TCCGGCGGCA AGTCTCGGCA TCTCCACCGG CGACGTGATC ACCGCGGTCT ACGGCGCTCC      1020
GATCAACTCG GCCACCGCCA TGGCGGACCG GCTTAACGGG CATCATCCCG GTGACGTCAT      1080
CTCGGTGAAC TGGCAAACA AGTCCGGCCG CACGCGTIACA GGAACGCTGA CATTGGCCGA      1140
GGGACCCCGG CCCTGATTTG TCGCGGATAC CACCCGCGCG CCGGCCAATT GGATTGGCGC      1200
CAGCCGTGAT TGCCCGGTGA GCCCCCGAGT TCCGTCTCCC GTGCGCGTGG CATGTGGGAA      1260
GCAATGAACG AGGAGAAACA CAGCCTTGAG CACCCTCCCG TGCAGGGCAG TTACGTGCGAA      1320
GGCGGTGTGG TCCGACATCC GGATGCCAAG GACTTCGGCA GCGCCGCCCG CCTGCCCGCC      1380
GATCCGACCT GGTTTAAGCA CGCCGTCTTC TACGAGGTGC TGGTCCGGGC GTTCTTCGAC      1440
GCCAGCGCGG ACGGTTCGCG CGATCTGCGT GGAATCATCG ATCGCCTCGA CTACCTGCAG      1500
TGGCTTGCCA TCGACTGCAT CTGTTGCCSC CGTTCCTAGC ACTCACCGCT GCGCGACGGC      1560
GGTTACGACA TTCGCGACTT CTACAAGGTG CTGCCCCGAA TCGGCACCGT CGACGATTTC      1620
GTGCCCTTGG TCGACACCGC TCACCGGCGA GGTATCCGCA TCATCACCGA CCTGGTGTATG      1680
AATCACACCT CGGAGTCCGA CCCCTGGTTT CAGGAGTCCC GCCCGGACCC AGACGGACCG      1740
TACGGTGACT ATTACGTGTG GAGCGACACC AGCGAGCGCT ACACCGACCG CCGGATCATC      1800
TTCGTGACA CCGAAGAGTC GAACTGTGTC TTCGATCCTG TCCGCCGACA GTTNTACTG      1860
GCACCGATTG TT                                     1872

```

10

20

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 355 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2: MTB32A (Ra35FL)

```

Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
1           5           10           15
Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
20           25           30
Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
35           40           45
Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val
50           55           60
Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
65           70           75           80

```

40

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val  
 85 90 95  
 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln  
 100 105 110  
 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala  
 115 120 125  
 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly  
 130 135 140  
 Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly  
 145 150 155 160  
 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu  
 165 170 175  
 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr  
 180 185 190  
 Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser  
 195 200 205  
 Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr  
 210 215 220  
 Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala  
 225 230 235 240  
 Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly  
 245 250 255  
 Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu  
 260 265 270  
 Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val  
 275 280 285  
 Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile  
 290 295 300  
 Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp  
 305 310 315 320  
 Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gln  
 325 330 335  
 Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly  
 340 345 350  
 Pro Pro Ala  
 355

10

20

<212> DNA  
 <213> Ra35 mature  
 <400> SEQ ID NO:3

catatgcatac accatcacca tcacgccccg ccggccttgt cgcaggaccg gttcgcggac 60  
 ttccccgcgc tgcocctcga cccgtcccgc atggtcgccc aagtggggcc acagggtggtc 120  
 aacatcaaca ccaaaactggg ctacaacaac gccgtggggc ccgggaccgg catcgtcacc 180  
 gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cggggcggcc cgacatcaat 240  
 gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc 300  
 caggatgtcg cgggtctgca gctgcgcggt gccggtggcc tgcctcggc gccgatcggg 360  
 ggcggcgtcg cggttggtga gccctcgtc gcgatgggca acagcgtgg gcagggcgga 420  
 acgcccctg cggtgccctgg cagggtggtc gcgctcggcc aaaccctgca ggcgtcggat 480  
 tcgctgaccg gtgccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag 540  
 cccggtgagg cggggcggcc cgtcgtcaac ggcctaggac agtggtcgg tatgaacacg 600  
 gccgcgtccg ataacttcca gctgtcccag ggtgggcagg gattcggcat tccgatcggg 660  
 caggcgatgg cgatcgggg ccagatccga tcgggtgggg ggtcaccac cgttcataac 720  
 gggcctaccg ccttcctcgg cttgggtgtt gtcgacaaca acggcaacgg cgcacgagtc 780  
 caacgcgtgg tcgggagcgc tccggcgga agtctcggca tctccaccgg cgacgtgac 840  
 accgcggtcg acggcgctcc gatcaactcg gccaccgca tggcggacgc gcttaacggg 900  
 catcatcccg gtgacgtcat ctccgtgacc tggcaaacca agtcgggccc caccgctaca 960  
 gggaaactga cattggccga gggacccccg gcctgagaat tc 1002

30

<212> PRT  
 <213> Ra35 mature  
 <400> SEQ ID NO:4

Met His His His His His His Ala Pro Pro Ala Leu Ser Gln Asp Arg  
 5 10 15

Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala  
 20 25 30  
 Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn  
 35 40 45  
 Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val  
 50 55 60  
 Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala  
 65 70 75 80  
 Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr  
 85 90 95  
 Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly  
 100 105 110  
 Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val  
 115 120 125  
 Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val  
 130 135 140  
 Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser  
 145 150 155 160  
 Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala  
 165 170 175  
 Ala Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu Gly  
 180 185 190  
 Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser  
 195 200 205  
 Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile  
 210 215 220  
 Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly  
 225 230 235 240  
 Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly  
 245 250 255  
 Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly  
 260 265 270  
 Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn  
 275 280 285  
 Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp  
 290 295 300  
 Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly  
 305 310 315 320  
 Asn Val Thr Leu Ala Glu Gly Pro Pro Ala  
 325 330

10

20

30

<212> DNA  
 <213> Ra35FLMutSA  
 <400> SEQ ID NO:5

catatgcatac accatcacca tcacgcccgc cgggccttgt cgcaggaccg gttcgcggac 60  
 ttccccgcgc tgcccctoga cccgtccgcg atggtcgccc aagtggggcc acaggtggtc 120  
 aacatcaaca ccaaactggg ctacaacaac gccgtggggc ccgggaccgg catcgtcacc 180

40

```

gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat 240
gcggttcagcg tccggtcccg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc 300
caggatgtcg cgggtgctgca gctgcgcggt gccggtggcc tgcctcggc ggcgatcgg 360
ggcggcgctcg cggttggtga gcccgctctc gcgatgggca acagcgggtgg gcaggggcga 420
acgccccgtg cgggtgctgg cagggtggtc gcgctcggcc aaaccgtgca ggcgtcggat 480
tcgctgaccg gtgccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag 540
cccgtgatg cgggcccggc cgtcgtcaac ggcctaggac aggtggtcgg tatgaacacg 600
cccgcgtccg ataacttcca gctgtcccag ggtgggcagg gattcggccat tccgatcggg 660
caggcgatgg cgatcggcgg ccagatccga tcgggtgggg ggtcaccac cgttcatatc 720
gggcctaccg ccttcctcgg cttgggtggt gtcgacaaca acggcaacgg cgcacgagtc 780
caacgcgtgg tcgggagcgc tccggcggca agtctcggca tctccacgg cgacgtgatc 840
accgcggtcg acggcgtccc gatcaactcg gccaccgcga tggcggacgc gcttaacggg 900
catcatcccg gtgacgtcat ctccgtgacc tggcaaacca agtcgggcgg cacgcgtaca 960
gggaacgtga cattggccga gggacccccg gcttgagaat tc 1002

```

```

<212> PRT
<213> Ra35FLMutSA
<400> SEQ ID NO:6

```

10

```

Met His His His His His Ala Pro Pro Ala Leu Ser Gln Asp Arg
      5              10              15

Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala
      20              25              30

Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn
      35              40              45

Asn Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val
      50              55              60

Val Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala
      65              70              75              80

Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr
      85              90              95

Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly
      100             105             110

Leu Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val
      115             120             125

Val Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val
      130             135             140

Pro Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser
      145             150             155             160

Leu Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala
      165             170             175

Ala Ile Gln Pro Gly Asp Ala Gly Gly Pro Val Val Asn Gly Leu Gly
      180             185             190

Gln Val Val Gly Met Asn Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser
      195             200             205

Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile
      210             215             220

Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly
      225             230             235             240

Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly
      245             250             255

```

20

30

Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly  
 260 265 270

Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn  
 275 280 285

Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp  
 290 295 300

Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly  
 305 310 315 320

Asn Val Thr Leu Ala Glu Gly Pro Pro Ala  
 325 330

(2) INFORMATION FOR SEQ ID NO:7: Ra35 (MTB32A N-term)

10

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 615 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

gccccgcccgccttctgctgcaggaccggttcgcccgaactccccgcgctgcccctcgaccgcctccgcg  
 atggctcgcccaagtggggccacaggtggtcaacatcaacaccaaactgggctacaacaacgcccgtg  
 ggcgcccgggaccggcatcgatcgatcccaacgggtgctggtgaccacaacaaccgctgatcgcg  
 ggcgccaccgacatcaatgcgttcagcgtcggtccggccaaaacctacggcgctcgatggtggtcggg  
 tatgaccgcacccaggatgctgcgggtgctgcagctgcgcgggtgccgggtggcctgcccgtcggcggcg  
 atcgggtggcggcgtcgcggttggtgagcccctcgtcgcgatgggcaacagcggtagggcaggcggga  
 acgccccgtgcgggtgctggcagggtggtcgcgctcggccaaaccgtgcaggcgtcggattcgtg  
 accggtgccgaagagacattgaacgggttgatccagttcgatgcccgcgatccagcccggtagggcg  
 ggcgggcccctcgtcaacggcctaggacaggtggtcgggtatgaacacggcccgcgtcc

20

(2) INFORMATION FOR SEQ ID NO:8: Ra35 (MTB32A N-term)

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 205 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu  
 Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val  
 Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr  
 Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val  
 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln  
 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala  
 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly  
 Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly  
 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu  
 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr  
 Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser

30

40

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr  
Ala Ala Ser

(2) INFORMATION FOR SEQ ID NO:9: Ra12

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 447 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGTATGAAC	ACGGCCCGGT	CCGATAACTT	CCAGCTGTCC	CAGGGTGGGC	AGGGATTCCG	60
CATTCCGATC	GGGCAGGCGA	TGGCGATCGC	GGGCCAGATC	CGATCGGGTG	GGGGGTCACC	120
CACCGTTCAT	ATCGGGCCTA	CCGCCTTCCT	CGGCTTGGGT	GTTGTCGACA	ACAACGGCAA	180
CGGCGCACGA	GTCCAACGCG	TGGTCCGGGAG	CGCTCCGGCG	GCAAGTCTCG	GCATCTCCAC	240
CGGCGACGTG	ATCACCCGGG	TCGACGGGCG	TCCGATCAAC	TCGGCCACCG	CGATGGCGGA	300
CGCGCTTAAC	GGGCATCATC	CCGGTGACGT	CATCTCGGTG	AACTGGCAAA	CCAAGTCCGG	360
CGGCACGCGT	ACAGGGAACG	TGACATTGCG	CGAGGGACCC	CCGGCCTGAT	TTCTGCGYGG	420
ATACCACCCG	CCGGCCGGCC	AATTGGA				447

10

(2) INFORMATION FOR SEQ ID NO:10: Ra12

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 132 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu	Ser	Gln	Gly	Gly	Gln	Gly	Phe
1				5					10					15	
Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala	Ile	Ala	Gly	Gln	Ile	Arg	Ser
		20					25						30		
Gly	Gly	Gly	Ser	Pro	Thr	Val	His	Ile	Gly	Pro	Thr	Ala	Phe	Leu	Gly
		35					40					45			
Leu	Gly	Val	Val	Asp	Asn	Asn	Gly	Asn	Gly	Ala	Arg	Val	Gln	Arg	Val
	50				55					60					
Val	Gly	Ser	Ala	Pro	Ala	Ala	Ser	Leu	Gly	Ile	Ser	Thr	Gly	Asp	Val
65					70					75				80	
Ile	Thr	Ala	Val	Asp	Gly	Ala	Pro	Ile	Asn	Ser	Ala	Thr	Ala	Met	Ala
			85					90						95	
Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly	Asp	Val	Ile	Ser	Val	Asn	Trp
			100					105						110	
Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr	Gly	Asn	Val	Thr	Leu	Ala	Glu
		115					120						125		
Gly	Pro	Pro	Ala												
			130												

20

30

(2) INFORMATION FOR SEQ ID NO:11: TbH9

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 851 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTGCAGGGTG	GCGTGGATGA	GCGTCACCGC	GGGGCAGGCC	GAGCTGACCG	CCGCCAGGT	60
CCGGGTGCT	GCGGCGGCT	ACGAGACGGC	GTATGGGCTG	ACGGTGCCCC	CGCCGGTGAT	120

```

CGCCGAGAAC CGTGTGAAC TGATGATTCT GATAGCGACC AACCTCTTGG GGCAAACAC 180
CCCCGGGATC GCGGTCAACG AGGCCGAATA CGCGGAGATG TGGGCCCAAG ACGCCGCCGC 240
GATGTTTGGC TAGCCCGCGG CGACGGCGAC GCGGACGGCG ACCTTGCTGC CGTTGAGGGA 300
GGCGCCGGAG ATGACCAGCG CGGGTGGGCT CCTCGAGCAG GCGCCGCGGG TCGAGGAGGC 360
CTCCGACACC GCCGCGGCGA ACCAGTTGAT GAACAATGTG CCCCAGGCGC TGAACACAGT 420
GGCCCCAGCCC ACGCAGGGCA CCACGCCTTC TTCCAAGCTG GGTGGCCTGT GGAAGACGGT 480
CTCGCCGCAT CGGTCGCCGA TCAGCAACAT GGTGTGCGAT GCCAACAAAC ACATGTGCGAT 540
GACCAACTCG GGTGTGTGCGA TGACCAACAC CTTGAGCTCG ATGTTGAAGG GCTTTGCTCC 600
GGCGGCGGCC GCCCAGGCGG TGCAAACCGC GCGCAAAAC GGGTCCGGG CGATGAGCTC 660
GCTGGGCAGC TCGCTGGGTT CTTCGGGTCT GGGCGGTGGG GTGGCCGCCA ACTTGGGTG 720
GGCGGCCTCG GTACGGIATG GTCACCGGGA TGGCGGAAAA TATGCANAGT CTGGTCCGGC 780
GAACCGTGGT CCGGCGTAAG GTTTACCCCC GTTTTCTGGA TCGCGTGAAC TTCGTCAACG 840
GAAACAGTTA C 851

```

(2) INFORMATION FOR SEQ ID NO:12:

10

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 263 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12: Tbh9

