

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-533236

(P2004-533236A)

(43) 公表日 平成16年11月4日(2004.11.4)

(51) Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A 4 B O 2 4
A 6 1 K 39/00	A 6 1 K 39/00	H 4 B O 6 3
A 6 1 K 39/39	A 6 1 K 39/39	4 B O 6 4
A 6 1 K 39/395	A 6 1 K 39/395	D 4 B O 6 5
A 6 1 K 48/00	A 6 1 K 39/395	N 4 C O 8 4
	審査請求 未請求 予備審査請求 有	(全 186 頁) 最終頁に続く

(21) 出願番号	特願2002-582198 (P2002-582198)	(71) 出願人	302042209
(86) (22) 出願日	平成14年4月12日 (2002. 4. 12)		ワイエス
(85) 翻訳文提出日	平成15年10月14日 (2003. 10. 14)		アメリカ合衆国ニュージャージー州 O 7 9
(86) 国際出願番号	PCT/US2002/011610		4 O, マジソン, ファイブ・ジラルダ・フ
(87) 国際公開番号	W02002/083859		ァームス
(87) 国際公開日	平成14年10月24日 (2002. 10. 24)	(74) 代理人	100089705
(31) 優先権主張番号	60/283, 358		弁理士 社本 一夫
(32) 優先日	平成13年4月13日 (2001. 4. 13)	(74) 代理人	100076691
(33) 優先権主張国	米国 (US)		弁理士 増井 忠武
		(74) 代理人	100075270
			弁理士 小林 泰
		(74) 代理人	100080137
			弁理士 千葉 昭男
		(74) 代理人	100096013
			弁理士 富田 博行
			最終頁に続く

(54) 【発明の名称】 化膿性連鎖球菌 (*Streptococcus pyogenes*) の表面タンパク質

(57) 【要約】

β 溶血性連鎖球菌ポリヌクレオチド、ポリペプチド、特に化膿性連鎖球菌ポリペプチドおよびポリヌクレオチド、並びにこれらのポリペプチドの抗体を記載する。本発明のポリヌクレオチド、ポリペプチド、および抗体を、免疫原組成物として使用するために配合することが可能である。やはり開示するのは、β 溶血性連鎖球菌感染に対して免疫し、そして該感染を減少させる方法、および生物学的試料中の β 溶血性連鎖球菌を検出する方法である。

【特許請求の範囲】

【請求項 1】

(i) 偶数番号の配列番号 2～668 のいずれかのアミノ酸配列に、少なくとも 70 % の同一性を有するアミノ酸配列；

(ii) 偶数番号の配列番号 2～668 のいずれかのアミノ酸配列；

(iii) (i) または (ii) のアミノ酸配列いずれかの免疫原断片；

(iv) (i) または (ii) のアミノ酸配列いずれかの少なくとも 7 の隣接するアミノ酸残基；あるいは

(v) 感受性被験者において、β溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善する (ameliorate) のに有効な、(i)、(ii)、(iii) または (iv) のいずれかの生物学的均等物
を含んでなる、単離ポリペプチド。 10

【請求項 2】

OPA 中のコロニー形成単位 (CFU) の減少によって測定される際、生物学的均等物が、陰性対照に対して、少なくとも約 15 % の殺細菌パーセントのオプソニン食作用 (opsonophagocytic) 活性を示す、請求項 1 の単離ポリペプチド。

【請求項 3】

OPA 中のコロニー形成単位 (CFU) の減少によって測定される際、生物学的均等物が、陰性対照に対して、少なくとも約 20 % の殺細菌パーセントのオプソニン食作用活性を示す、請求項 1 の単離ポリペプチド。 20

【請求項 4】

OPA 中のコロニー形成単位 (CFU) の減少によって測定される際、生物学的均等物が、陰性対照に対して、少なくとも約 40 % の殺細菌パーセントのオプソニン食作用活性を示す、請求項 1 の単離ポリペプチド。

【請求項 5】

OPA 中のコロニー形成単位 (CFU) の減少によって測定される際、生物学的均等物が、陰性対照に対して、少なくとも約 50 % の殺細菌パーセントのオプソニン食作用活性を示す、請求項 1 の単離ポリペプチド。

【請求項 6】

OPA 中のコロニー形成単位 (CFU) の減少によって測定される際、生物学的均等物が、陰性対照に対して、少なくとも約 60 % の殺細菌パーセントのオプソニン食作用活性を示す、請求項 1 の単離ポリペプチド。 30

【請求項 7】

生物学的均等物が、β溶血性連鎖球菌に対して、望ましいレベルの防御を提供する、請求項 1 の単離ポリペプチド。

【請求項 8】

(i) 偶数番号の配列番号 2～668 のいずれかのアミノ酸配列に、少なくとも 70 % の同一性を有するアミノ酸配列、(ii) 偶数番号の配列番号 2～668 のいずれかのアミノ酸配列、(iii) (i) または (ii) のアミノ酸配列いずれかの免疫原断片、あるいは (iv) (i) または (ii) のアミノ酸配列いずれかの少なくとも 7 の隣接するアミノ酸残基のいずれかを含んでなる、請求項 7 の単離ポリペプチド。 40

【請求項 9】

偶数番号の配列番号 2～668 のいずれかのアミノ酸配列に、少なくとも 85 % の同一性を有するアミノ酸配列を含んでなる、請求項 1 の単離ポリペプチド。

【請求項 10】

生物学的均等物が、β溶血性連鎖球菌の少なくとも 2 つの株に渡る交差反応性を提供する、請求項 1 の単離ポリペプチド。

【請求項 11】

アミノ酸配列が、配列番号 6、配列番号 46、配列番号 48、配列番号 80、配列番号 88 およびその組み合わせからなる群より選択される、請求項 1 の単離ポリペプチド。 50

【請求項 1 2】

アミノ酸配列が、配列番号 1 6、配列番号 6 4、配列番号 1 2 8、配列番号 1 4 0、配列番号 1 8 2 およびその組み合わせからなる群より選択される、請求項 1 の単離ポリペプチド。

【請求項 1 3】

アミノ酸配列が、配列番号 3 2、配列番号 5 8、配列番号 6 0、配列番号 1 0 4、配列番号 1 3 8 およびその組み合わせからなる群より選択される、請求項 1 の単離ポリペプチド。

【請求項 1 4】

アミノ酸配列が、配列番号 6 6、配列番号 8 2、配列番号 7 8、配列番号 1 4 2、配列番号 1 4 6、配列番号 1 6 2、配列番号 1 8 6、配列番号 3 4 2 およびその組み合わせからなる群より選択される、請求項 1 の単離ポリペプチド。 10

【請求項 1 5】

アミノ酸配列が、配列番号 2、配列番号 8、配列番号 2 0、配列番号 2 2、配列番号 2 4、配列番号 3 4、配列番号 4 0、配列番号 4 2、配列番号 4 4、配列番号 5 4、配列番号 6 2、配列番号 6 8、配列番号 9 0、配列番号 9 6、配列番号 9 8、配列番号 1 0 0、配列番号 1 0 6、配列番号 1 1 8、配列番号 1 2 4、配列番号 1 3 0、配列番号 1 4 8、配列番号 1 5 8、配列番号 3 7 6 およびその組み合わせからなる群より選択される、請求項 1 の単離ポリペプチド。

【請求項 1 6】

アミノ酸配列が配列番号 6 である、請求項 2 の単離ポリペプチド。 20

【請求項 1 7】

アミノ酸配列が配列番号 4 6 である、請求項 2 の単離ポリペプチド。

【請求項 1 8】

アミノ酸配列が配列番号 4 8 である、請求項 2 の単離ポリペプチド。

【請求項 1 9】

アミノ酸配列が配列番号 8 0 である、請求項 2 の単離ポリペプチド。

【請求項 2 0】

アミノ酸配列が配列番号 8 8 である、請求項 2 の単離ポリペプチド。

【請求項 2 1】

偶数番号の配列番号 2～6 6 8 のいずれかのアミノ酸配列の成熟ポリペプチドである、請求項 1 の単離ポリペプチド。 30

【請求項 2 2】

(i) 請求項 1 の単離ポリペプチドをコードするヌクレオチド配列；

(i i) 成熟ポリペプチドである、請求項 1 の単離ポリペプチドをコードする、ヌクレオチド配列；

(i i i) 奇数番号の配列番号 1～1 4 7 のいずれかのヌクレオチド配列；

(i v) 奇数番号の配列番号 1 4 9～1 8 1 のいずれかのヌクレオチド配列；

(v) 奇数番号の配列番号 1 8 3～1 8 7 のいずれかのヌクレオチド配列；

(v i) 奇数番号の配列番号 1 8 9～6 6 7 のいずれかのヌクレオチド配列； 40

(v i i) 請求項 1 のポリペプチドをコードするヌクレオチド配列に、少なくとも 7 0 % の同一性を有する、ヌクレオチド配列；

(v i i i) 奇数番号の配列番号 1～6 6 7 のいずれかのヌクレオチド配列に、少なくとも 7 0 % の同一性を有する、ヌクレオチド配列；

(i x) 請求項 1 のポリペプチドをコードするヌクレオチド配列に、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、ヌクレオチド配列；

(x) 奇数番号の配列番号 1～6 6 7 のいずれかのヌクレオチド配列に、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、ヌクレオチド配列；または

(x i) (i)～(x) のいずれかのヌクレオチド配列に完全に相補的であるヌクレオチド配列 50

を含んでなる、単離ポリヌクレオチド。

【請求項 23】

ヌクレオチド配列が、配列番号 5、配列番号 45、配列番号 47、配列番号 79、配列番号 87 およびその組み合わせからなる群より選択される、請求項 22 の単離ポリヌクレオチド。

【請求項 24】

ヌクレオチド配列が、配列番号 15、配列番号 63、配列番号 127、配列番号 139、配列番号 181 およびその組み合わせからなる群より選択される、請求項 22 の単離ポリヌクレオチド。

【請求項 25】

ヌクレオチド配列が、配列番号 31、配列番号 57、配列番号 59、配列番号 103、配列番号 137 およびその組み合わせからなる群より選択される、請求項 22 の単離ポリヌクレオチド。

【請求項 26】

ヌクレオチド配列が、配列番号 65、配列番号 81、配列番号 77、配列番号 141、配列番号 145、配列番号 161、配列番号 185、配列番号 341 およびその組み合わせからなる群より選択される、請求項 22 の単離ポリヌクレオチド。

【請求項 27】

ヌクレオチド配列が、配列番号 1、配列番号 7、配列番号 19、配列番号 21、配列番号 23、配列番号 33、配列番号 39、配列番号 41、配列番号 43、配列番号 53、配列番号 61、配列番号 67、配列番号 89、配列番号 95、配列番号 97、配列番号 99、配列番号 105、配列番号 117、配列番号 123、配列番号 129、配列番号 147、配列番号 157、配列番号 375 およびその組み合わせからなる群より選択される、請求項 22 の単離ポリヌクレオチド。

【請求項 28】

ヌクレオチド配列が、25、131、147、149、151、153、155、159、163、165、169、171、173、175、177、179、183、187、215、243、301、327、331、463、541、579、617、619、665、669 およびその組み合わせからなる群より選択される、請求項 22 の単離ポリヌクレオチド。