```

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
1 5 10 15
Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
20 25 30
Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
35 40 45
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
50 55 60
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
65 70 75 80
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
85 90 95
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
100 105 110
Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
115 120 125
Asn Asn Val Pro Gln Ala Leu Lys Gln Leu Ala Gln Pro Thr Gln Gly
130 135 140
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
145 150 155 160
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
165 170 175
Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
180 185 190
Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
195 200 205
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
210 215 220
Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
225 230 235 240
Ser Val Arg Tyr Gly His Arg Asp Gly Gly Lys Tyr Ala Xaa Ser Gly
245 250 255
Arg Arg Asn Gly Gly Pro Ala
260

```

20

30

(2) INFORMATION FOR SEQ ID NO:13: TBH9FL

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3058 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

40



(x1) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GATCGTACCC GTGCGAGTGC TCGGGCCGTT TGAGGATGGA GTGCACGFGT CTTTCGTGAT	60
GGCATAACCA GAGATGTTGG CGGCCGCGGC TGACACCCCTG CAGAGCATCG GTGCTACCAC	120
TGTGGCTAGC AATGCCGCTG CGGCCGCCCC GACGACTGGG GTGGTGGCCC CCGCTGCCGA	180
TGAGGTGTCG GCGCTGACTG CGGCGCACTT CGCCGCACAT GCGGCGATGT ATCAGTCCGT	240
GAGCGTCCG GCTGCTGCGA TTCATGACCA GTTCGTGGCC ACCCTTGCCA GCAGCGCCAG	300
CTCGTATGCG GCCACTGAAG TCGCCAATGC GCGGCGGCC AGCTAAGCCA GGAACAGTCG	360
GCACGAGAAA CCACGAGAAA TAGGGACACG TAATGGTGA TTTCCGGGCG TTACCACCGG	420
AGATCAACTC CGCGAGGATG TACGCCGGCC CGGGTTCGGC CTCGCTGGTG GCCCGCGCTC	480
AGATGTGGGA CAGCGTGGCG AGTGACCTGT TTTCCGGCCG GTCGGCGTTT CAGTCCGTGG	540
TCTGGGTCT GACGGTGGGG TCGTGATAG GTTCGTGGC GGGTCTGATG GTGGCGGCGG	600
CCTCGCCGTA TGTGGCTGG ATGAGCGTCA CGCGGGGCA GCGCGAGCTG ACCCGCGCCC	660
AGGTCCGGGT TGCTGCGGCG GCCTACGAGA CGCGGTATGG GCTGACGGTG CCCCAGCCGG	720
TGATCGCCGA GAACCGTGCT GAACTGATGA TTCTGATAGC GACCAACCTC TTGGGGCAAA	780
ACACCCCGGC GATCGCGGTC AACGAGGCCG AATACGGCGA GATGTGGGCC CAAGACGCCG	840
CCGCGATGTT TGGCTACGCC GCGGCGACGG CGACGGGAC GCGGACGTTG CTGCCGTTCG	900
AGGAGGCGCC GGAGATGACC AGCGCGGGTG GGCTCCTCGA GCAGGCCGCC GCGGTGAGG	960
AGGCCTCCGA CACCGCCGCG GCGAACCAAT TGATGAACAA TGTGCCCCAG GCGCTGCAAC	1020
AGCTGGCCCA GCCCACGCGG GGCACCACGC CTCTTCCAA GCTGGGTGGC CTGTGAAGA	1080
CGGTCTCGCC GCATCGGTGCG CCGATCAGCA ACATGGTGTG GATGSCCAAC AACCACATGT	1140
CGATGACCAA CTCGGGTGTG TCGATGACCA ACACCTTGAG CTCGATGTTG AAGGGCTTTG	1200
CTCCGGCGGC GCGCGCCAG GCCGTGCAAA CGCGGCGCA AAACGGGGTC CGGGCGATGA	1260
GCTCGCTGGG CAGCTCGCTG GGTTCCTCGG GTCGCGCGG TGGGGTGGCC GCCAACTTGG	1320
GTCGGGCGGC CTCGGTCCGT TCGTTGTCCG TGCCGCGGC CTGGGCGCG GCCAACCGG	1380
CAGTACCCC GCGGCGCGG GCGCTGCCGC TGACCAGCCT GACCAGCGCC GCGGAAAGAG	1440
GGCCCGGCA GATGCTGGGC GGGCTGCCGG TGGGGCAGAT GGGCGCCAGG GCCGTTGGTG	1500
GGCTCAGTGG TGTGCTGGT GTTCCGCCGC GACCCTATGT GATGCCGAT TCTCCGGCGG	1560
CCGGCTAGGA GAGGGGGCG AGACTGTCGT TATTTGACCA GTGATCGCGG GTCTCGGTGT	1620
TTCCGCGGCC GGCTATGACA ACAGTCAATG TGATGACAA GTTACAGGTA TTAGGTCCAG	1680
GTTCAACAAG GAGACAGGCA ACATGGCCTC ACGTTTATG ACGGATCCGC ACGCGATGCG	1740
GGACATGGCG GCGGTTTTG AGGTGCACGC CCAGACGGTG GAGGACGAGG CTCGCCGGAT	1800
GTGGGCGTCC GCGCAAAACA TTTCCGGTGC GGGCTGGAGT GGATGGCCG AGGCGACCTC	1860
GCTAGACACC ATGCCCCAGA TGAATCAGGC GTTTCGCAAC ATCGTGAACA TGCTGCACGG	1920
GGTGCCTGAC GGGCTGGTTC GCGACGCCAA CAACTACGAG CAGCAAGAGC AGGCCTCCCA	1980

10

20

30

GCAGATCCCTC AGCAGCTAAC GTCAGCCGCT GCAGCACAAT ACTTTTACAA GCGAAGGAGA 2040  
ACAGGTTCGA TGACCATCAA CTATCAATTC GGGGATGTCG ACGCTCACGG CGCCATGATC 2100  
CGCGCTCAGG CCGGGTTGCT GGAGGCCGAG CATCAGGCCA TCATTCGTGA TGTGTTGACC 2160  
GCGAGTGA CT TTTGGGGCGG CGCCGGTTCG GCGGCCTGCC AGGGGTTCAT TACCCAGTTG 2220  
GGCCGTA ACT TCCAGGTGAT CTACGAGCAG GCCAACGCC ACCGGCAGAA GGTGCAGGCT 2280  
GCCGGCAACA ACATGGCGCA AACCGACAGC GCGTCCGCT CCAGCTGGGC CTGACACCAG 2340  
GCCAAGGCCA GGGACGTGGT GTACGAGTGA AGTTCCTCGC GTGATCCTTC GGGTGGCAGT 2400  
CTAAGTGGTC AGTGCTGGGG TGTGGTGGT TTGCTGCTTG GCGGGTTCTT CGGTGTGGT 2460  
CAGTGTGCT CCGGCTCGGG TGAGGACCTC GAGGCCCAGG TAGCGCCGTC CTTCGATCCA 2520  
TTCGTGCTGT TGTTCGGCGA GGACGGCTCC GACGAGGCCG ATGATCGAGG CCGGTCGGG 2580  
GAAGATGCCC ACGACGTCGG TTCGGCTCG TACCTCTCGG TTGAGGCGTT CCTGGGGGTT 2640  
GTTGGACCAG ATTTGGCGCC AGATCTGCTT GGGAAGGCG GTGAACGCCA GCAGTCCGTT 2700  
CGGGCGGGT TCGAGGTGCT CCGCCACCGC GGGGAGTTT TCGTCCAGAG CGTCCAGTAC 2760  
CCGATCATAT TGGCAACAA CTGATTCGGC GTCGGGCTGG TCGTAGATGG AGTGCAGCAG 2820  
GGTGCACACC CACGGCCAGG AGGGCTTCGG GGTGGCTGCC ATCAGATTGG CTGCCGAGTG 2880  
GGTCTGCGAG CGCTGCCAG CCGCTCGGG CAGGGTGGCG CCGATCGCGG CCACCAGGCC 2940  
GGCGTGGCG TCGTGGTGA CCAGCGCGAC CCCGGACAGG CCGCGGGCGA CCAGTCCGCG 3000  
GAAGAAGGCC AGCCAGCCGG CCCCCTCCT GCGGAGGTT ACCTGGATGC CCAGGATC 3058

10

20

(2) INFORMATION FOR SEQ ID NO:14: TbH9FL

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 391 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met  
1 5 10 15  
Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp  
20 25 30  
Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser  
35 40 45  
Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly  
50 55 60  
Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr  
65 70 75 80  
Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala  
85 90 95  
Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala  
100 105 110

30

Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly  
115 120 125

Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met  
130 135 140

Trp Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala  
145 150 155 160

Thr Ala Thr Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr  
165 170 175

Ser Ala Gly Gly Leu Leu Glu Gln Ala Ala Val Glu Glu Ala Ser  
180 185 190

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu  
195 200 205

Gln Gln Leu Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu  
210 215 220

Gly Gly Leu Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn  
225 230 235 240

Met Val Ser Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val  
245 250 255

Ser Met Thr Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala  
260 265 270

Ala Ala Ala Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala  
275 280 285

Met Ser Ser Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly  
290 295 300

Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val  
305 310 315 320

Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg  
325 330 335

Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly  
340 345 350

Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly  
355 360 365

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met  
370 375 380

Pro His Ser Pro Ala Ala Gly  
385 390

10

20

30

<210> SEQ ID NO:15  
<211> 2287  
<212> DNA  
<213> Artificial Sequence  
<223> Description of Artificial Sequence:tri-fusion  
protein Mtb72F(Ra12-TbH9-Ra35 or Mtb32-Mtb39  
fusion)

tctagaata attttgttta ctttaagaan ganatataca 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

gga ttc gcc att ccg atc ggg cag gcg atg gcg atc gcg ggc cag atc 152  
 Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile  
                   25                                  30                                  35

cga tcg ggt ggg ggg tca ccc acc gtt cat atc ggg cct acc gcc ttc 200  
 Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe  
                   40                                  45                                  50

ctc ggc ttg ggt gtt gtc gac aac aac ggc aac ggc gca cga gtc caa 248  
 Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln  
                   55                                  60                                  65

cgc gtg gtc ggg agc gct ccg gcg gca agt ctc ggc atc tcc acc ggc 296  
 Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly  
                   70                                  75                                  80                                  85

gac gtg atc acc gcg gtc gac ggc gct ccg atc aac tcg gcc acc gcg 344  
 Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala  
                   90                                  95                                  100

atg gcg gac gcg ctt aac ggg cat cat ccc ggt gac gtc atc tcg gtg 392  
 Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val  
                   105                                  110                                  115

acc tgg caa acc aag tcg ggc ggc acg cgt aca ggg aac gtg aca ttg 440  
 Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu  
                   120                                  125                                  130

gcc gag gga ccc ccg gcc gaa ttc atg gtg gat ttc ggg gcg tta cca 488  
 Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro  
                   135                                  140                                  145

ccg gag atc aac tcc gcg agg atg tac gcc ggc ccg ggt tcg gcc tcg 536  
 Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser  
                   150                                  155                                  160                                  165

ctg gtg gcc gcg gct cag atg tgg gac agc gtg gcg agt gac ctg ttt 584  
 Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe  
                   170                                  175                                  180

tcg gcc gcg tcg gcg ttt cag tcg gtg gtc tgg ggt ctg acg gtg ggg 632  
 Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly  
                   185                                  190                                  195

tcg tgg ata ggt tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tcg 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 ccg 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

10

20

30

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 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 tcg 1160  
 Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser  
 360 365 370

ccg cat cgg tcg ccg atc agc aac atg gtg tcg 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 tcg atg acc aac tcg ggt gtg tcg atg acc aac acc ttg agc tcg 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 cgc cag gcc gtg caa acc 1304  
 Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Arg Gln Ala Val Gln Thr  
 410 415 420

gcg gcg caa aac ggg gtc ccg gcg atg agc tcg ctg gcc agc tcg ctg 1352  
 Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu  
 425 430 435

ggt tct tcg 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 tcg gtc ggt tcg ttg tcg 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 ccg 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 ggc 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 tcg cag gac ccg 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

10

20

30

ctg ccc ctc gac ccg 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 tcc 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 Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala  
 665 670 675

ctc ggc caa acc gtg cag gcg tcc gat tcc 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

tcc 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 agccccgaaa 2287

10

20

<210> SEQ ID NO:16  
 <211> 729  
 <212> PRT  
 <213> Artificial Sequence  
 <223> Description of Artificial Sequence:tri-fusion  
 protein Mtb72F (Ra12-TbH9-Ra35 or Mtb32-Mtb39  
 fusion)

30

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

40

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	
						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	Arg	
					405				410					415		
Gln	Ala	Val	Gln	Thr	Ala	Ala	Gln	Asn	Gly	Val	Arg	Ala	Met	Ser	Ser	

10

20

30

40

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

10

20

30

<210> SEQ ID NO:17  
<211> 2190  
<212> DNA  
<213> Mtb72FMutSA

atgcatcacc atcaccatca cacggcgcg tccgataact tccagctgtc ccagggtggg 60  
caggattcgc ccattccgat cgggcaggcg atggcgatcg cgggccagat ccgatcgggt 120  
gggggggtcac ccaccgttca tatcgggcct accgccttcc tcggcttggg tgttgtcgac 180

40



```

aacaacggca acggcgacg agtccaacgc gtggtcggga gcgctccggc ggcaagtctc 240
ggcatctcca cggcgacagt gatcaccgcg gtcgacggcg ctccgatcaa ctccggccacc 300
gggatggcgg acgcgcttaa cgggcatacat cccggtgacg tcactctcggg gacctggcaa 360
accaagtccg gcggcacgcg tacagggaaac gtgacattgg ccgagggacc cccggccgaa 420
ttcatggtgg atttcggggc gttaccaccg gagatcaact ccgagggat gtaccgccc 480
ccgggttcgg cctcgcctgg ggcccggct cagatgtggg acagcgtggc gagtgacctg 540
ttttcggcgg cgtcggcggt tcagtcgggt gtctggggtc tgacgggtgg gtctgggata 600
ggttcgctcg cgggtctgat ggtggcggcg gcctcggcgt atgtggcgtg gatgagcgtc 660
accgcggggc aggcggagct gaccgcccgc caggtccggg ttgctcggc ggctacgag 720
acggcgatg ggctgacggt gcccgcggcg tgatcgccg agaaccgtgc tgaactgatg 780
attctgatag cgaccaacct ctgggggcaa aacaccccgg cgatcgggt caacgaggcc 840
gaatacggcg agatgtgggc ccaagacgcc gccgcgatgt ttggctacgc cgcggcgacg 900
gcgacggcga cggcgacggt gctgcccgtt caggagggcg cggagatgac cagcggcgg 960
gggctcctcg agcaggccgc cgcggtcggg gaggcctccg acaccgcccg ggcgaaccag 1020
ttgatgaaca atgtgcccga ggcgctgcaa cagctggccc agcccacgca gggcaccacg 1080
ccttcttcca agctgggtgg cctgtggaag acggtctcgc cgcactcggc gccgatcagc 1140
aacatggtgt cgatggccaa caaccacatg tcgatgacca actcgggtgt gtcgatgacc 1200
aacaccttga gctcgatggt gaaggcctt gctccggcgg cggccgcccc ggccgtgcaa 1260
accgcggcgc aaaacggggc cggggcgatg agctcgcctg gcagctcgtt gggttctctg 1320
ggtctggggc gtgggggtgg cgcacaactg ggtcggggcg cctcgggtcgg ttcgttctcg 1380
gtgcccagcg cctggggcgc ggccaaccag gcagtcaccc cggcggcggc ggcgctgccc 1440
ctgaccagcc tgaccagcgc cgcggaaga gggcccggcg agatgctggg cgggctgccc 1500
gtggggcaga tgggcgcag ggcgggtggt gggctcagtg gttgctgctg tttccgccc 1560
cgaccctatg tgatgcccga ttctcggca gccggcgata tgcggcccgc ggccttctg 1620
caggaccggt tcgcccactt ccccgcgctg cccctcgacc cgtcccgat ggtcggccaa 1680
gtggggccac aggtggtcaa catcaacacc aaactgggt acaacaacgc cgtggggcgc 1740
gggaccggca tcgtcatcga tcccacggg tccgtcctga ccaacaacca cgtgatcgcg 1800
ggcggcaccg acatcaatgc gttcagcgtc ggctccggcc aaacctacgg cgtcgatgtg 1860
gtcgggtatg accgcaccca ggatgtcgcg gtgctgcagc tcgcccgtgc cgggtggcctg 1920
ccgtcggcgg cgatcgggtg cggcgtcgcg gttggtgagc ccgtcgtcgc gatgggcaac 1980
agcgggtggc agggcggaac gcccgcgtgc gtgcctggca ggggtggtgc gctcggccaa 2040
accgtcagg cgtcggatc gctgaccggt gccgaagaga cattgaacgg gttgatccag 2100
ttcgtgccc cgatccagcc cgggtgatgc ggcgggcccg tcgtcaacgg cctaggacag 2160
gtggtcggta tgaacacggc cgcgtcctag 2190

```

10  
20

```

<210> SEQ ID NO:18
<211> 729
<212> PRT
<213> Mtb72FMutSA

```

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
      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

```

30

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 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

10

20

30

40

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

Val Val Gly Met Asn Thr Ala Ala Ser  
725

10

20

<210> SEQ ID NO:19  
<211> 1797  
<212> DNA  
<213> Artificial Sequence  
<223> Description of Artificial Sequence:bi-fusion  
protein TbH9-Ra35 (designated Mtb59f)  
<222> (1)..(1791)

30

cat atg cat cac cat cac cat cac atg gtg gat ttc ggg gcg tta cca 48  
His Met His His His His His His Met Val Asp Phe Gly Ala Leu Pro  
1 5 10 15

ccg gag atc aac tcc gcg agg atg tac gcc ggc ccg ggt tcg gcc tcg 96  
Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser  
20 25 30

ctg gtg gcc gcg gct cag atg tgg gac agc gtg gcg agt gac ctg ttt 144  
Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe  
35 40 45

tcg gcc gcg tcg gcg ttt cag tcg gtg gtc tgg ggt ctg acg gtg ggg 192  
Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly  
50 55 60

tcg tgg ata ggt tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tcg ccg 240  
 Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro  
 65 70 75 80

tat gtg gcg tgg atg agc gtc acc gcg ggg cag gcc gag ctg acc gcc 288  
 Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala  
 85 90 95

gcc cag gtc cgg gtt gct gcg gcg gcc tac gag acg gcg tat ggg ctg 336  
 Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu  
 100 105 110

acg gtg ccc ccg ccg gtg atc gcc gag aac cgt gct gaa ctg atg att 384  
 Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile  
 115 120 125

ctg ata gcg acc aac ctc ttg ggg caa aac acc ccg gcg atc gcg gtc 432  
 Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val  
 130 135 140

aac gag gcc gaa tac ggc gag atg tgg gcc caa gac gcc gcc gcg atg 480  
 Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met  
 145 150 155 160

ttt ggc tac gcc gcg gcg acg gcg acg gcg acg gcg acg ttg ctg ccg 528  
 Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Leu Leu Pro  
 165 170 175

ttc gag gag gcg ccg gcg gag atg acc agc gcg ggt ggg ctc ctc gag cag 576  
 Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln  
 180 185 190

gcc gcc gcg gtc gag gag gcc tcc gac acc gcc gcg gcg aac cag ttg 624  
 Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu  
 195 200 205

atg aac aat gtg ccc ccg gcg ctg caa cag ctg gcc cag ccc acg cag 672  
 Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln  
 210 215 220

ggc acc acg cct tct tcc aag ctg ggt ggc ctg tgg aag acg gtc tcg 720  
 Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser  
 225 230 235 240

ccg cat cgg tcg ccg atc agc aac atg gtg tcg atg gcc aac aac cac 768  
 Pro His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His  
 245 250 255

atg tcg atg acc aac tcg ggt gtg tcg atg acc aac acc ttg agc tcg 816  
 Met Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser  
 260 265 270

atg ttg aag ggc ttt gct ccg gcg gcg gcc gcc cag gcc gtg caa acc 864  
 Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr  
 275 280 285

gcg gcg caa aac ggg gtc ccg gcg atg agc tcg ctg ggc agc tcg ctg 912  
 Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu  
 290 295 300

ggt tct tcg ggt ctg ggc ggt ggg gtg gcc gcc aac ttg ggt cgg gcg 960  
 Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala  
 305 310 315 320

gcc tcg gtc ggt tcg ttg tcg gtg ccg cag gcc tgg gcc gcg gcc aac 1008  
 Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn  
 325 330 335

10

20

30

cag gca gtc acc ccg gcg gcg cgg gcg ctg ccg ctg acc agc ctg acc 1056  
 Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr  
 340 345 350

agc gcc gcg gaa aga ggg ccc ggg cag atg ctg ggc ggg ctg ccg gtg 1104  
 Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val  
 355 360 365

ggg cag atg ggc gcc agg gcc ggt ggt ggg ctc agt ggt gtg ctg cgt 1152  
 Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg  
 370 375 380

gtt ccg ccg cga ccc tat gtg atg ccg cat tct ccg gca gcc ggc gat 1200  
 Val Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp  
 385 390 395 400

atc gcc ccg ccg gcc ttg tcg cag gac cgg ttc gcc gac ttc ccc gcg 1248  
 Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala  
 405 410 415

ctg ccc ctc gac ccg tcc gcg atg gtc gcc caa gtg ggg cca cag gtg 1296  
 Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val  
 420 425 430

gtc aac atc aac acc aaa ctg ggc tac aac aac gcc gtg ggc gcc ggg 1344  
 Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly  
 435 440 445

acc ggc atc gtc atc gat ccc aac ggt gtc gtg ctg acc aac aac cac 1392  
 Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His  
 450 455 460

gtg atc gcg ggc gcc acc gac atc aat gcg ttc agc gtc ggc tcc ggc 1440  
 Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly  
 465 470 475 480

caa acc tac ggc gtc gat gtg gtc ggg tat gac cgc acc cag gat gtc 1488  
 Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val  
 485 490 495

gcg gtg ctg cag ctg cgc ggt gcc ggt ggc ctg ccg tcg gcg gcg atc 1536  
 Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile  
 500 505 510

ggt ggc ggc gtc gcg gtt ggt gag ccc gtc gtc gcg atg ggc aac agc 1584  
 Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser  
 515 520 525

ggt ggg cag ggc gga acg ccc cgt gcg gtg cct ggc agg gtg gtc gcg 1632  
 Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala  
 530 535 540

ctc ggc caa acc gtg cag gcg tcg gat tcg ctg acc ggt gcc gaa gag 1680  
 Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu  
 545 550 555 560

aca ttg aac ggg ttg atc cag ttc gat gcc gcg atc cag ccc ggt gat 1728  
 Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp  
 565 570 575

tcg ggc ggg ccc gtc gtc aac ggc cta gga cag gtg gtc ggt atg aac 1776  
 Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn  
 580 585 590

acg gcc gcg tcc taggatatc 1797  
 Thr Ala Ala Ser  
 595

10

20

30

&lt;210&gt; SEQ ID NO:20

&lt;211&gt; 596

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

<223> Description of Artificial Sequence:bi-fusion  
protein TbH9-Ra35 (designated Mtb59f)

His Met His His His His His His Met Val Asp Phe Gly Ala Leu Pro  
 1 5 10 15  
 Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser  
 20 25 30  
 Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe  
 35 40 45  
 Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly  
 50 55 60  
 Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro  
 65 70 75 80  
 Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala  
 85 90 95  
 Ala Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu  
 100 105 110  
 Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile  
 115 120 125  
 Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val  
 130 135 140  
 Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met  
 145 150 155 160  
 Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Leu Leu Pro  
 165 170 175  
 Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln  
 180 185 190  
 Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu  
 195 200 205  
 Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln  
 210 215 220  
 Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser  
 225 230 235 240  
 Pro His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His  
 245 250 255  
 Met Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser  
 260 265 270  
 Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr  
 275 280 285  
 Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu  
 290 295 300  
 Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala  
 305 310 315 320  
 Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn

10

20

30

325 330 335

Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr  
340 345 350

Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val  
355 360 365

Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg  
370 375 380

Val Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp  
385 390 395 400

Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala  
405 410 415

Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val  
420 425 430

Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly  
435 440 445

Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His  
450 455 460

Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly  
465 470 475 480

Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val  
485 490 495

Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile  
500 505 510

Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser  
515 520 525

Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala  
530 535 540

Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu  
545 550 555 560

Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp  
565 570 575

Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn  
580 585 590

Thr Ala Ala Ser  
595

10

20

30

(2) INFORMATION FOR SEQ ID NO:21: DPV (MTB8.4)

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 500 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGTGGCAATG TCGTTGACCG TCGGGGCCGG GGTGCTCTCC GCAGATCCCG TGGACGCGGT 60  
CATTAAACACC ACCGTCAATT ACGGGCAGGT AGTAGCTGCG CTC AACCGCA CGGATCCGGG 120  
GGCTGCCGCA CAGTTCACG CCTCACCGGT GGCGCAGTCC TATTGCGCA ATTCCTCGC 180  
CGCACCGCCA CCTCAGCGCG CTGCCATGGC CGCGCAATTG CAAGCTGTGC CGGGGCCGGC 240

40

```

ACAGTACATC GGCCTTGTCG AGTCGGTTGC CGGCTCCTGC AACAACTATT AAGCCCATGC 300
GGGCCCCATC CCGCGACCCG GCATCGTTCG CCGGGCTAGG CCAGATTGCC CCGCTCCTCA 360
ACGGGGCGCA TCCCGCGACC CCGCATCGTC GCCGGGGCTA GGCCAGATTG CCCCCTCCT 420
CAACGGGCGG CATTCGTCGC CGAATTCCTG CAGCCCGGGG GATCCACTAG TTCTAGAGCG 480
GCCGCCACCG CGGTGGAGCT 500

```

(2) INFORMATION FOR SEQ ID NO:22: DPV (MTB8.4)

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 96 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```

Val Ala Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro
1          5          10          15
Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala
20          25          30
Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser
35          40          45
Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro
50          55          60
Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala
65          70          75          80
Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr
85          90          95

```

10

(2) INFORMATION FOR SEQ ID NO:23: MSL (MTB9.8)

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 585 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```

TGGATTCCGA TAGCGGTTTC GGCCCTCGA CCGGCGACCA CCGCGCGCAG GCCTCCGAAC 60
GGGGGGCCGG GACGCTGGGA TTCGCCGGA CCGCAACCA AGAACGCCGG GTCCGGGGCGG 120
TCGGGCTGAC CGCACTGGCC GGTGATGAGT TCGGCAACCG CCCCCTGGATG CCGATGGTGC 180
CGGGGACCTG GGAGCAGGGC AGCAACGAGC CCGAGGCGCC CGACGGATCG GGGAGAGGGG 240
GAGGCGACCG CTTACCGCAC GACAGCAAGT AACCGAATTC CGAATCACGT GGACCCGTAC 300
GGGTCGAAAG GAGAGATGTT ATGAGCCTTT TGATGCTCA TATCCACAG TTGGTGGCCT 360
CCCAGTCGGC GTTTGCGGCC AAGGCGGGGC TGATGCGGCA CACGATCGGT CAGGCCGAGC 420
AGGCGGCGAT GTCGGCTCAG GCGTTTCACC AGGGGGAGTC GTCGGCGGCG TTTCAGGCCG 480
CCCATGCCCG GTTTGTGGCG GCGGCCGCCA AAGTCAACAC CTTGTTGGAT GTCGCGCAGG 540
CGAATCTGGG TGAGGCCGCC GGTACCTATG TGGCCGCGGA TGCTG 585

```

30

(2) INFORMATION FOR SEQ ID NO:24: MSL (MTB9.8)

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 97 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser  
 1 5 10 15  
 Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala  
 20 25 30  
 Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser  
 35 40 45  
 Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys  
 50 55 60  
 Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala  
 65 70 75 80  
 Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ser Thr Tyr Thr Gly  
 85 90 95  
 Phe

10

(2) INFORMATION FOR SEQ ID NO:25: MTI (MTB9.9A)