【請求項 29】

請求項 22 のポリヌクレオチドを含んでなる組換え宿主細胞。

【請求項 30】

請求項 22 のポリヌクレオチドを含んでなる組換え発現ベクター。

【請求項 31】

請求項 22 のベクターを含んでなる組換え宿主細胞。

【請求項 32】

(a) (i) 請求項 22 のポリヌクレオチドまたは (ii) 請求項 22 のポリヌクレオチドを含んでなる組換え発現ベクターを含んでなる、組換え宿主細胞を、ポリヌクレオチドにコードされるポリペプチドを産生するのに適した条件下で培養し；そして

(b) 培養からポリペプチドを回収する

ことを含んでなる、ポリペプチド産生法。

【請求項 33】

請求項 1 のポリペプチドに免疫特異的に結合する抗体。

【請求項 34】

配列番号 6、配列番号 46、配列番号 48、配列番号 80、配列番号 88 およびその組み合わせからなる群より選択されるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 33 の抗体。

【請求項 35】

配列番号 16、配列番号 64、配列番号 128、配列番号 140、配列番号 182 および

10

20

30

40

50

その組み合わせからなる群より選択されるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 33 の抗体。

【請求項 36】

配列番号 32、配列番号 58、配列番号 60、配列番号 104、配列番号 138 およびその組み合わせからなる群より選択されるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 33 の抗体。

【請求項 37】

配列番号 66、配列番号 82、配列番号 78、配列番号 142、配列番号 146、配列番号 162、配列番号 186、配列番号 342 およびその組み合わせからなる群より選択されるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 33 の抗体。

10

【請求項 38】

配列番号 2、配列番号 8、配列番号 20、配列番号 22、配列番号 24、配列番号 34、配列番号 40、配列番号 42、配列番号 44、配列番号 54、配列番号 62、配列番号 68、配列番号 90、配列番号 96、配列番号 98、配列番号 100、配列番号 106、配列番号 118、配列番号 124、配列番号 130、配列番号 148、配列番号 158、配列番号 376 およびその組み合わせからなる群より選択されるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 33 の抗体。

【請求項 39】

26、132、148、150、152、154、156、160、164、166、170、172、174、176、178、180、184、188、216、244、302、328、332、464、542、580、618、620、666、670 およびその組み合わせからなる群より選択されるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 33 の抗体。

20

【請求項 40】

配列番号 6 であるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 34 の抗体。

【請求項 41】

配列番号 46 であるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 34 の抗体。

【請求項 42】

配列番号 48 であるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 34 の抗体。

30

【請求項 43】

配列番号 80 であるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 34 の抗体。

【請求項 44】

配列番号 88 であるアミノ酸配列を有するポリペプチドに免疫特異的に結合する、請求項 34 の抗体。

【請求項 45】

請求項 1 のポリペプチドを含んでなる構成要素の免疫原量を含んでなる免疫原組成物であって、前記構成要素が、感受性哺乳動物において、 β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効な量である、前記免疫原組成物。

40

【請求項 46】

ペプチド、ポリペプチド、またはタンパク質にコンジュゲート化するかまたは連結した、前記ポリペプチドの少なくとも部分を含んでなる、請求項 45 の免疫原組成物。

【請求項 47】

多糖にコンジュゲート化するかまたは連結した、前記ポリペプチドの少なくとも部分を含んでなる、請求項 45 の免疫原組成物。

【請求項 48】

生理学的に許容しうるビヒクルをさらに含んでなる、請求項 45 の免疫原組成物。

50

【請求項 49】

有効量のアジュバントをさらに含んでなる、請求項 45 の免疫原組成物。

【請求項 50】

哺乳動物がヒト、イヌ、ウシ、ブタ、またはウマである、請求項 45 の免疫原組成物。

【請求項 51】

哺乳動物がヒトである、請求項 50 の免疫原組成物。

【請求項 52】

請求項 1 のポリペプチドを含んでなる構成要素の免疫原量を含んでなる免疫原組成物であって、ポリペプチドが、前記ポリペプチドを特異的に認識する抗体を生成することが可能であり、そして前記構成要素の量が、感受性哺乳動物において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効である、前記免疫原組成物。 10

【請求項 53】

請求項 22 のポリヌクレオチドを含んでなる構成要素の免疫原量を含んでなる免疫原組成物であって、前記構成要素が、感受性哺乳動物において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効な量である、前記免疫原組成物。

【請求項 54】

請求項 22 のポリヌクレオチドを含んでなる組換え発現ベクターを含んでなる、請求項 53 の免疫原組成物。

【請求項 55】

生理学的に許容しうるビヒクルをさらに含んでなる、請求項 53 の免疫原組成物。 20

【請求項 56】

有効量のアジュバントをさらに含んでなる、請求項 53 の免疫原組成物。

【請求項 57】

哺乳動物がヒト、イヌ、ウシ、ブタ、またはウマである、請求項 53 の免疫原組成物。

【請求項 58】

哺乳動物がヒトである、請求項 57 の免疫原組成物。

【請求項 59】

β 溶血性連鎖球菌が A 群連鎖球菌、B 群連鎖球菌、C 群連鎖球菌、または G 群連鎖球菌である、請求項 53 の免疫原組成物。

【請求項 60】

β 溶血性連鎖球菌が化膿性連鎖球菌 (*Streptococcus pyogenes*) である、請求項 59 の免疫原組成物。 30

【請求項 61】

(i) β 溶血性連鎖球菌の株に渡って実質的に保存され、そして感受性被験者において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効であり、偶数番号の配列番号 2～668 のいずれかのアミノ酸配列に、少なくとも 70% の同一性を有する、単離ポリペプチド；

(ii) (i) の免疫原断片；あるいは

(iii) (i) または (ii) に免疫特異的に結合する抗体を含んでなる、免疫原組成物。 40

【請求項 62】

単離ポリペプチドが、ヒト血清に対して試験した際、反応性に関して陽性の試験結果になる、請求項 61 の免疫原組成物。

【請求項 63】

β 溶血性連鎖球菌が A 群連鎖球菌、B 群連鎖球菌、C 群連鎖球菌、または G 群連鎖球菌である、請求項 61 の免疫原組成物。

【請求項 64】

β 溶血性連鎖球菌が化膿性連鎖球菌である、請求項 61 の免疫原組成物。

【請求項 65】

単離ポリペプチドが、株に渡って、少なくとも約 80% 保存されている、請求項 61 の免 50

疫原組成物。

【請求項 66】

単離ポリペプチドが、株に渡って、少なくとも約 85% 保存されている、請求項 61 の疫原組成物。

【請求項 67】

単離ポリペプチドが、株に渡って、少なくとも約 90% 保存されている、請求項 61 の疫原組成物。

【請求項 68】

単離ポリペプチドが、株に渡って、少なくとも約 95% 保存されている、請求項 61 の疫原組成物。

【請求項 69】

β 溶血性連鎖球菌コロニー形成または感染に対して、感受性哺乳動物を防御する方法であって、請求項 1 のポリペプチドを含んでなる構成要素の免疫原量を含んでなる免疫原組成物を哺乳動物に投与することを含んでなり、その量が、該感受性哺乳動物において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効である、前記方法。

【請求項 70】

免疫原組成物が、場合によって、ペプチド、ポリペプチド、またはタンパク質にコンジュゲート化するかまたは連結した、前記ポリペプチドの少なくとも部分を含んでなる、請求項 69 の方法。

【請求項 71】

免疫原組成物が、場合によって、多糖にコンジュゲート化するかまたは連結した、前記ポリペプチドの少なくとも部分を含んでなる、請求項 69 の方法。

【請求項 72】

β 溶血性連鎖球菌コロニー形成または感染に対して、感受性哺乳動物を防御する方法であって、請求項 1 のポリペプチドを含んでなる免疫原組成物の有効量を哺乳動物に投与することを含んでなり、ポリペプチドが前記ポリペプチドに特異的な抗体を生成することが可能であり、そしてその量が、該感受性哺乳動物において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効である、前記方法。

【請求項 73】

ポリペプチドが配列番号 2～670 のいずれかのアミノ酸配列の成熟ポリペプチドを含んでなる、請求項 72 の方法。

【請求項 74】

免疫原組成物が生理学的に許容しうるビヒクルをさらに含んでなる、請求項 72 の方法。

【請求項 75】

免疫原組成物が皮下または筋肉注射によって投与される、請求項 72 の方法。

【請求項 76】

免疫原組成物が経口摂取によって投与される、請求項 72 の方法。

【請求項 77】

免疫原組成物が鼻内投与される、請求項 72 の方法。

【請求項 78】

β 溶血性連鎖球菌が A 群連鎖球菌、B 群連鎖球菌、C 群連鎖球菌、または G 群連鎖球菌である、請求項 72 の方法。

【請求項 79】

β 溶血性連鎖球菌が化膿性連鎖球菌である、請求項 72 の方法。

【請求項 80】

哺乳動物がヒト、イヌ、ウシ、ブタ、またはウマである、請求項 72 の方法。

【請求項 81】

哺乳動物がヒトである、請求項 80 の免疫原組成物。

【請求項 82】

10

20

30

40

50

β溶血性連鎖球菌コロニー形成または感染に対して、感受性哺乳動物を防御する方法であって、請求項22のポリヌクレオチドを含んでなる免疫原組成物の有効量を哺乳動物に投与することを含んでなり、その量が、該感受性哺乳動物において、β溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効である、前記方法。

【請求項83】

前記免疫原組成物が請求項22のポリヌクレオチドを含んでなる組換え発現ベクターを含んでなる、請求項82の方法。

【請求項84】

免疫原組成物が生理学的に許容しうるビヒクルをさらに含んでなる、請求項82の方法。

【請求項85】

免疫原組成物が皮下または筋肉注射によって投与される、請求項82の方法。

【請求項86】

免疫原組成物が経口摂取によって投与される、請求項82の方法。

【請求項87】

免疫原組成物が鼻内投与される、請求項82の方法。

【請求項88】

β溶血性連鎖球菌がA群連鎖球菌、B群連鎖球菌、C群連鎖球菌、またはG群連鎖球菌である、請求項82の方法。

【請求項89】

β溶血性連鎖球菌が化膿性連鎖球菌である、請求項82の方法。

【請求項90】

哺乳動物がヒト、イヌ、ウシ、ブタ、またはウマである、請求項82の方法。

【請求項91】

哺乳動物がヒトである、請求項90の方法。

【請求項92】

β溶血性連鎖球菌コロニー形成または感染を有する哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるための組成物であって、請求項1のポリペプチドに免疫特異的に結合する抗体を含んでなる、前記組成物。

【請求項93】

β溶血性連鎖球菌コロニー形成または感染を有する哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるための組成物であって、請求項1のポリペプチドをコードするヌクレオチド配列の発現を遮断可能なアンチセンスオリゴヌクレオチドを含んでなる、前記組成物。

【請求項94】

ポリペプチドが配列番号2～668のいずれかのアミノ酸配列の成熟ポリペプチドを含んでなる、請求項93の組成物。

【請求項95】

β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるための組成物であって、請求項1のポリペプチドをコードするヌクレオチド配列の発現を遮断可能なアンチセンスオリゴヌクレオチドを含んでなる、前記組成物。

【請求項96】

ポリペプチドが配列番号2～668のいずれかのアミノ酸配列の成熟ポリペプチドを含んでなる、請求項95の組成物。

【請求項97】

β溶血性連鎖球菌コロニー形成または感染を有する哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させる方法であって、請求項1のポリペプチドに免疫特異的に結合する抗体を含んでなる組成物の有効量を哺乳動物に投与することを含んでなり、その量が、該哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるのに有効である、前記方法。