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1742 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCGCTCTCTT TCAACGTCAT AAGTTCGGTG GGCCAGTCGG CCGCGCGTGC ATATGGCACC 60  
 AATAACGCGT GTCCCATGGA TACCCGGACC GCACGACGGT AGAGCGGATC AGCGCAGCCG 120  
 GTGCCGAACA CTACCGCGTC CACGCTCAGC CTTGCCGCGT TCGGGAAGAT CGAGCCCBAG 180  
 TTCTCATGGT CGTTAACGCC TTCCARCACT GCGACGGTGC GCGCCCCGGC GACCACCTGA 240  
 GCAACGCTCG GCTCCGGCAC CCGGCGCGCG GCTGCCAACA CCCCACGATT GAGATGGAAG 300  
 CCGATCACCC GTGCCATGAC ATCAGCCGAC GCTCGATAGT ACGGCGCGCC GACACCGGCC 360  
 AGATCATCCT TGAGCTCGGC CAGCCGGCGG TCGGTGCCGA ACAGCGCCAG CGGCGTGAAC 420  
 CGTGAAGCCA GCATCGCGTG CACCACCAGC ACACCCFCGG CGATCACCAA CGCCTTGCCG 480  
 FTGGGCAGAT CGGGACNACN GTCGATGCTG TTCAGGTCAC GAAATCGTC GAGCCGTGGG 540  
 TCGTCCGGAT CGCAGACGTC CTGAACATCG AGGCCGTCGG GTTGCTGGGC ACAACGGCCT 600  
 TCGGTCACGG GCTTTCGTCG ACCAGAGCCA GCATCAGATC GCGGCGGCTG CGCAGGATGT 660  
 CACGCTCGCT CCGGTTCCAG GTCGCGAGCC GCTCAGCCAG CCACTCTTGC AGAGAGCCGT 720  
 TGCTGGGATT AATTGGGAGA GGAAGACAGC ATGTCTGTTC TGACCACACA GCCGGAAGCC 780  
 CTGGCAGCTG CCGCGCGGAA CCTACAGGGT ATTGGCACGA CAATGAACGC CCAGAACCGC 840  
 GCCGCGGCTG CTCCAACCAC CGGAGTAGTG CCGCAGCCG CCGATGAAGT ATCAGCGCTG 900  
 ACCGCGGCTC AGTTTGTCTG GCACGCGCAG ATGTACCAA CCGTCAGCGC CCAGGCCGCG 960  
 GCCATTCACG AATGTTCTGT GAACACGCTG GTGGCCAGTT CTGGCTCATA CGCGGCCACC 1020  
 GAGGCGGCCA ACGCAGCCGC TGCCCGCTGA ACGGGCTCGC ACGAACCTGC TGAAGGAGAG 1080  
 GGGGAACATC CGGAGTTCTC GGGTCAGGGG TTGGCCAGC GCCCAGCCGA TTCAGNTATC 1140  
 GGCGTCCATA ACAGCAGACG ATCTAGGCAT TCAGTACTAA GGAGACAGGC AACATGGCCT 1200  
 CACGTTTTAT GACGGATCCG CATGCGATGC GGGACATGGC GGGCCGTTTT GAGGTGCACG 1260  
 CCCAGACGGT GGAGGACGAG GCTCGCCGGA TGTGGGCGTC CGCGCAAAC ATTTCCGGTG 1320  
 CCGGCTGGAG TGGCATGCC GAGGCGACCT CGCTAGACAC CATGACCTAG ATGAATCAGG 1380  
 CGTTTCGCAA CATCGTGAAC ATGCTGCACG GGTGCGTGA CCGGCTGGTT CGCGACGCCA 1440  
 ACAANTACGA ACAGCAAGAG CAGGCCTCCC AGCAGATCCT GAGCAGNTAG CGCCGAAAGC 1500  
 CACAGCTGNG TACGNTTCT CACATTAGGA GAACACCAAT ATGACGATTA ATTACCAGTT 1560  
 CGGGGACGTC GACGCTCATG GCGCCATGAT CCGCGCTCAG GCGGCGTCGC TTGAGGCGGA 1620  
 GCATCAGGCC ATCGTTCGTG ATGTGTTGGC CGCGGCTGAC TTTTGGGGCG GCGCCGGTTC 1680  
 GGTGGCTTGC CAGGAGTTCA TTACCCAGTT GGGCCGTAACT TTCCAGGTGA TCTACGAGCA 1740  
 GG 1742

20

30

(2) INFORMATION FOR SEQ ID NO:26: MTI (MTB9.9A)

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2836 base pairs

40

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

GTGTGATTCOG TTCGCGGGCCG CGCCGAAGAC CACCAACTCC GCTGGGGTGG TCGCACAGGC      60
GGTTGCGTCCG GTCAGCTGGC CGAATCCCAA TGATTGGTGG CTGNGTGC GG TGTCTGGGCT      120
CGATTACCCC CACGGAAAAG ACGACGATCG TTCGTTTGCT CGGTCAGTCG TACTTGGCGA      180
CGGGCATGGC GCGGTTTTCT ACCTCGATCG CACAGCAGCT GACCTTCGGC CCAGGGGGCA      240
CAACGGGTGG CTCGCGCGGA GCCTGGTACC CAACGCCACA ATTCGCGCGC CTGGGTGCAG      300
GCCCGGGCGT GTCGCGGAGT TTGGCGCGGG CGGAGCCGGT CGGAGGTTG TCGGTGCCCG      360
CAAGTTGGGC CGTCGCGGCT CCGGCCTTCG CGGAGAAGCC TGAGGGGGGC ACGCCGATGT      420
COGTCAATCG CGAAGCGTCC AGCTGCGGTC AGGGAGGCC T GCTTCGAGGC ATACCGCTGG      480
CGAGAGCGGG GCGGCGTACA GCGGCCTTCG CTCACCGATA CCGGTTCGCG CACAGCGTGA      540
TTACCCGGTC TCCGTCGAGT GGATAGCTTT CGATCCGGTC TGGCGGGCGC CCGGAAATGC      600
TGCAGATAGC GATCGACCGC GCCGGTCCGT AAACGCCGCA CACGGCACTA TCAATGCGCA      660
CGGCGGGCGT TGATGCCAAA TTGACCGTCC CGACGGGGCT TTATCTCGCG CAAGATTTCA      720
TCCCCAGCCC GGTCCGTGGG CCGATAAATA CGCTGGTCAG CGCGACTCTT CCGGCTGAAT      780
TCGATGCTCT GGGCGCCCGC TCGACGCCGA STATCTCGAG TGGGCCGCAA ACCCGGTCAA      840
AGCTGTATTG TGTGGCGTTA CCACAGGTGA ATTTGCGGGT CCAACTGGTG AACACTTGCG      900
AACGGGTGGC ATCGAAATCA ACTTGTTCGC TTGCAGTGAT C TACTCTCTT GCAGAGAGCC      960
GTTCCTGGGA TTAATTGGGA GAGGAAGACA GCATGTCTGT CGTGACCACA CAGCCGGAAG     1020
CCCTGGCAGC TGGCGCGCGG AACCTACAGG GTATTGGCAC GACAATGAAC GCCCAGAACG     1080
CGGCCGCGGC TGCTCCAACC ACCGGAGTAG TGCCCGCAGC CGCGATGAA GTATCAGCGC     1140
TGACCCGCGC TCAGTTTGGT GCGCACGCGC AGATGTACCA AACGGTCAGC GCCCAGGCGC     1200
CGGCCATTCA CGAAATGTTT GTGAACACGC TGGTGGCCAG TTCTGGCTCA TACGCGGCCA     1260
CCGAGGCGCG CAACGCAGCC GCTGCCGGCT GAACGGGCTC GCACGAACCT GCTGAAGGAG     1320
AGGGGGAAAC TCCGGAGTTC TCGGGTCAGG GGTTCGSCCA GCGCCAGCC GATTCACTA     1380
TCGGCGTCCA TAAACAGAGA CGATCTAGGC ATTCAGTACT AAGGAGACAG GCAACATGGC     1440
CTCACGTTTT ATGACGGATC CGCATGCGAT GCGGGACATG CCGGGCCGTT TTGAGGTGCA     1500
CGCCAGACG GTGGAGGACG AGSCTCGCCG GATGTGGGCG TCCGCGCAA ACATTTCCGG     1560
TGCGGGCTGG AGTGGCATGG CCGAGGCGAC CTCGCTAGAC ACCATGACCT AGATGAATCA     1620
GGCGTTTTCC AACATCGTGA ACATGCTGCA CCGGGTGGCT GACGGGCTGG TTCGCGACGC     1680
CAACAACCTC GAACAGCAAG AGCAGGCCCT CCACAGATC CTGAGCAGCT AGCCCGGAAA     1740
GCCACAGCTG CGTACGCTTT CTCACATTAG GAGAACACCA ATATGACGAT TAATTACCGA     1800
TTGGGGGACG TCGACGCTCA TGGCGCCATG ATCCGCGCTC AGGCGGCGTC GCTTGAGGCG     1860
GAGCATCAGG CCATCGTTCG TGATGTGTTG GCCGCGGGTG ACTTTCGGGG CGGGCCCGGT     1920
TCGGTGGCTT GCCAGGAGTT CATTAACCCAG TTGGGCCGTA ACTTCCAGGT GATCTACGAG     1980
CAGGCCAACC CCCACGGGCA GAAGTGGCAG GCTGCCGGCA ACAACATGGC GCAAACCGAC     2040
AGCGCCGTCG GCTCCAGTGG GGCCTAAAC TGAACATCAG TCGGGCAGC ACACCAACCA     2100
GCCGTTGTGC TGCTGTGTC TGCAGTTAAC TAGCACTCGA CCGCTGAGGT AGCGATGGAT     2160
CAACAGAGTA CCCGCACCGA CATCACCGTC AACGTCCGAC GCTTCTGGAT GCTTCAGGCG     2220
CTACTGGATA TCCGCCACGT TCGCCCTGAG TTAGCTTGCC GGCCGTACGT CTCCACCGAT     2280
TCCAATGACT GGCTAAACGA GCACCCGGGG ATGGCGGTCA TGCCCGAGCA GGGCATTTGT     2340
GTCAACGACG CGGTCAACGA ACAGTTCGCT GCCCGGATGA AGTGTCTTGC CGCACCTGAT     2400
CTTGAAGTCG TCGCCCTGCT GTCACGCGCG AAGTTGCTGT ACCGGGTCAT AGACGACGAG     2460
AACCAGCCGC CGGTTTCGCG TGACATCCCT GACAATGAGT TCCGGTGGT GTTGCCCGG     2520
CGAGGCCAGC ACTGGGTGTC GCGGTACCG GTTGGCAATG ACATCACCGT CGATGACGTG     2580
ACGGTCTCGG ATAGCGCTC GATCGCCGCA CTGTAATGG ACGTCTGGA GTCGATTCAC     2640
CACGCCGACC CAGCCCGAT CAACGCGGTC AACGTGCCAA TGGAGGAGAT CTCGTGCCGA     2700
ATTCGCGACG AGGCACGAGG CGGTGTTCGCT GACGACGGGA TCGATCACGA TCATCGACCG     2760
GCCGGGATCC TTGGCGATCT CGTTGAGCAC GACCCGGGCC CGCGGGAAGC TCTGCGACAT     2820
CCATGGGTTT TTCCCG

```

10

20

30

(2) INFORMATION FOR SEQ ID NO:27: MTL (MTB9.9A)

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
1      5      10      15
Ile Arg Ala Leu Ala Gly Leu Leu Glu Ala Glu His Gln Ala Ile Ile
      20      25      30
Ser Asp Val Leu Thr Ala Ser Asp Phe Trp Gly Gly Ala Gly Ser Ala
      35      40      45
Ala Cys Gln Gly Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile
      50      55      60
Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn
65      70      75      80
Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala
      85      90

```

10

(2) INFORMATION FOR SEQ ID NO:28: HTCC#1

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```

CAGGCATGAG CAGAGCGTTC ATCATCGATC CAACGATCAG TGCCATTGAC GGCTTGTACG      60
ACCTTCTGGG GATTGGAATA CCAACCAAG GGGGATCCT TACTCCTCA CTAGAGTACT      120
TCGAAAAGC CCTGGAGGAG CTGGCAGCAG CGPTTCCGGG TGATGGCTGG TTAGGTTCCG      180
CCGCGGACAA ATACCCCGC AAAAACCACA ACCACGTGAA TTTTTCAG GAACTGGCAG      240
ACCTCGATCG TCAGTCAATC AGCCTGATCC ACGACCAGGC CAACGCGGTC CAGACGACCC      300
GCGACATCCT GGAGGGCGCC AAGAAAGGTC TCGAGTTCGT GCGCCCGGTC GCTGTGGACC      360
TGACCTACAT CCCGTCGTC GGGCACGCC TATCGGCCGC CTTCCAGGCG CCGTPTTGGC      420
CGGGCGGAT GGCCTAGTGC GCGCGCGCG TTGCCTACTT GGTCTGAAA ACGCTGATCA      480
ACGCGACTCA ACTCCTCAA TTGCTTGCCA AATGGCGGA GTTGGTCGCG GCCGCCATTG      540
CGGACATCAT TTCGGATGTG GCGGACATCA TCAAGGGCAC CCTCGGAGAA GTGTGGGAGT      600
TCATCACAAA CGCGTCAAC GGCCTGAAAG AGCTTTGGGA CAAGCTCACG GGGTGGGTGA      660
CCGACTGTT CTCTGAGGG TGGTCGAACC TGGAGTCCTT CTTGCGGGC GTCCC CGGCT      720
TGACCGGCGC GACCAGCGGC TTPFCGCAAG TGA CTGGCTT GTTCGGTGGC GCCGGTCTGT      780
CCGCATCGTC GGGCTTGGCT CACGCGGATA GCCTGGCGAG CTCAGCCAGC TTGCCCGCCC      840
TGGCCGGCAT TGGGGCGGG TCCGGTTTTG GGGCTTGCC GAGCCTGGCT CAGGTCCATG      900
CCGCTCAAC TCGGCAGGCG CTACGGCCCC GAGCTGATGG CCCGTCGCG GCCGCTGCCG      960
AGCAGSTCGG CCGGCAGTCG CAGCTGGTCT CCGCGCAGGG TTCCAAGST ATGGGOGGAC      1020
CCGTAGGCAT GGGCGGCATG CACCCCTCTT CCGGGCGGTC GAAAGGGACG ACGACGAAGA      1080
AGTACTCGGA AGGCGCGGCG GCGGGCACTG AAGACGCCGA GCGCGCGCCA GTCGAAGCTG      1140
ACGCGGGCGG TGGGCAAAAG GTGCTGGTAC GAAACGTCGT CTAACGGCAT GCGGAGCCAA      1200

```

20

30

(2) INFORMATION FOR SEQ ID NO:29: HTCC#1

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 392 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly  
 1 5 10 15  
 Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu  
 20 25 30  
 Tyr Ser Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala  
 35 40 45  
 Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala  
 50 55 60  
 Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu  
 65 70 75 80  
 Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln  
 85 90 95  
 Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val  
 100 105 110  
 Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala  
 115 120 125  
 Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val  
 130 135 140  
 Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala  
 145 150 155 160  
 Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala  
 165 170 175  
 Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Thr  
 180 185 190  
 Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys  
 195 200 205  
 Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg  
 210 215 220  
 Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr  
 225 230 235 240  
 Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala  
 245 250 255  
 Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser  
 260 265 270  
 Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe  
 275 280 285  
 Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln  
 290 295 300  
 Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln  
 305 310 315 320  
 Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met  
 325 330 335  
 Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser  
 340 345 350  
 Lys Gly Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr  
 355 360 365  
 Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln  
 370 375 380  
 Lys Val Leu Val Arg Asn Val Val  
 385 390

10

20

30

(2) INFORMATION FOR SEQ ID NO:30: MTCC#2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1441 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GAGGTTGCTG GCAATGGATT TCGGGCTTTT ACCTCCGGAA GTGAATTCAA GCCGAATGTA 60  
 TTCCGGTCCG GGGCCGGAGT CGATGCTAGC CGCCGCGGCC GCCTGGGACG GTGTGGCCGC 120  
 GGAGTTGACT TCCGCCCGGG TCTCGTATGG ATCGGTGGTG TCGACGCTGA TCGTTGAGCC 180  
 GTGGATGGGG CCGGCGCGGG CCGCGATGCG GGCCGCGGCA ACGCCGTATG TGGGGTGGCT 240

40

GGCCGCCACG GCGGCGCTGG CGAAGGAGAC GGCCACACAG GCGAGGGCAG CCGCGGAAGC 300  
 GTTTGGGACG GCGTTCGCGA TGACCGTGCC ACCATCCCTC GTCGCGGCCA ACCGCGGCCG 360  
 GTTGATGTCG CTGGTCGCGG CGAACATTCT GGGCAAAAAC AGTGCGGCGA TCGCGGCTAC 420  
 CCAGGCCGAG TATGCGGAAA TGTTGGGCCA AGACGCTGCC GTGATGTACA GCTATGAGGG 480  
 GGCATCTGCG GCCGCGTCCG CFTTGCCGCC GTTCACTCCA CCCGTGCAAG GCACCGGCC 540  
 GCGCGGGCCC GCGGCGCGAG CCGCGGCGAC CCAAGCCGCC GGTGCGGGCG CCGTTGCGGA 600  
 TGACAGGCGG ACGTGGCCCC AGCTGCCCCC GGGGATCCTG AGCGACATTC TGTCCGCATT 660  
 GGCCGCCAAC GCTGATCCCG TGACATCGGG ACTGTTGGGG ATCGCGTCCA CCTCAACCC 720  
 GCAAGTCGGA TCCGCTCAGC CGATAGTGAT CCCACCCCG ATAGGGGAAT TGGACGTGAT 780  
 CGCCTCTAC ATTGCATCCA TCGCGACCGG CAGCATTGCG CTCGCGATCA CGAACACGGC 840  
 CAGACCCJGG CACATCGGCC TATACGGGAA CGCCGCGGGG CTGGGACCGA CGCAGGGCCA 900  
 TCCACTGAGT TCGGCGACCG ACGAGCCGGA GCCGCATGCG GGCCCTTTCG GGGCGCGGCG 960  
 GCGGTGTCC GCGGGCGTCC GCCACGCAGC ATTAGTCGGA GCGTTGTCCG TGCCGCACAG 1020  
 CTGGACCCAG GCCGCCCCGG AGATCCAGCT CGCCGTTTCG GCAACACCCA CCTTCAGCTC 1080  
 CAGCGCCCGC GCCGACCCGA CGGCCCTAAA CCGGATGCGG CGAGGCCTGC TCAGCGGGAT 1140  
 GSCTTTGGCG AGCCTGGCCG CACGCGGCAC GACGGGCGGT GCGGCACACC GTAGCGGCAC 1200  
 CAGCACTGAC GGCCAAGAGG ACGGCCGCAA ACCCCCGGTA GTTGTGATTA GAGAGCAGCC 1260  
 GCCGCCCGGA AACCCCCCGC GGTAAAAGTC CGGCAACCGT TCGTCCCGC GCGGAAAATG 1320  
 CCTGTTGAGC GTGGCTATCC GACGGGCGGT TCACACCGCT TGTAGTAGCG TACGGCTATG 1380  
 GACGACGGTG TCTGGATTCT CCGCGGCTAT CAGAGCGATT TTGCTCGCAA CCTCAGCAAA 1440  
 G 1441

10

(2) INFORMATION FOR SEQ ID NO:31: MTCC#2

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 423 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

20

Met Asp Phe Gly Leu Leu Pro Pro Glu Val Asn Ser Ser Arg Met Tyr  
 1 5 10 15  
 Ser Gly Pro Gly Pro Glu Ser Met Leu Ala Ala Ala Ala Ala Trp Asp  
 20 25 30  
 Gly Val Ala Ala Glu Leu Thr Ser Ala Ala Val Ser Tyr Gly Ser Val  
 35 40 45  
 Val Ser Thr Leu Ile Val Glu Pro Trp Met Gly Pro Ala Ala Ala Ala  
 50 55 60  
 Met Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp Leu Ala Ala Thr Ala  
 65 70 75 80  
 Ala Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg Ala Ala Ala Glu Ala  
 85 90 95  
 Phe Gly Thr Ala Phe Ala Met Thr Val Pro Pro Ser Leu Val Ala Ala  
 100 105 110  
 Asn Arg Ser Arg Leu Met Ser Leu Val Ala Ala Asn Ile Leu Gly Gln  
 115 120 125  
 Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu Tyr Ala Glu Met Trp  
 130 135 140  
 Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu Gly Ala Ser Ala Ala  
 145 150 155 160  
 Ala Ser Ala Leu Pro Pro Phe Thr Pro Val Gln Gly Thr Gly Pro  
 165 170 175  
 Ala Gly Pro Ala Ala Ala Ala Ala Thr Gln Ala Ala Gly Ala Gly  
 180 185 190  
 Ala Val Ala Asp Ala Gln Ala Thr Leu Ala Gln Leu Pro Pro Gly Ile  
 195 200 205  
 Leu Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn Ala Asp Pro Leu Thr  
 210 215 220  
 Ser Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn Pro Gln Val Gly Ser  
 225 230 235 240  
 Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly Glu Leu Asp Val Ile  
 245 250 255  
 Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser Ile Ala Leu Ala Ile

30

40

```

                260                265                270
Thr Asn Thr Ala Arg Pro Trp His Ile Gly Leu Tyr Gly Asn Ala Gly
   275                280                285
Gly Leu Gly Pro Thr Gln Gly His Pro Leu Ser Ser Ala Thr Asp Glu
   290                295                300
Pro Glu Pro His Trp Gly Pro Phe Gly Gly Ala Ala Pro Val Ser Ala
   305                310                315
Gly Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser
   325                330                335
Trp Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro
   340                345                350
Thr Phe Ser Ser Ser Ala Gly Ala Asp Pro Thr Ala Leu Asn Gly Met
   355                360                365
Pro Ala Gly Leu Leu Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg
   370                375                380
Gly Thr Thr Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly
   385                390                395
Gln Glu Asp Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro
   405                410                415
Pro Pro Gly Asn Pro Pro Arg
   420

```

10

## (2) INFORMATION FOR SEQ ID NO:32: ESAT-6

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 154 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```

ATGACAGAGC AGCAGTGGAA TTTCGCGGGT ATCGAGGCCG CCGCAAGCGC AATCCAGGGA    60
AATGTCACGT CCATTCATTC CCTCCTTGAC GAGGGGAAGC AGTCCCTGAC CAAGCTCGCA    120
GCGGCTGGG GCGGTAGCGG TTCGGAAGCG TACC                                     154

```

20

## (2) INFORMATION FOR SEQ ID NO:33: ESAT-6

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 51 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```

Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser
 1      5      10      15
Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly
 20      25      30
Lys Gln Ser Leu Thr Lys Leu Ala Ala Trp Gly Gly Ser Gly Ser
 35      40      45
Glu Ala Tyr
 50

```

30

## (2) INFORMATION FOR SEQ ID NO:34: Tb38-1

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 327 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```

CGGCACGAGA GACCGATGCC GCTACCCCTCG CGCAGGAGGC AGGTAATTTC GAGCGGATCT 60
CCGGCGACCT GAAAACCCAG ATCGACCAGG TGGAGTCGAC GGCAGGTTTC TTGCAGGGCC 120
AGTGGCGCGG CGCGGCGGGG ACGGCCGCC AGGCCGCGGT GGTGCGCTTC CAAGAAGCAG 180
CCAATAAGCA GAAGCAGGAA CTCGACGAGA TCTCGACGAA TATTCTGTCAG GCCGGCGTCC 240
AATACTCGAG GCGCCGACGAG GAGCAGCAGC AGGCGCTGTC CTCGCAAATG GGCTTCTGAC 300
CCGCTAATAC GAAAAGAAAC GGAGCAA 327

```

(2) INFORMATION FOR SEQ ID NO:35: Tb38-1

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 95 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```

Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile
1          5          10          15
Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly
20          25          30
Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala
35          40          45
Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys Gln Lys Gln Glu Leu
50          55          60
Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg
65          70          75          80
Ala Asp Glu Glu Gln Gln Ala Leu Ser Ser Gln Met Gly Phe
85          90          95

```

(2) INFORMATION FOR SEQ ID NO:36: TbRa3

20

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 542 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```

GAATTCGGCA CGAGAGGTGA TCGACATCAT CGGGACCAGC CCCACATCCT GGGAACAGGC 60
GGCGGCGGAG GCGGTCCAGC GGGCGCGGGA TAGCGTCGAT GACATCCGCG TCGCTCGGGT 120
CATTGAGCAG GACATGGCCG TGGACAGCGC CGGCAAGATC ACCTACCGCA TCAAGCTCGA 180
AGTGTCTGTT AAGATGAGGC CGGCGCAACC GCGCTAGCAC GGGCCGGCGA GCAAGACGCA 240
AAATCGCACG GTTTCGCGTT GATTCTGTGG ATTTTGTGTC TGCTCGCCGA GGCCTACCAG 300
GGCGGCGCCA GGTCCGCGTG CTGCCGTATC CAGGCCTGCA TCGCGATTCC GCGGGCCACG 360
CCGGAGTTAA TGCTTCGCGT CGACCCGAAC TGGGCGATCC GCCGNGAGC TGATCGATGA 420
CCGTGGCCAG CCCGTCGATG CCCGAGTTGC CCGAGGAAAC GTGCTGCCAG GCCGGTAGGA 480
AGCGTCCGTA GCGGCGCGTG CTGACCGGCT CTGCCTGCGC CCTCAGTCCG GCCAGCGAGC 540
GG 542

```

30

(2) INFORMATION FOR SEQ ID NO:37: TbRa3

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 66 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```