【請求項98】

10

20

30

40

50

β溶血性連鎖球菌コロニー形成または感染を有する哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させる方法であって、請求項1のポリペプチドをコードするヌクレオチド配列の発現を遮断可能なアンチセンスオリゴヌクレオチドを含んでなる組成物の有効量を哺乳動物に投与することを含んでなる、前記方法。

【請求項99】

哺乳動物において、β溶血性連鎖球菌感染によって引き起こされる副作用を減少させる方法であって、請求項1のポリペプチドに免疫特異的に結合する抗体を含んでなる組成物の有効量を哺乳動物に投与することを含んでなり、その量が、該哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるのに有効である、前記方法。

【請求項100】

哺乳動物において、β溶血性連鎖球菌感染によって引き起こされる副作用を減少させる方法であって、請求項1のポリペプチドをコードするヌクレオチド配列の発現を遮断可能なアンチセンスオリゴヌクレオチドを含んでなる組成物の有効量を哺乳動物に投与することを含んでなり、その量が、該哺乳動物において、β溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるのに有効である、前記方法。

【請求項101】

生物学的試料において、β溶血性連鎖球菌を検出し、そして／または同定する方法であって：

(a) 請求項22のポリヌクレオチドと生物学的試料を、相補塩基対のハイブリダイゼーションを可能にする条件下で接触させ；そして

(b) 試料中のハイブリダイゼーション複合体の存在を検出する、ここでハイブリダイゼーション複合体の検出が、生物学的試料にβ溶血性連鎖球菌が存在する指標となることを含んでなる、前記方法。

【請求項102】

前記β溶血性連鎖球菌が化膿性連鎖球菌である、請求項101の方法。

【請求項103】

生物学的試料において、β溶血性連鎖球菌を検出し、そして／または同定する方法であって：

(a) 請求項1のポリペプチドに免疫特異的に結合する抗体と生物学的試料を、免疫複合体形成に適した条件下で接触させ；そして

(b) 試料中の免疫複合体の存在を検出する、ここで免疫複合体の検出が、生物学的試料にβ溶血性連鎖球菌が存在する指標となることを含んでなる、前記方法。

【請求項104】

前記β溶血性連鎖球菌が化膿性連鎖球菌である、請求項103の方法。

【請求項105】

生物学的試料において、β溶血性連鎖球菌に対する抗体を検出し、そして／または同定する方法であって：

(a) 請求項1のポリペプチドと生物学的試料を、免疫複合体形成に適した条件下で接触させ；そして

(b) 試料中の免疫複合体の存在を検出する、ここで免疫複合体の検出が、生物学的試料にβ溶血性連鎖球菌に対する抗体が存在する指標となることを含んでなる、前記方法。

【請求項106】

ポリペプチドが偶数の配列番号2～670のいずれかのアミノ酸配列の成熟ポリペプチドを含んでなる、請求項105の方法。

【請求項107】

前記β溶血性連鎖球菌が化膿性連鎖球菌である、請求項105の方法。

【請求項108】

請求項1のポリペプチドを含んでなる、免疫原組成物。

10

20

30

40

50

【請求項 109】

前記ポリペプチドが偶数番号の配列番号 2～668 のアミノ酸配列の成熟ポリペプチドである、請求項 108 の免疫原組成物。

【請求項 110】

請求項 22 のポリヌクレオチドを含んでなる、免疫原組成物。

【請求項 111】

請求項 30 の発現ベクターを含んでなる、請求項 110 の免疫原組成物。

【請求項 112】

請求項 1 の単離ポリペプチドに免疫特異的に結合する抗体を含んでなる、免疫原組成物。

【請求項 113】

単離ポリペプチドが、偶数番号の配列番号 2～668 のいずれかのアミノ酸配列の成熟ポリペプチドを含んでなる、請求項 112 の免疫原組成物。

【請求項 114】

配列番号 2～668 のアミノ酸をコードするヌクレオチド配列に、少なくとも 70 % の同一性を有するヌクレオチド配列を含んでなる単離ポリヌクレオチドであって、前記ポリヌクレオチドが：

(a) 配列番号 2～668 の成熟ポリペプチドをコードするヌクレオチドから第一の PCR プライマーおよび第二の PCR プライマーを得て、ここで第一のプライマーおよび第二のプライマーは、PCR 条件下で、外側に向かう方式で、核酸合成を開始することが可能であり、そして第一のプライマーはアンチセンス方向に伸長可能であり、そして第二のプライマーはセンス方向に伸長可能である；そして

(b) 前記ポリヌクレオチドを含有する cDNA ライブラリーと、前記の第一の PCR プライマーおよび第二の PCR プライマーを、第一のプライマーおよび第二のプライマーから前記ヌクレオチド配列を合成するのに適した PCR 条件下で合わせることを含んでなる工程によって同定される、前記単離ポリヌクレオチド。

【請求項 115】

ポリメラーゼ連鎖反応 (PCR) を用いて、請求項 22 のポリヌクレオチドを伸長する方法であって：

(a) 前記ポリヌクレオチドから第一の PCR プライマーおよび第二の PCR プライマーを得て、ここで第一の PCR プライマーおよび第二の PCR プライマーは、PCR 条件下で、外側に向かう方式で、核酸合成を開始することが可能であり、そして第一の PCR プライマーはアンチセンス方向に伸長可能であり、そして第二の PCR プライマーはセンス方向に伸長可能である；そして

(b) cDNA ライブラリーに含有される前記ポリヌクレオチドと、前記の第一の PCR プライマーおよび第二の PCR プライマーを、第一の PCR プライマーおよび第二の PCR プライマーからヌクレオチド配列を合成するのに適した PCR 条件下で合わせて、それによって前記ポリヌクレオチドを伸長する

工程を含んでなる、前記方法。

【発明の詳細な説明】

【技術分野】

【0001】

発明の分野

本発明は、一般的に、β溶血性連鎖球菌ポリペプチドおよびポリヌクレオチド、特に化膿性連鎖球菌 (*Streptococcus pyogenes*) ポリペプチドおよびポリヌクレオチドに関する。より具体的には、本発明は、表面に局在する化膿性連鎖球菌ポリペプチド、およびこれらのポリペプチドの抗体に関する。本発明はまた、化膿性連鎖球菌のポリペプチドをコードするヌクレオチド配列、およびこれらのヌクレオチド配列を含む発現ベクターにも関する。本発明はさらに、β溶血性連鎖球菌感染に対して免疫し、そして該感染を減少させる免疫原組成物および方法に関する。本発明はまた、生物学的試料において、これらのヌクレオチドおよびポリペプチドを検出し、そしてβ溶血性連鎖球菌お

10

20

30

40

50

よび化膿性連鎖球菌を検出する方法にも関する。

【背景技術】

【0002】

発明の背景

連鎖球菌分類の伝統的な表現型規準には、溶血反応およびランスフィールド血清学的分類両方が含まれる。しかし、分類学が進歩して、関連しないβ溶血性（寒天プレートにおけるヒッジ赤血球の完全な溶解と定義される）連鎖球菌種が同一のランスフィールド抗原を産生する可能性があり、そして種レベルで遺伝的に関連する株が異種ランスフィールド抗原を有する可能性があることが、現在知られる。連鎖球菌分類の伝統的な規則に対するこれらの例外にもかかわらず、溶血反応およびランスフィールド血清学的試験は、臨床的単離体同定の最初の工程として、連鎖球菌を広いカテゴリーに分類するのに、今なお使用可能である。Ruoff, K. L., R. A. WhitleyおよびD. Beighton. 1999. *Streptococcus*. P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. TenoverおよびR. H. Tenover (監修), *Manual of Clinical Microbiology*. American Society of Microbiology Press, ワシントンD. C.

ランスフィールド群A、C、またはG抗原を持つβ溶血性単離体は、2つの群：大コロニー（直径>0.5mm）形成体および小コロニー（直径<0.5mm）形成体に細分可能である。大コロニー形成群A（化膿性連鎖球菌）、C、およびG株は、多様な有効病原性機構を十分に備えた「化膿性」連鎖球菌である。ストレプトコッカス・アガラクティエ（*Streptococcus agalactiae*）（B群）は、今なおランスフィールドB群抗原産生または他の表現型特質によって、信頼性をもって同定される。

【0003】

A、B、CおよびG群を含む、β溶血性連鎖球菌によって引き起こされる感染を改善（ameliorate）、そして防御する組成物および方法を開発する必要性が存在する。これらの種間の類似性には病原性因子だけでなく、疾患発症も含まれる。後者に含まれるのは、肺炎、関節炎、膿瘍、鼻咽頭炎、子宮筋層炎、産褥期敗血症、新生児敗血症、創傷感染、髄膜炎、腹膜炎、蜂巣炎、膿皮症、壊疽性筋膜炎、トキシックショック症候群、敗血症、感染性心内膜炎、心外膜周囲炎、糸球体腎炎、および骨髄炎である。

【0004】

化膿性連鎖球菌は、ヒトの咽頭および皮膚にコロニー形成するグラム陽性双球菌であり、これらの部位はその後、この生物の主な貯蔵庫になる。偏性寄生虫であるこの細菌は、呼吸分泌物の直接接触によるか、または手から口によるか、いずれかで伝染する。大部分の化膿性連鎖球菌感染は、咽頭炎または膿疱疹など、比較的穏やかな疾病である。現在、米国において、咽頭炎のみで、2000万～3500万の間の症例があり、医師の診察および他の関連する費用に約20億ドルの費用がかかっている。さらに、化膿性連鎖球菌感染からは、リウマチ熱、しょう紅熱、および糸球体腎炎などの非化膿性続発症が生じる。全世界的に、急性リウマチ熱（ARF）は、小児科心臓疾患の最も一般的な原因である（参考文献エントリー1）。

【0005】

化膿性連鎖球菌は、最初の進入口である咽頭および皮膚から、細菌が通常は見られない、血液、深筋および脂肪組織、または肺などの体の他の部分に広がることが可能であり、そして侵襲性感染を引き起こすことが可能である。最も重症であるが、最も一般的でない型の侵襲性化膿性連鎖球菌疾患の2つは、壊疽性筋膜炎および連鎖球菌トキシックショック症候群（STSS）である。壊疽性筋膜炎（メディアでは「人喰い細菌」と記載される）は、筋肉および脂肪組織の破壊性感染である。STSSは迅速に進行する感染であり、腎臓、肝臓、および肺などの内臓にショックおよび傷害を引き起こす。この損傷の多くは細菌増殖による局所損傷よりむしろ、毒素血症によるものである。

【0006】

10

20

30

40

50

1995年、侵襲性化膿性連鎖球菌感染およびSTSSは、報告義務のある疾患となった。咽頭炎および膿痂疹を得る数百万の個体とは対照的に、米国疾病対策予防センター（CDC）義務症例報告は、1997年、米国では、15,000～20,000例の侵襲性化膿性連鎖球菌疾患があり、2,000を超える死が生じたことを示す（1）。他の報告は、侵襲性疾患が、年間、100,000人あたり、10～20例と同程度に高いと概算する（62）。より具体的には、侵襲性疾患の15,000～20,000例のうち、1,100～1,500が壊疽性筋膜炎症例であり、そして1,000～1,400がSTSS症例であり、これらの死亡率はそれぞれ、20%および60%である。やはり深刻な侵襲性疾患に含まれるのは、筋炎の症例であり、これは80%～100%の致死率を持つ。侵襲性A群連鎖球菌疾患の他の型の個体のさらに10%～15%が死ぬ。症例報告は1995年に開始され、そして過去10年または20年に渡って起こった一般的な傾向を反映するため、これらの数字は増加してきている。さらに、症例定義が厳しい結果、より少数の、そしてしたがって誤解を招く数字になっており、定義を満たす前に、初期診断および治療によって、多くの症例が首尾よく解決していることが、一般的に認められている。