Val Ile Asp Ile Ile Gly Thr Ser Pro Thr Ser Trp Glu Gln Ala Ala
1          5          10          15
Ala Glu Ala Val Gln Arg Ala Arg Asp Ser Val Asp Asp Ile Arg Val
20          25          30

```

40

Ala Arg Val Ile Glu Gln Asp Met Ala Val Asp Ser Ala Gly Lys Ile  
 35 40 45  
 Thr Tyr Arg Ile Lys Leu Glu Val Ser Phe Lys Met Arg Pro Ala Gln  
 50 55 60  
 Pro Arg  
 65

(2) INFORMATION FOR SEQ ID NO:38: 38 kD

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1993 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGTTCPTCGA CGGCAGGCTG GTGGAGGAAG GGCCCACCGA ACAGCTGTTC TCCTCGCCGA 60  
 AGCATGCGGA AACCGCCCGA TACGTGCGCG GACTGTGCGG GCACGTCAAG GACGCCAAGC 120  
 GCGGAAATTG AAGAGCACAG AAAGGTATGG CGTGAAAATT CGTTTGATA CGCTGTTGGC 180  
 CGTGTGACC GCTGCGCCGC TGCTGCTAGC AGCGGCGGGC TGTGGCTCGA AACCACCGAG 240  
 CGGTTCGCCT GAAACGGGCG CCGGCGCCGG TACTGTGCGC ACTACCCCGC CGTCGTCGCC 300  
 GGTGACGTTG GCGGAGACCG GTAGCACGCT GCTCTACCCG CTGTTCAACC TGTGGGGTCC 360  
 GGCCTTTCAC GAGAGGTATC CGAACCTCAC GATCACCGCT CAGGGCACCG GTTCTGGTGC 420  
 CCGGATCGCG CAGGCCGCGC CCGGACGGT CAACATGGG GCCTCCGACG CCTATCTGTC 480  
 GGAAGGTGAT ATGCCCAGCG ACAAGGGGCT GATGAACATC GCGCTAGCCA TCTCCGCTCA 540  
 GCAGGTCAAC TACAACCTGC CCGGAGTGAG CGAGCACCTC AAGCTGAACG GAAAAGTCCT 600  
 GCGGCCCATG TACCAGGGCA CCATCAA AAC CTGGGACGAC CCGCAGATCG CTGCGCTCAA 660  
 CCCCAGCGTG AACCTGCCCG GCACCGCGGT AGTTCGCTG CACCGCTCCG ACGGGTCCGG 720  
 TGACACCTTC TTGTTACCC AGTACCTGTC CAAGCAAGAT CCCGAGGCT GGGGCAAGTC 780  
 GCCCGCTTC GGCACCACCG TCGACTTCCC GCGGTGCGG GGTGCGCTGG GTGAGAACGG 840  
 CAACGCGCGC ATGTTGACCG GTTGCGCCGA GACACCGGGC TGCCTGGCCT ATATCGGCAT 900  
 CAGCTTCCTC GACCAGGCCA GTCAACGGGG ACTCGGCGAG GCCCAACTAG GCAATAGCTC 960  
 TGGCAATTC TTGTTGCCCG ACGCGCAAAG CATTCAAGCC GCGGCGGCTG GCTTCGCATC 1020  
 GAAAACCCCG GCGAACCCAG CGATTTGAT GATCGACGGG CCCGCCCCGG ACGGCTACCC 1080  
 GATCATCAAC TACGAGTACG CCATCGTCAA CAACCGCAA AAGGACGCGC CCACCGCGCA 1140  
 GACCTTGACG GCATTTCTGC ACTGGGCGAT CACCGACGGC AACAAAGCCT CGTTCCTCGA 1200  
 CCAGGTTCA TCCAGCCGC TGCCGCCCGC GGTGGTGAAG TTGTTGACG CGTTGATCGC 1260  
 GACGATTTCC AGCTAGCCTC GTTGACCACC ACGCGACAGC AACCTCCGTC GGGCCATCGG 1320  
 GCTGCTTTGC GGAGCATGCT GGCCCGTGCC GGTGAAGTCG GCCGCGCTGG CCCGGCCATC 1380  
 CGGTGTTGG GTGGGATAGG TGCGGTGATC CCGCTGCTTG CGCTGGTCTT GGTGCTGGTG 1440  
 GTGCTGGTCA TCGAGGCGAT GGGTGCATC AGGCTCAACG GGTGCAATT CTTACCGGCC 1500

10

20

30



ACCGAATGGA ATCCAGGCAA CACCTACGGC GAAACCGTTG TCACCGACGC GTCGCCCATC 1560  
 CGGTCCGGCG CTACTACGGG GCGTTGCCGC TGATCGTCGG GACGCTGGCG ACCTCGGCAA 1620  
 TCGCCCTGAT CATCGCGGTG CCGGTCTCTG TAGGAGCGGC GCTGGTGATC GTGGAACGGC 1680  
 TGCCGAAACG GTTGGCCGAG GCTGTGGGAA TAGTCTTGGA ATTGCTCGCC GGAATCCCA 1740  
 GCGTGGTGGT CGGTTTGTGG GGGGCAATGA CGTTCGGGCC GTTCATCGCT CATCACATCG 1800  
 CTCCGGTGTGAT CGTTCACAAC GCTCCCGATG TGCCGGTGTCT GAACTACTTG CGCGGCGACC 1860  
 CGGGCAACGG GGAGGGCATG TTGGTGTCCG GTCTGGTGTG GCGGGTGATG GTCGTTCCCA 1920  
 TTATCGCCAC CACCACTCAT GACCTGTTCG GGCAGGTGCC GGTGTTGCCG CGGGAGGGCG 1980  
 CGATCGGGAA TTC 1993

10

## (2) INFORMATION FOR SEQ ID NO:39: 38 kD

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 374 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Met Lys Ile Arg Leu His Thr Leu Leu Ala Val Leu Thr Ala Ala Pro  
 1 5 10 15  
 Leu Leu Leu Ala Ala Ala Gly Cys Gly Ser Lys Pro Pro Ser Gly Ser  
 20 25 30  
 Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro Ala Ser  
 35 40 45  
 Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr Pro Leu  
 50 55 60  
 Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn Val Thr  
 65 70 75 80  
 Ile Thr Ala Gln Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln Ala Ala  
 85 90 95  
 Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser Glu Gly  
 100 105 110  
 Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala Ile Ser  
 115 120 125  
 Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His Leu Lys  
 130 135 140  
 Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile Lys Thr  
 145 150 155 160  
 Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn Leu Pro  
 165 170 175  
 Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly Asp Thr  
 180 185 190  
 Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Glu Gly Trp Gly  
 195 200 205

20

30

Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val Pro Gly  
 210 215 220

Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys Ala Glu  
 225 230 235 240

Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala  
 245 250 255

Ser Gln Arg Gly Leu Gly Glu Ala Gln Leu Gly Asn Ser Ser Gly Asn  
 260 265 270

Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala Gly Phe  
 275 280 285

Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro  
 290 295 300

Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile Val Asn  
 305 310 315 320

Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu  
 325 330 335

His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val  
 340 345 350

His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu  
 355 360 365

Ile Ala Thr Ile Ser Ser  
 370

10

(2) INFORMATION FOR SEQ ID NO:40: DPEP

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 999 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGCATCACC ATCACCATCA CATGCATCAG GTGGACCCCA ACTTGACACG TCGCAAGGGA 60  
 CGATTGGCGG CACTGGCTAT CGCGGCGATG GCCAGCGCCA GCCTGGTGAC CGTTGCCGTTG 120  
 CCCGCGACCG CCAACGCCGA TCCGGAGCCA GCGCCCCCGG TACCCACAAC GGCCGCCTCG 180  
 CCGCCGTCGA CCGCTGCAGC GCCACCCGCA CCGGCGACAC CTGTTGCCCC CCCACCACCG 240  
 GCGCGCGCCA ACACGCCGAA TGCCAGCCCG GCGGATCCCA ACGCAGCACC TCCGCGCGCC 300  
 GACCCGAACG CACCGCCGCC ACCTGTTCAT GCCCCAAACG CACCCCAACC TGTCCGGATC 360  
 GACAACCCCG TTGGAGGATT CAGCTTCGCG CTGCCGTGCTG GCTGGGTGGA GTCTGACGCC 420  
 GCCCACTTCG ACTACGGTTC AGCACTCCTC AGCAAAACCA CCGGGGACCC GCCATTTCCT 480  
 GGACAGCCCG CGCCGGTGGC CAATGACACC CGTATCGTGC TCGGCCGGCT AGACCAAAG 540  
 CTTTACGCCA GCGCCGAAGC CACCGACTCC AAGGCCGCGG CCGGTTGGG CTCGGACATG 600  
 GGTGAGTTCT ATATGCCCTA CCCGGGCACC CGGATCAACC AGGAAACCGT CTCGCTCGAC 660  
 GCCAACGGGG TGTCTGGAAG CCGCTCGTAT TACGAAGTCA AGTTCAGCGA TCCGAGTAAG 720  
 CCGAACGGCC AGATCTGGAC GGGCGTAATC GGCTCGCCCG CGGCGAACGC ACCGGACGCC 780  
 GGGCCCCCTC AGCGCTGGTT TGTGGTATGG CTCGGGACCG CCAACAACCC GGTGGACAAG 840  
 GGCGCGGCCA AGGCGCTGGC CGAATCGATC CGGCCTTTGG TCGCCCCGCC GCCGGCGCCG 900  
 GCACCGGCTC CTGCAGAGCC CGCTCCGGCG CCGGCGCCCG CCGGGGAAGT CGCTCCTACC 960  
 CCGACGACAC CGACCCGCA GCGGACCTTA CCGGCCTGA 999

30

(2) INFORMATION FOR SEQ ID NO:41: DPEP

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 332 amino acids
  - (B) TYPE: amino acid

40

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met His His His His His His Met His Gln Val Asp Pro Asn Leu Thr  
1 5 10 15  
Arg Arg Lys Gly Arg Leu Ala Ala Leu Ala Ile Ala Ala Met Ala Ser  
20 25 30  
Ala Ser Leu Val Thr Val Ala Val Pro Ala Thr Ala Asn Ala Asp Pro  
35 40 45  
Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro Ser Thr  
50 55 60  
Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro Val Ala Pro Pro Pro Pro  
65 70 75 80  
Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro Gly Asp Pro Asn Ala Ala  
85 90 95  
Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro Pro Pro Val Ile Ala Pro  
100 105 110  
Asn Ala Pro Gln Pro Val Arg Ile Asp Asn Pro Val Gly Gly Phe Ser  
115 120 125  
Phe Ala Leu Pro Ala Gly Trp Val Glu Ser Asp Ala Ala His Phe Asp  
130 135 140  
Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr Gly Asp Pro Pro Phe Pro  
145 150 155 160  
Gly Gln Pro Pro Pro Val Ala Asn Asp Thr Arg Ile Val Leu Gly Arg  
165 170 175  
Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu Ala Thr Asp Ser Lys Ala  
180 185 190  
Ala Ala Arg Leu Gly Ser Asp Met Gly Glu Phe Tyr Met Pro Tyr Pro  
195 200 205  
Gly Thr Arg Ile Asn Gln Glu Thr Val Ser Leu Asp Ala Asn Gly Val  
210 215 220  
Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys Phe Ser Asp Pro Ser Lys  
225 230 235 240  
Pro Asn Gly Gln Ile Trp Thr Gly Val Ile Gly Ser Pro Ala Ala Asn  
245 250 255  
Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp Phe Val Val Trp Leu Gly  
260 265 270  
Thr Ala Asn Asn Pro Val Asp Lys Gly Ala Ala Lys Ala Leu Ala Glu  
275 280 285  
Ser Ile Arg Pro Leu Val Ala Pro Pro Pro Ala Pro Ala Pro Ala Pro  
290 295 300  
Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala Gly Glu Val Ala Pro Thr  
305 310 315 320  
Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu Pro Ala  
325 330

10

20

(2) INFORMATION FOR SEQ ID NO:42: TbH4

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 702 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CGGCACGAGG ATCGGTACCC CGCGGCATCG GCAGCTGCCG ATTCGCGGGG TTTCCCACC 60  
CGAGGAAAGC CGCTACCAGA TGGCGCTGCC GAAGTAGGGC GATCCGFTCG CGATGCCGGC 120  
ATGAACGGGC GGCATCAAAT TAGTGCAGGA ACCTTTCAGT ITAGCGACGA TAATGGCTAT 180  
AGCACTAAGG AGGATGATCC GATATGACGC AGTCGCAGAC CGTGACGGTG GATCAGCAAG 240  
AGATTTTGAA CAGGGCCAAC GAGGTGGAG CCCCAGTGGC GGACCCACCG ACTGATGTCC 300  
CCATCACACC GTGCGAACTC ACGGNGGNTA AAAACGCCGC CCAACAGNTG GTNTGTCCG 360  
CCGACAACAT GCGGGAATAC CTGGCGGCCG GTGCCAAGA GCGGCAGCGT CTGGCGACCT 420  
CGCTCGCAA CGCGGCCAAG GNGTATGGCG AGGTTGATGA GGAGGCTGCG ACCGCGCTGG 480  
ACAACGACGG CGAAGGAACT GTGCAGGCAG AATCGGCCGG GCGCGTCGGA GGGGACAGTT 540

40

```

CGGCCGAAC T AACCGATACG CCGAGGGTGG CCACGGCCGG TGAACCCAAC TTCATGGATC 600
TCAAAGAAGC GGC AAGGAAG CTCGAAACGG GCGACCAAGG CGCATCGCTC GCGCACTGNG 660
GGGATGGGTG GAACACTTNC ACCCTGACGC TGCAAGGCGA CG 702

```

(2) INFORMATION FOR SEQ ID NO:43: TbH4

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 286 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```

Gly Asp Ser Phe Trp Ala Ala Ala Asp Gln Met Ala Arg Gly Phe Val
1      5      10      15
Leu Gly Ala Thr Ala Gly Arg Thr Thr Leu Thr Gly Glu Gly Leu Gln
20     25     30
His Ala Asp Gly His Ser Leu Leu Asp Ala Thr Asn Pro Ala Val
35     40     45
Val Ala Tyr Asp Pro Ala Phe Ala Tyr Glu Ile Gly Tyr Ile Xaa Glu
50     55     60
Ser Gly Leu Ala Arg Met Cys Gly Glu Asn Pro Glu Asn Ile Phe Phe
65     70     75     80
Tyr Ile Thr Val Tyr Asn Glu Pro Tyr Val Gln Pro Pro Glu Pro Glu
85     90     95
Asn Phe Asp Pro Glu Gly Val Leu Gly Gly Ile Tyr Arg Tyr His Ala
100    105    110
Ala Thr Glu Gln Arg Thr Asn Lys Xaa Gln Ile Leu Ala Ser Gly Val
115    120    125
Ala Met Pro Ala Ala Leu Arg Ala Ala Gln Met Leu Ala Ala Glu Trp
130    135    140
Asp Val Ala Ala Asp Val Trp Ser Val Thr Ser Trp Gly Glu Leu Asn
145    150    155    160
Arg Asp Gly Val Val Ile Glu Thr Glu Lys Leu Arg His Pro Asp Arg
165    170    175
Pro Ala Gly Val Pro Tyr Val Thr Arg Ala Leu Glu Asn Ala Arg Gly
180    185    190
Pro Val Ile Ala Val Ser Asp Trp Met Arg Ala Val Pro Glu Gln Ile
195    200    205
Arg Pro Trp Val Pro Gly Thr Tyr Leu Thr Leu Gly Thr Asp Gly Phe
210    215    220
Gly Phe Ser Asp Thr Arg Pro Ala Gly Arg Arg Tyr Phe Asn Thr Asp
225    230    235    240
Ala Glu Ser Gln Val Gly Arg Gly Phe Gly Arg Gly Trp Pro Gly Arg
245    250    255
Arg Val Asn Ile Asp Pro Phe Gly Ala Gly Arg Gly Pro Pro Ala Gln
260    265    270
Leu Pro Gly Phe Asp Glu Gly Gly Gly Leu Arg Pro Xaa Lys
275    280    285

```

(2) INFORMATION FOR SEQ ID NO:44: DPPD

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 339 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

ATGAGGTTGA AGTTTGCTCG CCTGAGTACT GCGATACTGG GTTGTGCAGC GCGCCTTGTG 60
TTTCCTGCCT CGGTTGCCAG CGCAGATCCA CCTGACCCGC ATCAGCCGGA CATGACGAAA 120
GGCTATTGCC CCGGTGGCCG ATGGGGTTTT GCGGACTTGG CCGTGTGCGA CCGCGAGAAG 180

```

TACCCCGACG GCTCGTTTGG GCACCAGTGG ATGCCAACGT GGTITACCGG CCCACAGTTT 240  
 TACTTCGATT GTGTCAGCGG CCGTGAGCCC CTCCCCGGCC CGCCGCCACC GGGTGGTTGC 300  
 GGTGGGGCAA TTCCGTCCGA GCAGCCCAAC GCTCCCTGA 339

(2) INFORMATION FOR SEQ ID NO:45: DPED

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 112 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Met Lys Leu Lys Phe Ala Arg Leu Ser Thr Ala Ile Leu Gly Cys Ala  
 1 5 10 15  
 Ala Ala Leu Val Phe Pro Ala Ser Val Ala Ser Ala Asp Pro Pro Asp  
 20 25 30  
 Pro His Gln Pro Asp Met Thr Lys Gly Tyr Cys Pro Gly Gly Arg Trp  
 35 40 45  
 Gly Phe Gly Asp Leu Ala Val Cys Asp Gly Glu Lys Tyr Pro Asp Gly  
 50 55 60  
 Ser Phe Trp His Gln Trp Met Gln Thr Trp Phe Thr Gly Pro Gln Phe  
 65 70 75 80  
 Tyr Phe Asp Cys Val Ser Gly Gly Glu Pro Leu Pro Gly Pro Pro Pro  
 85 90 95  
 Pro Gly Gly Cys Gly Gly Ala Ile Pro Ser Glu Gln Pro Asn Ala Pro  
 100 105 110

10

<210> SEQ ID NO:46  
 <211> 921  
 <212> DNA  
 <213> Artificial Sequence  
 <223> Description of Artificial Sequence:tri-fusion  
 protein DPV-MTI-MSL (designated Mtb31f)  
 <222> (1)..(900)

20

cat atg cat cac cat cac cat cac gat ccc gtg gac gcg gtc att aac 48  
 His Met His His His His His Asp Pro Val Asp Ala Val Ile Asn  
 1 5 10 15  
 acc acc tgc aat tac ggg cag gta gta gct gcg ctc aac gcg acg gat 96  
 Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu Asn Ala Thr Asp  
 20 25 30  
 ccg ggg gct gcc gca cag ttc aac gcc tca ccg gtg gcg cag tcc tat 144  
 Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val Ala Gln Ser Tyr  
 35 40 45  
 ttg cgc aat ttc ctc gcc gca ccg cca cct cag cgc gct gcc atg gcc 192  
 Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg Ala Ala Met Ala  
 50 55 60  
 gcg caa ttg caa gct gtg ccg ggg gcg gca cag tac atc ggc ctt gtc 240  
 Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr Ile Gly Leu Val  
 65 70 75 80  
 gag tcg gtt gcc gcc tcc tgc aac aac tat gag ctc atg acg att aat 288  
 Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr Glu Leu Met Thr Ile Asn  
 85 90 95  
 tac cag ttc ggg gac gtc gac gct cat ggc gcc atg atc cgc gct cag 336  
 Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met Ile Arg Ala Gln  
 100 105 110

30

40

gcg gcg tcg ctt gag gcg gag cat cag gcc atc gtt cgt gat gtg ttg 384  
Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val Arg Asp Val Leu  
115 120 125

gcc gcg ggt gac ttt tgg ggc ggc gcc ggt tcg gtg gct tgc cag gag 432  
Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val Ala Cys Gln Glu  
130 135 140

ttc att acc cag ttg ggc cgt aac ttc cag gtg atc tac gag cag gcc 480  
Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile Tyr Glu Gln Ala  
145 150 155 160

aac gcc cac ggg cag aag gtg cag gct gcc ggc aac aac atg gcg caa 528  
Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn Asn Met Ala Gln  
165 170 175

acc gac agc gcc gtc ggc tcc agc tgg gcc act agt atg agc ctt ttg 576  
Thr Asp Ser Ala Val Gly Ser Ser Trp Ala Thr Ser Met Ser Leu Leu  
180 185 190

gat gct cat atc cca cag ttg gtg gcc tcc cag tcg gcg ttt gcc gcc 624  
Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser Ala Phe Ala Ala  
195 200 205

aag gcg ggg ctg atg cgg cac acg atc ggt cag gcc gag cag gcg gcg 672  
Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala Glu Gln Ala Ala  
210 215 220

atg tcg gct cag gcg ttt cac cag ggg gag tcg tcg gcg gcg ttt cag 720  
Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser Ala Ala Phe Gln  
225 230 235 240

gcc gcc cat gcc cgg ttt gtg gcg gcg gcc gcc aaa gtc aac acc ttg 768  
Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys Val Asn Thr Leu  
245 250 255

ttg gat gtc gcg cag gcg aat ctg ggt gag gcc gcc ggt acc tat gtg 816  
Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala Gly Thr Tyr Val  
260 265 270

gcc gcc gat gct gcg gcc gcg tcg acc tat acc ggg ttc gat atc cat 864  
Ala Ala Asp Ala Ala Ala Ala Ser Thr Tyr Thr Gly Phe Asp Ile His  
275 280 285

cac act ggc ggc cgc tcg agc aga tcc ggc tgc taacaaagcc cgaaaggaag 917  
His Thr Gly Gly Arg Ser Ser Arg Ser Gly Cys  
290 295

ctga 921

10

20

30

<210> SEQ ID NO:47  
<211> 299  
<212> PRT  
<213> Artificial Sequence  
<223> Description of Artificial Sequence:tri-fusion  
protein DPV-MTI-MSL (designated Mtb31f)

His Met His His His His His His Asp Pro Val Asp Ala Val Ile Asn  
1 5 10 15

Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu Asn Ala Thr Asp  
20 25 30

Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val Ala Gln Ser Tyr  
35 40 45

Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg Ala Ala Met Ala  
50 55 60

Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr Ile Gly Leu Val  
65 70 75 80

Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr Glu Leu Met Thr Ile Asn  
85 90 95

Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met Ile Arg Ala Gln  
100 105 110

Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val Arg Asp Val Leu  
115 120 125

Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val Ala Cys Gln Glu  
130 135 140

Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile Tyr Glu Gln Ala  
145 150 155 160

Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn Asn Met Ala Gln  
165 170 175

Thr Asp Ser Ala Val Gly Ser Ser Trp Ala Thr Ser Met Ser Leu Leu  
180 185 190

Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser Ala Phe Ala Ala  
195 200 205

Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala Glu Gln Ala Ala  
210 215 220

Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser Ala Ala Phe Gln  
225 230 235 240

Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys Val Asn Thr Leu  
245 250 255

Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala Gly Thr Tyr Val  
260 265 270

Ala Ala Asp Ala Ala Ala Ala Ser Thr Tyr Thr Gly Phe Asp Ile His  
275 280 285

His Thr Gly Gly Arg Ser Ser Arg Ser Gly Cys  
290 295

10

20

<210> SEQ ID NO:48  
<211> 2168  
<212> DNA  
<213> Artificial Sequence  
<223> Description of Artificial Sequence:tetra-fusion  
protein DPV-MTI-MSL-MTCC2 (designated Mtb71f)  
<222> (1)..(2133)

30

cat atg cat cac cat cac cat cac gat ccc gtg gac gcg gtc att aac 48  
His Met His His His His His His Asp Pro Val Asp Ala Val Ile Asn  
1 5 10 15

acc acc tgc aat tac ggg cag gta gta gct gcg ctc aac gcg acg gat 96  
Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu Asn Ala Thr Asp  
20 25 30

ccg ggg gct gcc gca cag ttc aac gcc tca ccg gtg gcg cag tcc tat 144  
Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val Ala Gln Ser Tyr  
35 40 45

40

ttg cgc aat ttc ctc gcc gca ccg cca cct cag cgc gct gcc atg gcc 192  
 Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg Ala Ala Met Ala  
 50 55 60

gcg caa ttg caa gct gtg ccg ggg gcg gca cag tac atc ggc ctt gtc 240  
 Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr Ile Gly Leu Val  
 65 70 75 80

gag tcg gtt gcc ggc tcc tgc aac aac tat gag ctc atg acg att aat 288  
 Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr Glu Leu Met Thr Ile Asn  
 85 90 95

tac cag ttc ggg gac gtc gac gct cat ggc gcc atg atc cgc gct cag 336  
 Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met Ile Arg Ala Gln  
 100 105 110

gcg gcg tcg ctt gag gcg gag cat cag gcc atc gtt cgt gat gtg ttg 384  
 Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val Arg Asp Val Leu  
 115 120 125

gcc gcg ggt gac ttt tgg ggc ggc gcc ggt tcg gtg gct tgc cag gag 432  
 Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val Ala Cys Gln Glu  
 130 135 140

ttc att acc cag ttg ggc cgt aac ttc cag gtg atc tac gag cag gcc 480  
 Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile Tyr Glu Gln Ala  
 145 150 155 160

aac gcc cac ggg cag aag gtg cag gct gcc ggc aac aac atg gcg caa 528  
 Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn Asn Met Ala Gln  
 165 170 175

acc gac agc gcc gtc ggc tcc agc tgg gcc act agt atg agc ctt ttg 576  
 Thr Asp Ser Ala Val Gly Ser Ser Trp Ala Thr Ser Met Ser Leu Leu  
 180 185 190

gat gct cat atc cca cag ttg gtg gcc tcc cag tcg gcg ttt gcc gcc 624  
 Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser Ala Phe Ala Ala  
 195 200 205

aag gcg ggg ctg atg cgg cac acg atc ggt cag gcc gag cag gcg gcg 672  
 Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala Glu Gln Ala Ala  
 210 215 220

atg tcg gct cag gcg ttt cac cag ggg gag tcg tcg gcg gcg ttt cag 720  
 Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser Ala Ala Phe Gln  
 225 230 235 240

gcc gcc cat gcc cgg ttt gtg gcg gcg gcc gcc aaa gtc aac acc ttg 768  
 Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys Val Asn Thr Leu  
 245 250 255

ttg gat gtc gcg cag gcg aat ctg ggt gag gcc gcc ggt acc tat gtg 816  
 Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala Gly Thr Tyr Val  
 260 265 270

gcc gcc gat gct gcg gcc gcg tcg acc tat acc ggg ttc gat atc atg 864  
 Ala Ala Asp Ala Ala Ala Ala Ser Thr Tyr Thr Gly Phe Asp Ile Met  
 275 280 285

gat ttc ggg ctt tta cct ccg gaa gtg aat tca agc cga atg tat tcc 912  
 Asp Phe Gly Leu Leu Pro Pro Glu Val Asn Ser Ser Arg Met Tyr Ser  
 290 295 300

ggt ccg ggg ccg gag tcg atg cta gcc gcc gcg gcc gcc tgg gac ggt 960  
 Gly Pro Gly Pro Glu Ser Met Leu Ala Ala Ala Ala Ala Trp Asp Gly  
 305 310 315 320

10

20

30



gtg gcc gcg gag ttg act tcc gcc gcg gtc tcg tat gga tcg gtg gtg 1008  
Val Ala Ala Glu Leu Thr Ser Ala Ala Val Ser Tyr Gly Ser Val Val  
325 330 335

tcg acg ctg atc gtt gag ccg tgg atg ggg ccg gcg gcc gcg atg 1056  
Ser Thr Leu Ile Val Glu Pro Trp Met Gly Pro Ala Ala Ala Ala Met  
340 345 350

gcg gcc gcg gca acg ccg tat gtg ggg tgg ctg gcc gcc acg gcg gcg 1104  
Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp Leu Ala Ala Thr Ala Ala  
355 360 365

ctg gcg aag gag acg gcc aca cag gcg agg gca gcg gcg gaa gcg ttt 1152  
Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg Ala Ala Ala Glu Ala Phe  
370 375 380

ggg acg gcg ttc gcg atg acg gtg cca cca tcc ctc gtc gcg gcc aac 1200  
Gly Thr Ala Phe Ala Met Thr Val Pro Pro Ser Leu Val Ala Ala Asn  
385 390 395 400

cgc agc ccg ttg atg tcg ctg gtc gcg gcg aac att ctg ggg caa aac 1248  
Arg Ser Arg Leu Met Ser Leu Val Ala Ala Asn Ile Leu Gly Gln Asn  
405 410 415

agt gcg gcg atc gcg gct acc cag gcc gag tat gcc gaa atg tgg gcc 1296  
Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu Tyr Ala Glu Met Trp Ala  
420 425 430

caa gac gct gcc gtg atg tac agc tat gag ggg gca tct gcg gcc gcg 1344  
Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu Gly Ala Ser Ala Ala Ala  
435 440 445

tcg gcg ttg ccg ccg ttc act cca ccc gtg caa ggc acc ggc ccg gcc 1392  
Ser Ala Leu Pro Pro Phe Thr Pro Pro Val Gln Gly Thr Gly Pro Ala  
450 455 460

ggg ccc gcg gcc gca gcc gcg gcg acc caa gcc gcc ggt gcg ggc gcc 1440  
Gly Pro Ala Ala Ala Ala Ala Ala Thr Gln Ala Ala Gly Ala Gly Ala  
465 470 475 480

gtt gcg gat gca cag gcg aca ctg gcc cag ctg ccc ccg ggg atc ctg 1488  
Val Ala Asp Ala Gln Ala Thr Leu Ala Gln Leu Pro Pro Gly Ile Leu  
485 490 495

agc gac att ctg tcc gca ttg gcc gcc aac gct gat ccg ctg aca tcg 1536  
Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn Ala Asp Pro Leu Thr Ser  