【0007】

化膿性連鎖球菌は、ペニシリンおよびその誘導体にきわめて感受性なままである一方、治療は必ずしも該生物を根絶しない。抗生物質療法にもかかわらず、季節に応じて、ヒト集団のおよそ5%～20%がキャリアのままでありつづける（62）。この理由は完全には明らかでなく、そして多様な機構が関与する可能性がある。深刻な侵襲性感染の症例では、治療にはしばしば、積極的な外科的介入が必要である。STSSまたは関連疾患を伴う症例では、組織によく浸透し、そして外毒素産生を防御するため、クリンダマイシン（タンパク質合成阻害剤）が好ましい抗生物質である。テトラサイクリン、サルファ剤、および最近ではエリスロマイシンに対するある程度の耐性が報告されている。明らかに、β溶血性連鎖球菌感染を防御し、そして治療するための組成物に対する必要性が依然としてある。

【0008】

化膿性連鎖球菌に関して、多くの病原性因子が同定されてきており、分泌されるものもあり、そして表面に局在するものもある。化膿性連鎖球菌は被包されているが、莢膜はヒアルロン酸で構成され、そしてヒアルロン酸は哺乳動物細胞によって一般的に発現され、そして免疫原性でないため、免疫原組成物に包含する候補抗原としては適切でない（14）。T抗原および炭水化物基が他の候補であるが、これらは、心臓組織に対する交差反応性抗体もまた引き出す可能性がある。リポタイコ酸は、化膿性連鎖球菌の表面上に存在するが、LPSと同様の安全性の懸念を引き起こす。

【0009】

最も豊富な表面タンパク質は、その構造的類似性のため、Mまたは「M様」タンパク質と称される、タンパク質ファミリーに属する。この種のメンバーは、食作用を阻害する際と同様の生物学的役割を有する一方、各々、特有の基質結合特性を有する。このファミリーで最もよく性質決定されているタンパク質は、らせん状Mタンパク質である。同族（homologous）M株に対して向けられる抗体は、オプソニン性であり、防御性であることが示されてきている（12、13、16）。候補抗原としてのMタンパク質の使用を複雑にするのは、いくつかのより型決定されていないものと共に、同定されるMタンパク質のおよそ100の異なる血清型があるという事実である。典型的には、血清型M1、M3、M6、M12、およびM18に例示されるクラスI M血清型は、咽頭炎、しょう紅熱およびリウマチ熱と関連し、そして免疫グロブリン結合タンパク質を発現しない。M2およびM49などのクラスII M血清型は、より一般的な局所皮膚感染および続発性糸球体腎炎と関連し、そして免疫グロブリン結合タンパク質を発現する（54）。M血清型に対する抗体の異種交差反応性は、あるとしても、わずかであることに注目することが重要である。同じく重要なのは、これらの抗体がリウマチ熱において果たす役割である。Mタンパク質の特定の領域は、宿主心臓組織と交差反応する抗体を引き出し、細胞損傷を引き起こすか、または少なくとも細胞損傷と関連する（11、57）。

【0010】

MおよびM様タンパク質は、ソルターゼ (s o r t a s e) に標的とされる L P X T G モチーフによって定義される表面局在タンパク質の大きなファミリーに属する (38、64)。このモチーフは、タンパク質のカルボキシ末端近傍に位置し、まず、L P X T G モチーフのスレオニンおよびグリシン残基間で、ソルターゼに切断される。切断されると、該タンパク質は、スレオニンのカルボキシルを介して、ペプチドグリカン中のアミノ酸架橋の未結合 (f r e e) アミド基に共有結合し、こうして細菌細胞表面に該タンパク質が永続的に付着する。ソルターゼに標的とされるタンパク質のこのファミリーに含まれるのは、C5aペプチダーゼ (6、7)、フィブロネクチン (9、19、23、24)、ビトロネクチン、およびIV型コラーゲンのアドヘシン類、並びにプラスミノゲン、I g A、I g G、およびアルブミンと結合する他のM様タンパク質 (31) である。 10

【0011】

多くの分泌タンパク質が記載されてきており、このうちいくつかは毒素とみなされる。深刻な侵襲性疾患および連鎖球菌トキシックショック症候群 (S T S S) の症例由来の化膿性連鎖球菌単離体の大部分は、連鎖球菌発熱性外毒素 (S P E) A および C を生じる (8)。他の発熱性外毒素もまた、オクラホマ大学で完了し、G e n B a n k に提出され、そして寄託番号 A E 0 0 4 0 9 2 を割り当てられた、化膿性連鎖球菌ゲノム配列中で同定され、そして性質決定されてきている (55)。トキシックショック様症候群毒素、連鎖球菌スーパー抗原 (58)、および分裂促進因子 (66) などの他の毒素が、疾患において、より定義されていない役割を果たす。I L - β 放出を引き起こすため、連鎖球菌溶血素 O もまた、候補抗原の可能性があるとみなすことが可能である。さらに、システインプロテアーゼ (35、37)、ストレプトキナーゼ (26、48)、およびヒアルロニダーゼ (27、28) を含む、多様な分泌酵素もまた同定されてきている。 20

【0012】

化膿性連鎖球菌に産生される既知の病原性因子の数を考慮すると、成功する β 溶血性連鎖球菌免疫原組成物の重要な性質は、感染プロセス初期にコロニー形成を防御するかまたは制限するであろう反応を刺激する能力であろうことが明らかである。この防御反応は、接着を遮断し、そして/またはオプソニン食作用 (o p s o n o p h a g o c y t i c) を通じた細胞のクリアランスを増進するか、いずれかであろう。Mタンパク質に対する抗体はオプソニン性であり、そして抗B血清型莢膜抗体が、インフルエンザ菌 (H a e m o p h i l u s i n f l u e n z a e) B に引き起こされる疾患からの防御を示す (36) のとほぼ同じ方式で、タンパク質の抗食作用特性を克服する機構を提供することが示されてきている (30)。さらに、タンパク質Fに特異的な抗体が、接着および組織培養細胞による内在化を遮断することが示されてきている (43)。 30

【0013】

β 溶血性連鎖球菌コロニー形成または感染の防御または改善のための免疫原組成物および方法をさらに同定する必要性が依然としてある。また、化膿性連鎖球菌の表面タンパク質および化膿性連鎖球菌ポリペプチドをコードするポリヌクレオチドをさらに同定する必要性も依然としてある。また、 β 溶血性連鎖球菌および化膿性連鎖球菌コロニー形成または感染を検出する方法に対する必要性も依然としてある。 40

【発明の開示】

【発明が解決しようとする課題】

【0014】

発明の概要

これらの必要性および他の必要性を満たすため、そしてこの目的に鑑みて、本発明は、 β 溶血性連鎖球菌コロニー形成または感染の防御または改善のための組成物および方法を提供する。本発明はまた、化膿性連鎖球菌ポリペプチドおよびポリヌクレオチド、組換え材料、およびその産生法も提供する。本発明の別の側面は、こうした化膿性連鎖球菌ポリペプチドおよびポリヌクレオチドの使用法に関する。

【0015】

本発明のポリペプチドには、偶数番号の配列番号 2～668 のいずれかのアミノ酸配列の少なくとも 1 つを含んでなる単離ポリペプチドが含まれる。本発明にはまた、偶数番号の配列番号 2～668 のアミノ酸配列のいずれかに、少なくとも 70 % の同一性を有するアミノ酸配列、および偶数番号の配列番号 2～668 のいずれかのアミノ酸配列の成熟ポリペプチドも含まれる。本発明にはさらに、これらのポリペプチドの免疫原断片および生物学的均等物も含まれる。やはり提供されるのは、本発明のポリペプチドに免疫特異的に結合する抗体である。

【0016】

本発明のポリヌクレオチドには、本発明のポリペプチドをコードするヌクレオチド配列を含んでなる単離ポリヌクレオチドが含まれる。これらのポリヌクレオチドには、奇数番号の配列番号 1～667 のいずれかのヌクレオチド配列の少なくとも 1 つを含んでなる単離ポリヌクレオチドが含まれ、そしてまた、遺伝暗号の縮重の結果として、やはり本発明のポリペプチドをコードする、他のヌクレオチド配列も含まれる。本発明にはまた、本発明のポリペプチドをコードするヌクレオチド配列に、少なくとも 70 % の同一性を有するヌクレオチド配列を含んでなる単離ポリヌクレオチド、および奇数番号の配列番号 1～667 のいずれかのヌクレオチド配列に、少なくとも 70 % の同一性を有するヌクレオチド配列を含んでなる単離ポリヌクレオチドも含まれる。さらに、本発明の単離ポリヌクレオチドには、本発明のポリペプチドをコードするヌクレオチド配列に、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ヌクレオチド配列、奇数番号の配列番号 1～667 のいずれかのヌクレオチド配列に、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ヌクレオチド配列、およびこれらのポリヌクレオチドに完全に相補的であるヌクレオチド配列が含まれる。さらに、本発明には、これらのポリヌクレオチドを含んでなる発現ベクターおよび宿主細胞が含まれる。

10

20

【0017】

本発明はさらに、本発明のポリペプチドを産生する方法を提供する。1 つの態様において、該方法は、(a) 本発明の組換え宿主細胞を、本発明のポリペプチドを産生するのに適した条件下で培養し、そして (b) 培養からポリペプチドを回収する工程を含んでなる。

【0018】

本発明はまた、免疫原組成物も提供する。1 つの態様において、免疫原組成物は、感受性哺乳動物において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効な量で、本発明のポリペプチドを含んでなる、少なくとも 1 つの構成要素の免疫原量を含んでなる。構成要素は、ポリペプチドのみを含んでなることも可能であるし、またはポリペプチド、および β 溶血性連鎖球菌コロニー形成または感染の防御および／または改善を補助可能な他の物質いずれか（例えば 1 以上の化学薬品剤、タンパク質など）を含んでなることも可能である。これらの免疫原組成物はさらに、場合によって、ペプチド、ポリペプチド、またはタンパク質に、あるいは多糖にコンジュゲート化するかまたは連結した、該ポリペプチドの少なくとも部分を含んでなることが可能である。別の態様において、免疫原組成物は、本発明のポリヌクレオチドを含んでなる構成要素の免疫原量を含んでなり、該構成要素は、感受性哺乳動物において、β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効な量である。該構成要素は、ポリヌクレオチドのみを含んでなることが可能であるし、またはポリヌクレオチド、および β 溶血性連鎖球菌コロニー形成または感染の防御および／または改善を補助可能な他の物質いずれか（例えば 1 以上の化学薬品剤、タンパク質など）を含んでなることも可能である。さらに別の態様において、免疫原組成物は、本発明のポリヌクレオチドを含んでなるベクターを含んでなる。本発明の免疫原組成物にはまた、有効量のアジュバントが含まれることも可能である。