500 505 510

gga ctg ttg ggg atc gcg tcg acc ctc aac ccg caa gtc gga tcc gct 1584  
Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn Pro Gln Val Gly Ser Ala  
515 520 525

cag ccg ata gtg atc ccc acc ccg ata ggg gaa ttg gac gtg atc gcg 1632  
Gln Pro Ile Val Ile Pro Thr Pro Ile Gly Glu Leu Asp Val Ile Ala  
530 535 540

ctc tac att gca tcc atc gcg acc ggc agc att gcg ctc gcg atc acg 1680  
Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser Ile Ala Leu Ala Ile Thr  
545 550 555 560

aac acg gcc aga ccc tgg cac atc ggc cta tac ggg aac gcc ggc ggg 1728  
Asn Thr Ala Arg Pro Trp His Ile Gly Leu Tyr Gly Asn Ala Gly Gly  
565 570 575

ctg gga ccg acg cag ggc cat cca ctg agt tcg gcg acc gac gag ccg 1776  
Leu Gly Pro Thr Gln Gly His Pro Leu Ser Ser Ala Thr Asp Glu Pro  
580 585 590

10

20

30

gag ccg cac tgg ggc ccc ttc ggg ggc gcg gcg ccg gtg tcc gcg ggc 1824  
 Glu Pro His Trp Gly Pro Phe Gly Gly Ala Ala Pro Val Ser Ala Gly  
 595 600 605

gtc ggc cac gca gca tta gtc gga gcg ttg tcg gtg ccg cac agc tgg 1872  
 Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser Trp  
 610 615 620

acc acg gcc gcc ccg gag atc cag ctc gcc gtt cag gca aca ccc acc 1920  
 Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro Thr  
 625 630 635 640

ttc agc tcc agc gcc ggc gcc gac ccg acg gcc cta aac ggg atg ccg 1968  
 Phe Ser Ser Ser Ala Gly Ala Asp Pro Thr Ala Leu Asn Gly Met Pro  
 645 650 655

gca ggc ctg ctc agc ggg atg gct ttg gcg agc ctg gcc gca cgc ggc 2016  
 Ala Gly Leu Leu Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg Gly  
 660 665 670

acg acg ggc ggt ggc ggc acc cgt agc ggc acc agc act gac ggc caa 2064  
 Thr Thr Gly Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly Gln  
 675 680 685

gag gac ggc cgc aaa ccc ccg gta gtt gtg att aga gag cag ccg ccg 2112  
 Glu Asp Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro Pro  
 690 695 700

ccc gga aac ccc ccg ccg taagatttct aaatccatca cactggcggc cgctcgag 2168  
 Pro Gly Asn Pro Pro Arg  
 705 710

10

<210> SEQ ID NO:49  
 <211> 710  
 <212> PRT  
 <213> Artificial Sequence  
 <223> Description of Artificial Sequence:tetra-fusion  
 protein DPV-MTI-MSL-MTCC2 (designated Mtb71f)

20

His Met His His His His His Asp Pro Val Asp Ala Val Ile Asn  
 1 5 10 15

Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu Asn Ala Thr Asp  
 20 25 30

Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val Ala Gln Ser Tyr  
 35 40 45

Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg Ala Ala Met Ala  
 50 55 60

Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr Ile Gly Leu Val  
 65 70 75 80

Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr Glu Leu Met Thr Ile Asn  
 85 90 95

Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met Ile Arg Ala Gln  
 100 105 110

Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val Arg Asp Val Leu  
 115 120 125

Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val Ala Cys Gln Glu  
 130 135 140

30

Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile Tyr Glu Gln Ala  
 145 150 155 160  
 Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn Asn Met Ala Gln  
 165 170 175  
 Thr Asp Ser Ala Val Gly Ser Ser Trp Ala Thr Ser Met Ser Leu Leu  
 180 185 190  
 Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser Ala Phe Ala Ala  
 195 200 205  
 Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala Glu Gln Ala Ala  
 210 215 220  
 Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser Ala Ala Phe Gln  
 225 230 235 240  
 Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys Val Asn Thr Leu  
 245 250 255  
 Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala Gly Thr Tyr Val  
 260 265 270  
 Ala Ala Asp Ala Ala Ala Ala Ser Thr Tyr Thr Gly Phe Asp Ile Met  
 275 280 285  
 Asp Phe Gly Leu Leu Pro Pro Glu Val Asn Ser Ser Arg Met Tyr Ser  
 290 295 300  
 Gly Pro Gly Pro Glu Ser Met Leu Ala Ala Ala Ala Ala Trp Asp Gly  
 305 310 315 320  
 Val Ala Ala Glu Leu Thr Ser Ala Ala Val Ser Tyr Gly Ser Val Val  
 325 330 335  
 Ser Thr Leu Ile Val Glu Pro Trp Met Gly Pro Ala Ala Ala Ala Met  
 340 345 350  
 Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp Leu Ala Ala Thr Ala Ala  
 355 360 365  
 Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg Ala Ala Glu Ala Phe  
 370 375 380  
 Gly Thr Ala Phe Ala Met Thr Val Pro Pro Ser Leu Val Ala Ala Asn  
 385 390 395 400  
 Arg Ser Arg Leu Met Ser Leu Val Ala Ala Asn Ile Leu Gly Gln Asn  
 405 410 415  
 Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu Tyr Ala Glu Met Trp Ala  
 420 425 430  
 Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu Gly Ala Ser Ala Ala Ala  
 435 440 445  
 Ser Ala Leu Pro Pro Phe Thr Pro Pro Val Gln Gly Thr Gly Pro Ala  
 450 455 460  
 Gly Pro Ala Ala Ala Ala Ala Thr Gln Ala Ala Gly Ala Gly Ala  
 465 470 475 480  
 Val Ala Asp Ala Gln Ala Thr Leu Ala Gln Leu Pro Pro Gly Ile Leu  
 485 490 495  
 Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn Ala Asp Pro Leu Thr Ser  
 500 505 510

10

20

30

Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn Pro Gln Val Gly Ser Ala  
 515 520 525

Gln Pro Ile Val Ile Pro Thr Pro Ile Gly Glu Leu Asp Val Ile Ala  
 530 535 540

Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser Ile Ala Leu Ala Ile Thr  
 545 550 555 560

Asn Thr Ala Arg Pro Trp His Ile Gly Leu Tyr Gly Asn Ala Gly Gly  
 565 570 575

Leu Gly Pro Thr Gln Gly His Pro Leu Ser Ser Ala Thr Asp Glu Pro  
 580 585 590

Glu Pro His Trp Gly Pro Phe Gly Gly Ala Ala Pro Val Ser Ala Gly  
 595 600 605

Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser Trp  
 610 615 620

Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro Thr  
 625 630 635 640

Phe Ser Ser Ser Ala Gly Ala Asp Pro Thr Ala Leu Asn Gly Met Pro  
 645 650 655

Ala Gly Leu Leu Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg Gly  
 660 665 670

Thr Thr Gly Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly Gln  
 675 680 685

Glu Asp Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro Pro  
 690 695 700

Pro Gly Asn Pro Pro Arg  
 705 710

10

20

【 図 1 】

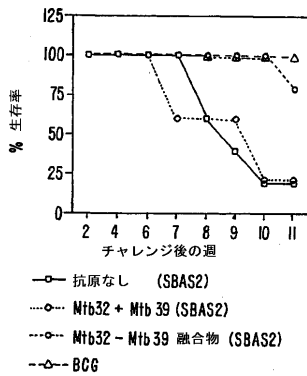


FIG. 1.

【 図 2 】

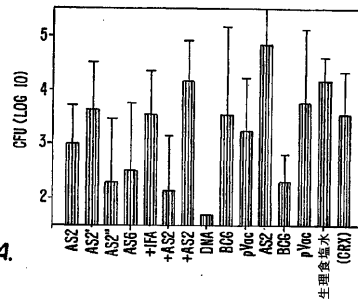


FIG. 2A.

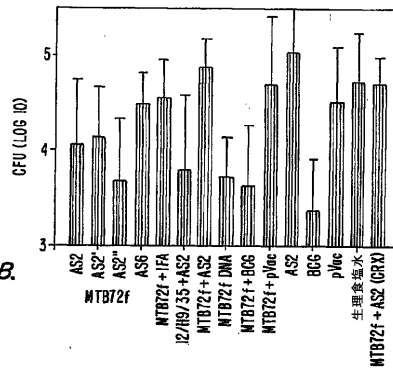


FIG. 2B.

【 3 】

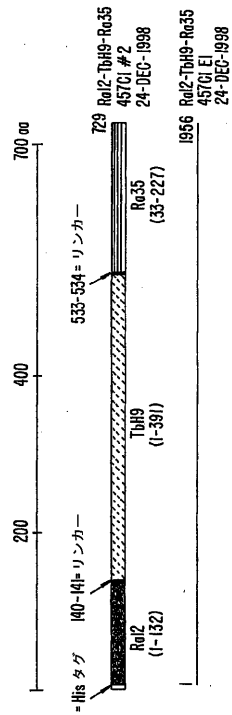


FIG. 3.

Ile Gly Gly Val Ala Val Cys Glu Pro Val Val Ala Met Cys Asn Ser Gly Gly Gln Gly Gly 115 120 125 130

Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu 135 140 145 150

Thr Gly Ala Cys Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Cys Asp Ser 155 160 165 170 175

Gly Gly Pro Val Val Asn Gly Cys Gln Val Val Cys Met Asn Thr Ala Ala Ser 180 185 190 195

FIG. 4. (続き)

【 4 】

Ra35 N末端 DNA

gccaccctgg cctttgtcncg agaccagttc accgacttcc ccgctctacc cctcaccg ccgctgctgg 70

tggccaagt ggggcaacag atgttcaaca tcaaccaca accgggttac aacaagccg tgggtgcccg 140

gaccgctc gctaccgac ccaaccgggt ggtctgacc acaaccacg tgaatggcg cgtaccacag 210

accatcgt tcaagctcg cctccgcaa acctaccgac tcatgtggt tgggtatgac cgtaccacag 280

atgctcgggt gctgcaatg cgtggcccg gttccctacc atcggggcg atcggggcg gctcggcgg 330

tgttgagccc tggctcaga tgggcaacag cgtggggacg gggggaacg cctcggcgt gctcggcgg 420

tgctcagct tgggcaaac cgtgcaagcg tggatttgc tgaaccgag cgaagagaca ttgaaaggt 490

ggtcgtatg aacaagcgg cgtccrag 588

Ra35 N末端アミノ酸配列

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu Pro Leu Asp Pro Ser Ala 5 10 15 20

Met Val Ala Gln Val Gly Pro Gln Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val 25 30 35 40

Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val Ile Ala 45 50 55 60 65

Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly 70 75 80 85

Tyr Asp Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Leu Pro Ser Ala Ala 90 95 100 105 110

FIG. 4.

【 5 】

Ra12

1 MHHHHHTAASDNFOLSGGGGFAIPICQAMATAQIRSGGGSPVHHIGTAPLG Wtb72f

1 MHHHHHTAASDNFOLSGGGGFAIPICQAMATAQIRSGGGSPVHHIGTAPLG Wtb72f-mutSA

56 LGVVDNNGGARVQRVVGSAFPAASLIGTGDVITAVDGPINSATAMADALNGHH Wtb72f

56 LGVVDNNGGARVQRVVGSAFPAASLIGTGDVITAVDGPINSATAMADALNGHH Wtb72f-mutSA

111 PGDVISVWTKSFFTRFNVTLAEGPPEFVWDEGALPFEINRMYAGPSSAS Wtb72f

111 PGDVISVWTKSFFTRFNVTLAEGPPEFVWDEGALPFEINRMYAGPSSAS Wtb72f-mutSA

166 LVAAQQMWDVSDLFSAASAFQSVVWGLTVGSWIGSAGLMVAAAFYVAMSV Wtb72f

166 LVAAQQMWDVSDLFSAASAFQSVVWGLTVGSWIGSAGLMVAAAFYVAMSV Wtb72f-mutSA

221 TAGQBELTAAQVRVAAAYETAYGLTVPVIAENRAELMILITNLLGONTPAI Wtb72f

221 TAGQBELTAAQVRVAAAYETAYGLTVPVIAENRAELMILITNLLGONTPAI Wtb72f-mutSA

276 AVNEAEYGEWQAQDAAMFGYAAATATATATLILPFEAPRMTSAGGLLEQAAVE Wtb72f

276 AVNEAEYGEWQAQDAAMFGYAAATATATATLILPFEAPRMTSAGGLLEQAAVE Wtb72f-mutSA

331 EASDTAAANQMNVPQALQAOPTQCTPSSKLGGLWKTYSRHSPISNMYSM Wtb72f

331 EASDTAAANQMNVPQALQAOPTQCTPSSKLGGLWKTYSRHSPISNMYSM Wtb72f-mutSA

386 ANNHSMINSGVSMNTLISMLKGFAPARAQVTAQNGVRMSSIGSSLGSS Wtb72f

386 ANNHSMINSGVSMNTLISMLKGFAPARAQVTAQNGVRMSSIGSSLGSS Wtb72f-mutSA

441 GLGGVYANLGRAASVGSLSVPOVAAAANQAVTPAARALPLTSAEREGQM Wtb72f

441 GLGGVYANLGRAASVGSLSVPOVAAAANQAVTPAARALPLTSAEREGQM Wtb72f-mutSA

FIG. 5.

Ra35  
 496 LGGLFVGMGARAGGGLSVLRVFRPVPVMPHSPAAGDPPALSDQDRFADFFAL Mtb72f  
 496 LGGLFVGMGARAGGGLSVLRVFRPVPVMPHSPAAGDPPALSDQDRFADFFAL Mtb72f-mutSA  
 551 PLDPAMVAQVGPQVNNIKLGYNNAYGAGTGIVIDPENGVLVLTNNHVIAGATDI Mtb72f  
 551 PLDPAMVAQVGPQVNNIKLGYNNAYGAGTGIVIDPENGVLVLTNNHVIAGATDI Mtb72f-mutSA  
 606 NAFSVGSGQTYGVVGVDRVQVAVLQIRGAGGLPFAAIGGGVAVGPPVVMGN Mtb72f  
 606 NAFSVGSGQTYGVVGVDRVQVAVLQIRGAGGLPFAAIGGGVAVGPPVVMGN Mtb72f-mutSA  
 661 SGQGGTFRVAVGQVQASDSLTAETLNLGLIQFDAAIQPQDGGGPV Mtb72f  
 661 SGQGGTFRVAVGQVQASDSLTAETLNLGLIQFDAAIQPQDGGGPV Mtb72f-mutSA  
 716 NGLGQVVMNTAAAS  
 716 NGLGQVVMNTAAAS

FIG. 5. (続き)

【 6 】

Ra35 N株編  
 1 MHHHHHPPALSDQDRFADFFALPLDPAMVAQVGPQVNNIKLGYNNA Tbra35\_mat  
 1 MHHHHHPPALSDQDRFADFFALPLDPAMVAQVGPQVNNIKLGYNNA Tbra35\_mutSA  
 51 VGAGTGIVDPNGVVLVLTNNHVIAGATDINAFSVGSGQTYGVVGVDRVQ Tbra35\_mat  
 51 VGAGTGIVDPNGVVLVLTNNHVIAGATDINAFSVGSGQTYGVVGVDRVQ Tbra35\_mutSA  
 101 DVAVLQIRGAGGLPFAAIGGGVAVGPPVVMGNSSGGGTPRAVGRVVA Tbra35\_mat  
 101 DVAVLQIRGAGGLPFAAIGGGVAVGPPVVMGNSSGGGTPRAVGRVVA Tbra35\_mutSA  
 Ra12 C株編  
 151 LGQTVQASDSLTAETLNLGLIQFDAAIQPQDGGGPVVMNGLGQVGMNTA Tbra35\_mat  
 151 LGQTVQASDSLTAETLNLGLIQFDAAIQPQDGGGPVVMNGLGQVGMNTA Tbra35\_mutSA  
 Ra35 N株編  
 201 ASDNFOLSQGGGFALPIGQAMAIAGQIRSGGSPVHICPTAFILGLGVV Tbra35\_mat  
 201 ASDNFOLSQGGGFALPIGQAMAIAGQIRSGGSPVHICPTAFILGLGVV Tbra35\_mutSA  
 251 DNNNGRVRQVVGSAFAASLIGTGDVITAVDGPINSATAMADALNGH Tbra35\_mat  
 251 DNNNGRVRQVVGSAFAASLIGTGDVITAVDGPINSATAMADALNGH Tbra35\_mutSA  
 301 HPGDVISVFWQTKSGGTRTGNVTLAEGEPPA 終点 Tbra35\_mat  
 301 HPGDVISVFWQTKSGGTRTGNVTLAEGEPPA 終点 Ra12 Tbra35\_mutSA

FIG. 6.

【 7 】

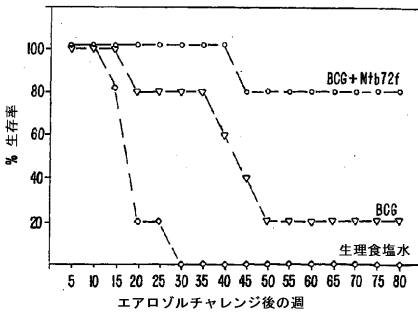
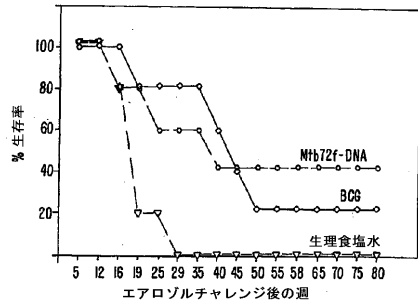


FIG. 7.

## フロントページの続き

(51)Int.Cl. F I  
 A 6 1 K 39/39 (2006.01) A 6 1 K 39/39  
 A 6 1 P 31/06 (2006.01) A 6 1 P 31/06

(72)発明者 スカイキー, ヤシル  
 アメリカ合衆国 ワシントン 9 8 1 0 7, シアトル, エヌダブリュー, 2 5 ティーエイチ  
 アベニュー 8 3 2 7

(72)発明者 リード, スティーブン  
 アメリカ合衆国 ワシントン 9 8 0 0 5, ベルビュー, 1 2 2 エヌディー プレイス エヌ  
 イー 2 8 4 3

(72)発明者 アルダーソン, マーク  
 アメリカ合衆国 ワシントン 9 8 1 1 6, ベインブリッジ アイランド, エヌダブリュー,  
 グロウ アベニュー 1 1 1 6

審査官 引地 進

(56)参考文献 特表2 0 0 2 - 5 1 0 4 9 4 ( J P , A )  
 国際公開第 0 1 / 0 2 4 8 2 0 ( W O , A 1 )  
 国際公開第 0 1 / 0 2 5 4 0 1 ( W O , A 1 )

(58)調査した分野(Int.Cl., D B 名)  
 C12N 15/00-15/90  
 CA/BIOSIS/MEDLINE/WPIDS(STN)  
 JSTPlus(JDreamII)  
 GenBank/EMBL/DDBJ/GeneSeq  
 UniProt/GeneSeq