30

40

【0019】

本発明にはまた、β 溶血性連鎖球菌コロニー形成または感染に対して、感受性哺乳動物を防御する方法も含まれる。1 つの態様において、該方法は、本発明のポリペプチドの免疫原量を含んでなる免疫原組成物の有効量を哺乳動物に投与することを含んでなり、その量

50

は、感受性哺乳動物において、 β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効である。別の態様において、該方法は、本発明のポリヌクレオチドを含んでなる免疫原組成物の有効量を哺乳動物に投与することを含んでなり、その量は、感受性哺乳動物において、 β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効である。本発明の免疫原組成物は、いかなる慣用的な経路によって、例えば皮下または筋肉注射、経口摂取、あるいは鼻内投与によって、投与することも可能である。

【0020】

本発明にはさらに、 β 溶血性連鎖球菌コロニー形成または感染を有する哺乳動物において、 β 溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるための組成物および方法も含まれる。1つの態様において、組成物は、本発明の抗体を含んでなる。別の態様において、組成物は、本発明のポリペプチドをコードするヌクレオチド配列の発現を遮断可能なアンチセンスオリゴヌクレオチドを含んでなる。

【0021】

やはり提供するの、哺乳動物において、 β 溶血性連鎖球菌感染によって引き起こされる副作用を減少させる方法である。1つの態様において、該方法は、本発明の抗体を含んでなる組成物の有効量を哺乳動物に投与することを含んでなり、その量は、該哺乳動物において、 β 溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるのに有効である。別の態様において、該方法は、本発明のポリペプチドをコードするヌクレオチド配列の発現を遮断可能なアンチセンスオリゴヌクレオチドを含んでなる組成物の有効量を哺乳動物に投与することを含んでなり、その量は、該哺乳動物において、 β 溶血性連鎖球菌の数および増殖の少なくとも1つを減少させるのに有効である。

【0022】

やはり提供するの、生物学的試料において、 β 溶血性連鎖球菌を検出し、そして／または同定する方法である。1つの態様において、該方法は、(a) 本発明のポリヌクレオチドと生物学的試料を、相補塩基対のハイブリダイゼーションを可能にする条件下で接触させ、そして(b) 試料中のハイブリダイゼーション複合体の存在を検出することを含んでなり、ここでハイブリダイゼーション複合体の検出が、生物学的試料に β 溶血性連鎖球菌が存在する指標となる。別の態様において、該方法は、(a) 本発明の抗体と生物学的試料を、免疫複合体形成に適した条件下で接触させ、そして(b) 試料中の免疫複合体の存在を検出することを含んでなり、ここで免疫複合体の検出が、生物学的試料に β 溶血性連鎖球菌が存在する指標となる。さらに別の態様において、該方法は、(a) 本発明のポリペプチドと生物学的試料を、免疫複合体形成に適した条件下で接触させ、そして(b) 試料中の免疫複合体の存在を検出することを含んでなり、ここで免疫複合体の検出が、生物学的試料に β 溶血性連鎖球菌に対する抗体が存在する指標となる。

【0023】

本発明はさらに、免疫原組成物を提供する。1つの態様において、免疫原組成物は、少なくとも1つの本発明のポリペプチドを含んでなる。別の態様において、免疫原組成物は、少なくとも1つの本発明のポリヌクレオチドを含んでなる。さらに別の態様において、免疫原組成物は、少なくとも1つの本発明の抗体を含んでなる。

【0024】

やはり提供するの、本発明のポリペプチドをコードするヌクレオチド配列に、少なくとも70%の同一性を有するヌクレオチド配列を含んでなる単離ポリヌクレオチドであって、該ポリヌクレオチドは、(a) 配列番号2～668のいずれかの成熟ポリペプチドをコードするヌクレオチドから第一のPCRプライマーおよび第二のPCRプライマーを得て、ここで第一のプライマーおよび第二のプライマーは、PCR条件下で、外側に向かう方式で、核酸合成を開始することが可能であり、そして第一のプライマーはアンチセンス方向に伸長可能であり、そして第二のプライマーはセンス方向に伸長可能である、そして(b) 該ポリヌクレオチドを含有するcDNAライブラリーと、第一のPCRプライマーおよび第二のPCRプライマーを、第一のプライマーおよび第二のプライマーからヌクレオ

チド配列を合成するのに適したPCR条件下で合わせることを含んでなる工程によって同定される。

【0025】

やはり提供するのは、ポリメラーゼ連鎖反応（PCR）を用いて、本発明のポリヌクレオチドを伸長する方法であって、（a）該ポリヌクレオチドから第一のPCRプライマーおよび第二のPCRプライマーを得て、ここで第一のPCRプライマーおよび第二のPCRプライマーは、PCR条件下で、外側に向かう方式で、核酸合成を開始することが可能であり、そして第一のPCRプライマーはアンチセンス方向に伸長可能であり、そして第二のPCRプライマーはセンス方向に伸長可能である、そして（b）cDNAライブラリーに含有される該ポリヌクレオチドと、第一のPCRプライマーおよび第二のPCRプライマーを、第一のPCRプライマーおよび第二のPCRプライマーからヌクレオチド配列を合成するのに適したPCR条件下で合わせて、それによって該ポリヌクレオチドを伸長する工程を含んでなる、前記方法である。

10

【0026】

前述の一般的な説明および続く詳細な説明は、本発明の例示であるが限定的でないことを理解すべきである。

【課題を解決するための手段】

【0027】

発明の詳細な説明

本発明は、A、B、CおよびG群を含む、β溶血性連鎖球菌によって引き起こされる感染を改善し、そして防御する組成物および方法を提供する。β溶血性連鎖球菌に引き起こされる感染の改善および防御に有用なポリヌクレオチドおよびポリペプチドを同定するため、2つの戦略、ゲノミクスアプローチおよびプロテオミクスアプローチを用いて、表面に局在する、化膿性連鎖球菌タンパク質を同定した。

20

【0028】

ゲノミクスアプローチには、表面に局在するタンパク質をコードするであろう遺伝子を同定し、そして性質決定するために設計した、いくつかのアルゴリズムを用いる、化膿性連鎖球菌ゲノムのin silicoでの広範なゲノム解析が含まれた。化膿性連鎖球菌表面に存在するタンパク質を同定するため、プロテオミクスアプローチに取り掛かった。各アプローチの欠陥を克服するため、両アプローチに対する信頼性が重要であった。ゲノムマイニングは、遺伝子の能力を提供するが、実際の表現型発現に関する情報はほとんど提供しない。逆に、プロテオミクス解析は、細胞表面に局在する実際のタンパク質を同定するが、タンパク質発現が制御されている可能性があり、そして細菌細胞を培養する特定の条件が、同定されるタンパク質セットに影響を及ぼしている可能性がある。

30

【0029】

ゲノミクスおよびプロテオミクスアプローチの結果を組み合わせ、そして目的のORFを4つの群の1つに分類した：（i）プロテオミクスによって同定される表面局在タンパク質をコードするORF（表I、奇数番号の配列番号1～147）；（ii）推定上のリポタンパク質をコードするORF（表II、奇数番号の配列番号149～181、669）；（iii）LPXTGモチーフを含有する推定上のポリペプチドをコードするORF（表III、奇数番号の配列番号183～187）；および（iv）他の推定上の表面局在ポリペプチドをコードするORF（表IV、奇数番号の配列番号189～667）。表I～IVに含有されるORFは重複せず、すなわち、多くのORFが別の表に一致する特性を所持するが、表I～IVに列挙されるORFは、各々、1度しか現れない。したがって、例えば表II～IVの1以上にも分類可能であるが、これらの表には含まれない、表Iに列挙されるORF（プロテオミクスによって同定される表面局在タンパク質をコードするORF）がある。

40

【0030】

表I. プロテオミクスによって同定される表面局在タンパク質をコードするオープンリーディングフレーム（ORF）

50

【 0 0 3 1 】

【 表 1 】

配列番号 1 (ORF 66)	配列番号 51 (ORF 1237)	配列番号 101 (ORF 1975)
配列番号 3 (ORF 102)	配列番号 53 (ORF 1238)	配列番号 103 (ORF 2019)
配列番号 5 (ORF 145)	配列番号 55 (ORF 1253)	配列番号 105 (ORF 2064)
配列番号 7 (ORF 232)	配列番号 57 (ORF 1284)	配列番号 107 (ORF 2086)
配列番号 9 (ORF 238)	配列番号 59 (ORF 1316)	配列番号 109 (ORF 2106)
配列番号 11 (ORF 436)	配列番号 61 (ORF 1330)	配列番号 111 (ORF 2116)
配列番号 13 (ORF 516)	配列番号 63 (ORF 1358)	配列番号 113 (ORF 2120)
配列番号 15 (ORF 554)	配列番号 65 (ORF 1487)	配列番号 115 (ORF 2123)
配列番号 17 (ORF 589)	配列番号 67 (ORF 1495)	配列番号 117 (ORF 2202)
配列番号 19 (ORF 661)	配列番号 69 (ORF 1557)	配列番号 119 (ORF 2214)
配列番号 21 (ORF 668)	配列番号 71 (ORF 1638)	配列番号 121 (ORF 2330)
配列番号 23 (ORF 678)	配列番号 73 (ORF 1650)	配列番号 123 (ORF 2354)
配列番号 25 (ORF 704)	配列番号 75 (ORF 1654)	配列番号 125 (ORF 2377)
配列番号 27 (ORF 743)	配列番号 77 (ORF 1659)	配列番号 127 (ORF 2379)
配列番号 29 (ORF 825)	配列番号 79 (ORF 1698)	配列番号 129 (ORF 2387)
配列番号 31 (ORF 850)	配列番号 81 (ORF 1788)	配列番号 131 (ORF 2417)
配列番号 33 (ORF 934)	配列番号 83 (ORF 1794)	配列番号 133 (ORF 2420)
配列番号 35 (ORF 993)	配列番号 85 (ORF 1816)	配列番号 135 (ORF 2422)
配列番号 37 (ORF 1036)	配列番号 87 (ORF 1818)	配列番号 137 (ORF 2450)
配列番号 39 (ORF 1140)	配列番号 89 (ORF 1819)	配列番号 139 (ORF 2459)
配列番号 41 (ORF 1157)	配列番号 91 (ORF 1850)	配列番号 141 (ORF 2477)
配列番号 43 (ORF 1191)	配列番号 93 (ORF 1854)	配列番号 143 (ORF 2586)
配列番号 45 (ORF 1218)	配列番号 95 (ORF 1878)	配列番号 145 (ORF 2593)
配列番号 47 (ORF 1224)	配列番号 97 (ORF 1902)	配列番号 147 (ORF 2601)
配列番号 49 (ORF 1234)	配列番号 99 (ORF 1943)	

10

20

30

表 I I . 推定上のリポタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 3 2 】

【 表 2 】

配列番号 149 (ORF 68)	配列番号 161 (ORF 685)	配列番号 173 (ORF 1789)
配列番号 151 (ORF 309)	配列番号 163 (ORF 729)	配列番号 175 (ORF 1882)
配列番号 153 (ORF 347)	配列番号 165 (ORF 747)	配列番号 177 (ORF 1918)
配列番号 155 (ORF 540)	配列番号 167 (ORF 1202)	配列番号 179 (ORF 1983)
配列番号 157 (ORF 601)	配列番号 169 (ORF 1723)	配列番号 181 (ORF 2452)
配列番号 159 (ORF 664)	配列番号 171 (ORF 1755)	配列番号 669 (ORF 1664)

40

表 I I I . L P X T G モチーフを含有する推定上のポリペプチドをコードするオープンリーディングフレーム (O R F)

【 0 0 3 3 】

50

【表 3】

配列番号 183 (ORF 433) 配列番号 185 (ORF 967) 配列番号 187 (ORF 2497)

表 I V. 他 の 推 定 上 の 表 面 局 在 ポ リ ペ プ チ ド を コ ー ド す る オ ー プ ン リ ー デ ィ ン グ フ レ ー ム (O R F)

【 0 0 3 4 】

【表 4 - 1】

配列番号 189 (ORF 4)	配列番号 349 (ORF 741)	配列番号 509 (ORF 1682)
配列番号 191 (ORF 5)	配列番号 351 (ORF 754)	配列番号 511 (ORF 1683)
配列番号 193 (ORF 11)	配列番号 353 (ORF 774)	配列番号 513 (ORF 1720)
配列番号 195 (ORF 17)	配列番号 355 (ORF 783)	配列番号 515 (ORF 1725)
配列番号 197 (ORF 18)	配列番号 357 (ORF 788)	配列番号 517 (ORF 1726)
配列番号 199 (ORF 20)	配列番号 359 (ORF 805)	配列番号 519 (ORF 1732)
配列番号 201 (ORF 25)	配列番号 361 (ORF 814)	配列番号 521 (ORF 1736)
配列番号 203 (ORF 49)	配列番号 363 (ORF 818)	配列番号 523 (ORF 1771)
配列番号 205 (ORF 64)	配列番号 365 (ORF 844)	配列番号 525 (ORF 1772)
配列番号 207 (ORF 65)	配列番号 367 (ORF 848)	配列番号 527 (ORF 1775)
配列番号 209 (ORF 67)	配列番号 369 (ORF 858)	配列番号 529 (ORF 1776)
配列番号 211 (ORF 69)	配列番号 371 (ORF 859)	配列番号 531 (ORF 1777)
配列番号 213 (ORF 72)	配列番号 373 (ORF 860)	配列番号 533 (ORF 1783)
配列番号 215 (ORF 73)	配列番号 375 (ORF 871)	配列番号 535 (ORF 1785)
配列番号 217 (ORF 75)	配列番号 377 (ORF 877)	配列番号 537 (ORF 1786)
配列番号 219 (ORF98)	配列番号 379 (ORF 896)	配列番号 539 (ORF 1814)
配列番号 221 (ORF 99)	配列番号 381 (ORF 908)	配列番号 541 (ORF 1820)
配列番号 223 (ORF 130)	配列番号 383 (ORF 909)	配列番号 543 (ORF 1828)
配列番号 225 (ORF 133)	配列番号 385 (ORF 910)	配列番号 545 (ORF 1833)
配列番号 227 (ORF 141)	配列番号 387 (ORF 920)	配列番号 547 (ORF 1834)
配列番号 229 (ORF 151)	配列番号 389 (ORF 921)	配列番号 549 (ORF 1839)
配列番号 231 (ORF 165)	配列番号 391 (ORF 926)	配列番号 551 (ORF 1873)
配列番号 233 (ORF 172)	配列番号 393 (ORF 928)	配列番号 553 (ORF 1875)
配列番号 235 (ORF 184)	配列番号 395 (ORF 929)	配列番号 555 (ORF 1876)
配列番号 237 (ORF 189)	配列番号 397 (ORF 933)	配列番号 557 (ORF 1888)
配列番号 239 (ORF 199)	配列番号 399 (ORF 952)	配列番号 559 (ORF 1909)
配列番号 241 (ORF 209)	配列番号 401 (ORF 961)	配列番号 561 (ORF 1917)
配列番号 243 (ORF 218)	配列番号 403 (ORF 975)	配列番号 563 (ORF 1931)
配列番号 245 (ORF 220)	配列番号 405 (ORF 983)	配列番号 565 (ORF 1970)
配列番号 247 (ORF 223)	配列番号 407 (ORF 991)	配列番号 567 (ORF 1972)
配列番号 249 (ORF 227)	配列番号 409 (ORF 1015)	配列番号 569 (ORF 1979)
配列番号 251 (ORF 241)	配列番号 411 (ORF 1018)	配列番号 571 (ORF 1987)

10

20

30

40

【表 4 - 2】

配列番号 253 (ORF 252)	配列番号 413 (ORF 1020)	配列番号 573 (ORF 1993)
配列番号 255 (ORF 264)	配列番号 415 (ORF 1021)	配列番号 575 (ORF 2013)
配列番号 257 (ORF 265)	配列番号 417 (ORF 1026)	配列番号 577 (ORF 2014)
配列番号 259 (ORF 291)	配列番号 419 (ORF 1058)	配列番号 579 (ORF 2015)
配列番号 261 (ORF 292)	配列番号 421 (ORF 1110)	配列番号 581 (ORF 2020)
配列番号 263 (ORF 306)	配列番号 423 (ORF 1132)	配列番号 583 (ORF 2023)
配列番号 265 (ORF 307)	配列番号 425 (ORF 1152)	配列番号 585 (ORF 2046)
配列番号 267 (ORF 313)	配列番号 427 (ORF 1156)	配列番号 587 (ORF 2048)
配列番号 269 (ORF 350)	配列番号 429 (ORF 1188)	配列番号 589 (ORF 2050)
配列番号 271 (ORF 352)	配列番号 431 (ORF 1200)	配列番号 591 (ORF 2069)
配列番号 273 (ORF 353)	配列番号 433 (ORF 1203)	配列番号 593 (ORF 2070)
配列番号 275 (ORF 368)	配列番号 435 (ORF 1205)	配列番号 595 (ORF 2091)
配列番号 277 (ORF 401)	配列番号 437 (ORF 1210)	配列番号 597 (ORF 2148)
配列番号 279 (ORF 405)	配列番号 439 (ORF 1216)	配列番号 599 (ORF 2170)
配列番号 281 (ORF 421)	配列番号 441 (ORF 1228)	配列番号 601 (ORF 2201)
配列番号 283 (ORF 491)	配列番号 443 (ORF 1231)	配列番号 603 (ORF 2222)
配列番号 285 (ORF 510)	配列番号 445 (ORF 1265)	配列番号 605 (ORF 2231)
配列番号 287 (ORF 511)	配列番号 447 (ORF 1267)	配列番号 607 (ORF 2236)
配列番号 289 (ORF 519)	配列番号 449 (ORF 1269)	配列番号 609 (ORF 2240)
配列番号 291 (ORF 523)	配列番号 451 (ORF 1272)	配列番号 611 (ORF 2245)
配列番号 293 (ORF 535)	配列番号 453 (ORF 1275)	配列番号 613 (ORF 2247)
配列番号 295 (ORF 551)	配列番号 455 (ORF 1292)	配列番号 615 (ORF 2250)
配列番号 297 (ORF 567)	配列番号 457 (ORF 1300)	配列番号 617 (ORF 2258)
配列番号 299 (ORF 570)	配列番号 459 (ORF 1310)	配列番号 619 (ORF 2266)
配列番号 301 (ORF 594)	配列番号 461 (ORF 1311)	配列番号 621 (ORF 2273)
配列番号 303 (ORF 597)	配列番号 463 (ORF 1318)	配列番号 623 (ORF 2289)
配列番号 305 (ORF 602)	配列番号 465 (ORF 1321)	配列番号 625 (ORF 2291)
配列番号 307 (ORF 613)	配列番号 467 (ORF 1362)	配列番号 627 (ORF 2300)
配列番号 309 (ORF 627)	配列番号 469 (ORF 1395)	配列番号 629 (ORF 2319)
配列番号 311 (ORF 639)	配列番号 471 (ORF 1497)	配列番号 631 (ORF 2342)
配列番号 313 (ORF 644)	配列番号 473 (ORF 1500)	配列番号 633 (ORF 2391)
配列番号 315 (ORF 650)	配列番号 475 (ORF 1512)	配列番号 635 (ORF 2398)
配列番号 317 (ORF 653)	配列番号 477 (ORF 1513)	配列番号 637 (ORF 2399)
配列番号 319 (ORF 665)	配列番号 479 (ORF 1525)	配列番号 639 (ORF 2411)
配列番号 321 (ORF 670)	配列番号 481 (ORF 1527)	配列番号 641 (ORF 2414)
配列番号 323 (ORF 671)	配列番号 483 (ORF 1548)	配列番号 643 (ORF 2428)

10

20

30

40

【表 4 - 3】

配列番号 325 (ORF 672)	配列番号 485 (ORF 1573)	配列番号 645 (ORF 2429)
配列番号 327 (ORF 674)	配列番号 487 (ORF 1585)	配列番号 647 (ORF 2437)
配列番号 329 (ORF 676)	配列番号 489 (ORF 1586)	配列番号 649 (ORF 2457)
配列番号 331 (ORF 688)	配列番号 491 (ORF 1593)	配列番号 651 (ORF 2458)
配列番号 333 (ORF 699)	配列番号 493 (ORF 1608)	配列番号 653 (ORF 2473)
配列番号 335 (ORF 702)	配列番号 495 (ORF 1661)	配列番号 655 (ORF 2482)
配列番号 337 (ORF 705)	配列番号 497 (ORF 1667)	配列番号 657 (ORF 2488)
配列番号 339 (ORF 706)	配列番号 499 (ORF 1671)	配列番号 659 (ORF 2508)
配列番号 341 (ORF 721)	配列番号 501 (ORF 1672)	配列番号 661 (ORF 2521)
配列番号 343 (ORF 731)	配列番号 503 (ORF 1678)	配列番号 663 (ORF 2534)
配列番号 345 (ORF 733)	配列番号 505 (ORF 1680)	配列番号 665 (ORF 2562)
配列番号 347 (ORF 737)	配列番号 507 (ORF 1681)	配列番号 667 (ORF 2583)

10

ゲノミクスアプローチ

現在、バイオインフォマティクスの情報プロセッシング能とカップリングした、ゲノミクス、転写プロファイリングおよびプロテオミクスを通じて、免疫原組成物候補を同定する際には、完全細菌ゲノム配列の入手可能性が、重要な役割を演じている（39～41、53、60、65）。

20

【0035】

ゲノミクスアプローチは、オクラホマ大学のウェブサイトからダウンロードした化膿性連鎖球菌の未注釈（unannotated）配列においてオープンリーディングフレーム（ORF）を同定することによって始まった。このゲノム配列は、GenBankに提出され、そして寄託番号AE004092を割り当てられたと報告された。株M1 GASは、ATCCに提出され、そして寄託番号ATCC 700294を与えられたと報告された。

【0036】

ORFは、3つのありうる開始部位コドン、ATG、GTG、またはTTGの1つ、および3つのありうる停止コドン、TAA、TAG、またはTGAの1つを有すると本明細書に定義される。ORFのこの定義を用いて、ORFを同定するため、3つのORFファインダーアルゴリズム、すなわちGLIMMER（59）、GeneMark（34）、および本発明者の譲受人に開発されたアルゴリズムを用いて、化膿性連鎖球菌ゲノムを解析した。3つのアルゴリズムすべてに共通に同定される736のORFがあった。異なるORFファインダー間の結果の相違は、主に、各プログラムが用いる特定の開始コドンによるものであるが、Glimmerはまた、シャイン・ダルガー・ボックスに関するある程度の評価も取り込む。共通の停止コドンを持つORFにはすべて、同一のORF名称を与え、そして同一のORFであるかのように扱った。

30

【0037】

決定したORFの正確さを評価するため、離散コサイン変換（Disction）として当該技術分野に知られる、離散数学的コサイン関数を使用して、各ORFにスコアを割り当てた。Disctionスコア>1.5のORFは、タンパク質産物をコードする高い確率を有するとみなした。3つのORF発見アルゴリズムに予測されるORFの最低の長さを、74アミノ酸のタンパク質をコードするであろう225ヌクレオチド（停止コドンを含む）に設定した。

40

【0038】

ORFの残りの最終検索として、フレームシフトを含有する遺伝子領域（42）または抗原変異を引き起こすのに役割を有する可能性がある遺伝子断片（21）を同定するため、tBLASTnを用いて、公的タンパク質データベースに対して、>75ヌクレオチドの

50

非コード領域すべてを検索した。これらの残りのORFをORFヒットに加えた。

【0039】

本発明者の譲受人に開発された図式的解析プログラムを用いて、ゲノム配列に比較して、予測されるORFの6つの読み枠および位置をすべて示した。これは、他のORFと長い重複を有するORFを除去するのを補助したが、他のORFに完全に埋まるORFの既知の例がある(25、33)。

【0040】

BLAST v2.0ギャップ化検索アルゴリズム、BLASTpを用いて、これらの化膿性連鎖球菌ORFの最初の注釈付けを行って、相同配列を同定した。 $< e^{-10}$ のいずれかのカットオフ「e」値を有意とみなした。FASTAおよびPSI-BLASTを含む他の検索アルゴリズムもまた用いた。相同検索に用いた非重複タンパク質配列データベースには、毎日アップデートされる、GenBank、SWISS-PROT、PIR、およびTREMBLデータベース配列が含まれた。 $> e^{-10}$ のBLASTp結果を持つORFを、化膿性連鎖球菌に特有であるとみなした。

【0041】

現在、細菌ゲノム内のすべてのORFの約60%が、その機能が決定されているタンパク質とある程度の一致を有する。これによって、約40%のゲノムORFがなお性質決定されずに残っている。既知のまたは推測される候補標的遺伝子と共に、タンパク質の位置または機能を同定するワードを用いて、全BLAST結果のキーワード検索を行った。さらに、最初のBLAST結果と関連するMEDLINE参考文献すべてのキーワード検索を行って、ORFに関するさらなる情報を探した。キーワード検索には、例えば、以下の検索用語が含まれた：アドヘシン（接着）；フィブロネクチン；フィブリノーゲン；コラーゲン；輸送体；搬出体（exporter）；細胞外；トランスフェラーゼ；表面；および結合。ORFのBLAST解析によって、1005のORFが未分類と指摘され、284のORFは、化膿性連鎖球菌由来のタンパク質とのみBLAST類似性を生じたため、この生物に特異的であるようであり、そして676のORFは、Medline参考文献と関連した。

【0042】

DNA解析のため、各遺伝子内の%G+C含量を同定した。ORFの%G+C含量は、ORF内のコドンすべての第三のヌクレオチド位の(G+C)含量として計算した。報告される値は、該生物で見られるすべてのORFに関して得られるこうした値の算術平均からのこの値の相違であった。これらのORFは、多くの他の病原性アイランドと調和するパターン(22)の、ピロリ菌(H. pylori)由来のcag病原性アイランドの場合に見られているような(2)、水平移行から生じる可能性があるため、さらなる解析には8以上の絶対値が重要とみなされた。G+C含量が有意に異なるORFは、総計289であった。水平移行によって別の生物から獲得した病原性因子に対する類似性に関して、これらのORFをさらに調べた。

【0043】

いくつかのパラメーターを用いて、予測されるタンパク質の分配(partitioning)を決定した。細胞膜を渡る転位置を運命付けるタンパク質は、N末端で陽性荷電残基が隣接する中央疎水性領域で構成されるリーダーシグナル(シグナル配列としても知られる)をコードする(56)。プログラムSignalPを用いて、シグナルペプチドおよびその切断部位を同定した(46)。発現中、シグナルペプチドが切断されて、成熟ペプチドを生じる。さらに、細菌中のタンパク質局在を予測するため、ソフトウェアPSORTを用いた(44)。PSORTは神経網アルゴリズムを用いて、グラム陽性細菌では、細胞質、周辺質、および／または細胞膜、それと共にグラム陰性細菌では外膜へのタンパク質の局在を予測する。PSORTは予測される40のORFが表面曝露されていると同定した(表V)。

【0044】

表V. 推定上の細胞外タンパク質をコードするオープンリーディングフレーム(ORF)

)

【 0 0 4 5 】

【表 5】

68	705	1202	1664	1723	2020	2385
165	729	1310	1667	1777	2046	2414
252	788	1358	1678	1909	2170	2437
510	1058	1362	1680	1972	2236	2601
601	1132	1573	1681	1975	2250	
668	1200	1638	1683	2014	2300	

さらに、ソフトウェアプログラム T o p P r e d 2 (1 0) を用いて、タンパク質の膜貫通 (T M) ドメインを解析した。このプログラムは、潜在的に膜の脂質二重層に渡る可能性がある、疎水性のタンパク質領域を予測する。T o p P r e d 2 による、潜在的に膜の脂質二重層に渡る可能性がある、タンパク質疎水性領域に関する解析によって、3 以上の膜貫通ドメインを持ち、そしてしたがって膜結合であるとみなされる、4 8 の O R F (表 V I) が同定された。

【 0 0 4 6 】

表 V I . 3 以上の膜貫通領域を持つ、推定上のタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 4 7 】

【表 6】

8	307	594	752	1222	1598	2069
73	312	613	844	1266	1657	2091
80	395	650	925	1317	1708	2227
95	508	672	975	1488	1726	2283
141	551	706	1018	1496	1779	2424
265	567	708	1152	1513	1999	2562
306	593	731	1156	1596	2002	

タンパク質ドメインまたは保存されるタンパク質領域の多数並列の隠れマルコフモデル (H M M) P f a m データベース (6 1) を用いて、現存するタンパク質ファミリーに属する可能性がある化膿性連鎖球菌タンパク質を同定した。このアウトプットのキーワード検索を用いて、B l a s t 検索規準では見逃した可能性があるタンパク質を同定した。本発明者の譲受人もまた、H M M モデルを開発した。コンピュータアルゴリズム、H M M L i p o を開発して、1 3 2 の生物学的に性質決定された、3 0 を超える生物由来の非化膿性連鎖球菌細菌リポタンパク質を用いて、リポタンパク質を予測した。このトレーニングセットは、実験的に立証された原核生物リポタンパク質から生成された。H M M L i p o は、推定上のリポタンパク質である、3 0 の O R F を同定した (表 V I I) 。

【 0 0 4 8 】

表 V I I . 推定上のリポタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 4 9 】

【表 7】

68	601	747	1659	1789	1983
309	678	1157	1664	1818	2417
347	685	1202	1723	1878	2452
540	704	1284	1755	1882	2459
554	729	1495	1788	1918	2601

さらに、1 5 の O R F が L P X T G モチーフを有すると予測され、そしてソルターゼに標的とされる可能性があるタンパク質として分類された (表 V I I I) 。

【 0 0 5 0 】

10

20

30

40

50

表 V I I I . L P X T G モチーフを含有する推定上のタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 5 1 】

【 表 8 】

433	1218	1854	2450
608	1316	2019	2477
967	1330	2434	2497
1191	1698	2446	

配列番号 6 6 9 ~ 6 7 4 は、それぞれ、タンパク質 G r a b (O R F 6 0 8)、M タンパク質 (O R F 2 4 3 4)、および S c p A (O R F 2 4 4 6) のヌクレオチドおよびアミノ酸配列を含有する。 10

【 0 0 5 2 】

さらに、L P X T G 細胞壁局所化シグナルを含有する、約 7 0 の既知の原核生物タンパク質を用いて、HMM (1 5) を開発して、ペプチドグリカン層に係留される細胞壁タンパク質を予測した (3 8、4 5)。このモデルは L P X T G 配列を用いるだけでなく、下流配列の 2 つの特徴、疎水性膜貫通ドメインおよび陽性荷電カルボキシ末端も含んだ。ソルターゼと独立した、非共有様式で、ペプチドグリカン層に潜在的に結合すると同定された 5 つのタンパク質があった (表 I X)。

【 0 0 5 3 】

表 I X . 推定上のペプチドグリカン結合タンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 5 4 】

【 表 9 】

898	1569	1675	2266	2311
-----	------	------	------	------

また、同定された O R F にコードされるタンパク質を他の特性に関しても評価した。タンデム反復ファインダー (5) は、M S C R A M M (2 0) および髄膜炎菌 (N e i s s e r i a m e n i n g i t i d i s) の相可変 (p h a s e v a r i a b l e) 表面タンパク質 (5 1) に見出されるものなどの反復 D N A 配列を含有する O R F を同定した。 30
こうした反復領域を含有するタンパク質をコードすることが見出された 2 3 の O R F があった (表 X)。

【 0 0 5 5 】

表 X . 反復領域を含有する推定上のタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 5 6 】

【 表 1 0 】

218	433	1149	1783	2422	2513
265	555	1562	1972	2434	2590
336	699	1583	2137	2437	2618
431	783	1683	2231	2477	

さらに、A r g - G l y - A s p (R G D) 接着モチーフを含有するタンパク質は、その受容体として作用するインテグリンと共に、細胞接着の主要認識系を構成する。R G D 認識は、微生物が真核生物組織への進入を獲得するのに用いる 1 つの機構である (2 9、6 3)。R G D 含有タンパク質をコードすると同定される、6 5 の O R F があった (表 X I)。

【 0 0 5 7 】

表 X I . R G D モチーフを含有する推定上のタンパク質をコードするオープンリーディングフレーム (O R F)

【0058】

【表11】

18	544	885	1149	1504	1957	2379
201	626	889	1161	1626	2042	2414
209	641	899	1200	1643	2054	2446
302	654	967	1274	1657	2082	2558
344	667	968	1313	1675	2148	2570
350	668	1010	1316	1773	2205	
396	695	1027	1373	1779	2247	
397	726	1074	1401	1885	2253	
413	787	1108	1416	1891	2287	
526	829	1110	1431	1901	2335	

10

ゲノム解析およびORF同定の結果の図式的表示を図1に示す。

【0059】

プロテオミクスアプローチ

上述のように、化膿性連鎖球菌表面局在タンパク質を同定するため、プロテオミクスアプローチも取った。

【0060】

細胞表面に局在するタンパク質のみを同定するため、トリプシンを用いて化膿性連鎖球菌細胞を調製し、そして消化する間、注意を払った。細胞試料は、トリプシン添加直前および消化完了時に採取し、そして生存数およびLV-SEMによって細胞完全性に関して試験した。消化後、未処理細胞は明らかに凝集し、そして試験管の側面に付着したが、処理細胞は均一な細胞懸濁物を形成した。生存数は試料間で有意な相違を示さず、そして実際は、未処理試料が凝集するため、処理細胞において、わずかにより高かった。LV-SEMによってこれらの結果を確認した(図2)。消化細胞は、均一に、そして個々に、カバーリップ上に分布し、一方、未処理試料は、より大きい細菌塊を示した。未処理細菌細胞の高倍率でのトポグラフィック検査は、化膿性連鎖球菌に特徴的な表面物質が多量にあることを示した。しかし、トリプシン消化試料中の個々の細胞は、細胞が擦り減って(bald)見え、そしていかなる表面物質も欠けているため、すべての観察可能な表面タンパク質の減少を示した。図3は、トリプシン消化前(左のパネル、パネルA)および消化後(右のパネル、パネルB)の化膿性連鎖球菌のLV-SEMを示す。トリプシン消化前の細胞(パネルA)はより大きく、そして表面物質を示す。消化後の細胞のLV-SEM(パネルB)はより小さく、そしていかなる表面タンパク質も欠けて見える。

20

30

【0061】

複雑な表面消化混合物のペプチド構成要素を同定するため、解析技術を用いて、広い濃度範囲に渡って高感度で多数のペプチドを分離し、そして配列決定した。タンデム質量分析(MS/MS)は、ゲルおよび溶液両方から、タンパク質を解析する強力なアプローチであることが示されてきた(17)。MS/MSは、まず、質量分析装置を用いて、イオン混合物からペプチドイオンを分離し、その後、第二の工程または質量分析装置を用いて、目的のイオンを活性化し、そして解離する。このプロセスは、衝突誘起解離(collision induced dissociation)(CID)として知られ、アミノ酸間のペプチド結合でのペプチドの断片化を引き起こし、そしてしたがって、ペプチドの断片化パターンを用いて、アミノ酸配列を決定する。

40

【0062】

さらに、SEQUESTコンピュータアルゴリズムを用いて、タンパク質に対して、または翻訳されたヌクレオチド配列データベースに対して、実験的断片化スペクトルを直接検索した。大きさおよそ800~900Daを超えるペプチドに関して、単一のスペクトルが、特有にタンパク質を同定可能である。

【0063】

複雑な混合物から多数のペプチドを配列決定するため、逆相クロマトグラフィー系をエレ

50

クトロスプレーイオントラップ質量分析計とカップリングした。この系では、流速およびカラム直径両方を最小化して、溶出体積を濃縮し、そして質量分析計検出装置の開口部に可能な限り多くのカラム流出物を導くことによって、高感度（サブフェムトモルレベル未満まで）に到達可能であることが知られる。最初の実験は、1%アセトニトリル／分の逆相勾配を用いて、ペプチドを分離した。その後、クロマトグラフィー分離を増加させるため、0.28%アセトニトリル／分までのより長い勾配およびより遅い流速（50 n l / 分）を使用した。試料に存在するタンパク質の範囲を最大にするため、イオントラップのデータ依存獲得特徴を使用した。

【0064】

動的排除（dynamic exclusion）を用いて、特定の m/z 値に関してひとたびスペクトルを得たら、イオンのタンデムマススペクトルの再獲得を防いだ。同位体排除機能は、MS/MSの候補としたイオンのリストから、ペプチドの ^{13}C 同位体と関連するイオンを排除した。この目的のため、3-uマス幅ウィンドウを選択した。これらのデータに依存する特徴を用いると、CID解析に選択されるペプチドイオンの数が劇的に増加した。

【0065】

上述のLC-MS/MSデータ獲得条件は、典型的には、各実行あたり、2000を超えるペプチドイオンに関する断片化データを生じた。SEQUESTアルゴリズムを用いて、非重複タンパク質配列データベースOWLと組み合わせて、化膿性連鎖球菌由来の翻訳されたORFを含有する合成タンパク質配列データベースに対して、このデータを検索した。SEQUEST検索条件は修飾トリプシン選択性を用い、そしてメチオニンに対する+16Daの示差検索を可能にして、メチオニン酸化を明らかにした。以下のマニュアル法を用いて、SEQUESTに同定された候補マッチを確認した。2.5より大きいXcorr値（実験ms/msデータと配列データベースから生成されるものの類似性の測定値）および0.1より大きいdelCn値（delCnは第一のマッチおよび第二のマッチのXcorr値間の規準化した相違を測定する）を持つマッチをさらなる選択のために選択した。優れたマッチ由来の断片化スペクトルを妥当なシグナル／ノイズに関してチェックし、そしてマッチしたイオンのリストを妥当な連続性に関して調べた。他の確認ms/msデータが同一試料によって生成されたならば、単独では許容しえないマッチのいくつかを含んだ。このプロテオミクスアプローチで得たORFを表XIIに示す。

【0066】

表XII. トリプシン消化によって同定されたオープンリーディングフレーム（ORF）

【0067】

【表12】

66	678	1224	1638	1878	2214	2459
102	704	1234	1650	1902	2330	2477
145	743	1237	1654	1943	2354	2586
232	825	1238	1659	1975	2377	2593
238	850	1253	1698	2019	2379	2601
436	934	1284	1788	2064	2387	
516	993	1316	1794	2086	2417	
554	1036	1330	1816	2106	2420	
589	1140	1358	1818	2116	2422	
608	1157	1487	1819	2120	2434	
661	1191	1495	1850	2123	2446	
668	1218	1557	1854	2202	2450	

同定されたORFのいくつかをクローニングし、そして発現した。精製タンパク質に対して生成したマウス抗血清をまず、コーティング抗原としてマウス免疫に用いるのと同じの調製を用いたELISAによって、反応性に関して解析した。化膿性連鎖球菌の表面上の

タンパク質発現を定量化するため、その後、これらの血清を全細胞 E L I S A で用いた。特定のタンパク質のタンパク質発現を認定するため、全化膿性連鎖球菌細胞をイムノゴールドによって標識し、そして L V - S E M によって視覚化した。

【0068】

同定した O R F のいくつかに関して、コードされるタンパク質は、増殖期（対数中期対静止期）に応じた方式で発現されることが観察された。この種類の例は、O R F 2 1 8（図4）、O R F 5 5 4（図5）、および O R F 1 1 9 1（図6）である。いくつかの場合、発現レベルは、対数増殖中期でより高く、一方、他の場合、静止細胞でより高かった。他の O R F にコードされるタンパク質は、増殖期にかかわらず、低レベルで発現され（O R F 2 0 6 4、2 6 0 1、および 1 3 1 6）（それぞれ図7～9に示す）、一方、他のものは、増殖期と独立して高レベルで発現された（O R F 1 2 2 4）（図10）。化膿性連鎖球菌の細胞壁で発現され、そしてそこに局在することが知られるため、陽性対照として、抗 C 5 a ペプチダーゼ血清を用いた。すべての抗血清は、それぞれの免疫前対照血清よりも、反応性の増加を示した。

【0069】

ゲノミクスアプローチおよびプロテオミクスアプローチの組み合わせ

表 V ～ X I I で同定された O R F を、その後、4つの群の1つに分類した：プロテオミクスによって同定される表面局在タンパク質をコードする O R F（表 I）；推定上のリポタンパク質をコードする O R F（表 I I）；L P X T G モチーフを含有する推定上のポリペプチドをコードする O R F（表 I I I）；および他の推定上の表面局在ポリペプチドをコードする O R F（表 I V）。表 I ～ I V を上記に提供する。表 I ～ I V に含有される O R F は重複せず、すなわち表 I ～ I V に列挙される O R F は、多くが別の表に一致する特性を所持するが、各々、1度しか現れないことが明らかなはずである。

【0070】

表 I のヌクレオチド配列は、プロテオミクスアプローチによって、表面に局在する化膿性連鎖球菌タンパク質であると同定されたポリペプチドをコードする。表 I I ～ I V のヌクレオチド配列は、記載されるゲノミクスアプローチによって、表面に局在する化膿性連鎖球菌タンパク質であると同定された推定上のポリペプチドをコードする。具体的には、表 I I のヌクレオチド配列は推定上のリポタンパク質をコードし、表 I I I のヌクレオチド配列は、L P X T G 細胞壁局所化シグナルを有する推定上のタンパク質をコードし、そして表 I V のヌクレオチド配列は、機能および細胞内の位置が先に同定されている他のタンパク質に対する類似性、タンパク質ファミリー（例えば P f a m）とのマッチ、並びに膜貫通ドメイン、P s o r t および s i g P 値、およびタンパク質の予測される分子量の組み合わせ解析を含む、本明細書に記載するような、いくつかの規準の少なくとも1つを含む、推定上の表面局在タンパク質をコードする。

【0071】

奇数番号の配列番号 1 ～ 6 6 7 は各々、ヌクレオチド配列の後に続いて番号付けされるアミノ酸配列をコードする。したがって、例えば、配列番号 1 のヌクレオチド配列は配列番号 2 のアミノ酸配列をコードし、そして配列番号 3 のヌクレオチド配列は配列番号 4 のアミノ酸配列をコードするなどである。

【0072】

ポリペプチド

本発明は、表面に局在する化膿性連鎖球菌ポリペプチドを提供する。具体的には、本発明のポリペプチドには、偶数番号の配列番号 2 ～ 6 6 8、すなわち、配列番号 2、4、6、8、10、12、14、16、18、20、22、24、26、26、30、32、34、36、38、40、42、44、46、48、50、52、54、56、58、60、62、64、66、68、70、72；74、76、78、80、82、84、86、88、90、92、94、96、98、100、102、104、106、108、110、112、114、116、118、120；122、124、126、128、130、132、134、136；138、140、142、144、146、148、150

、 1 5 2、 1 5 4、 1 5 6、 1 5 8、 1 6 0、 1 6 2、 1 6 4、 1 6 6、 1 6 8、 1 7 0
、 1 7 2、 1 7 4、 1 7 6、 1 7 8、 1 8 0、 1 8 2、 1 8 4、 1 8 6、 1 8 8、 1 9 0
、 1 9 2、 1 9 4、 1 9 6、 1 9 8、 2 0 0、 2 0 2、 2 0 4、 2 0 6、 2 0 8、 2 1 0
、 2 1 2、 2 1 4、 2 1 6、 2 1 8、 2 2 0、 2 2 2、 2 2 4、 2 2 6、 2 2 8、 2 3 0
、 2 3 2、 2 3 4、 2 3 6、 2 3 8、 2 4 0、 2 4 2、 2 4 4、 2 4 6、 2 4 8、 2 5 0
、 2 5 2、 2 5 4、 2 5 6、 2 5 8、 2 6 0、 2 6 2、 2 6 4、 2 6 6、 2 6 8、 2 7 0
、 2 7 2、 2 7 4、 2 7 6、 2 7 8、 2 8 0、 2 8 2、 2 8 4、 2 8 6、 2 8 8、 2 9 0
、 2 9 2、 2 9 4、 2 9 6、 2 9 8、 3 0 0、 3 0 2、 3 0 4、 3 0 6、 3 0 8、 3 1 0
、 3 1 2、 3 1 4、 3 1 6、 3 1 8、 3 2 0、 3 2 2、 3 2 4、 3 2 6、 3 2 8、 3 3 0
、 3 3 2、 3 3 4、 3 3 6、 3 3 8、 3 4 0、 3 4 2、 3 4 4、 3 4 6、 3 4 8、 3 5 0 10
、 3 5 2、 3 5 4、 3 5 6、 3 5 8、 3 6 0、 3 6 2、 3 6 4、 3 6 6、 3 6 8、 3 7 0
、 3 7 2、 3 7 4、 3 7 6、 3 7 8、 3 8 0、 3 8 2、 3 8 4、 3 8 6、 3 8 8、 3 9 0
、 3 9 2、 3 9 4、 3 9 6、 3 9 8、 4 0 0、 4 0 2、 4 0 4、 4 0 6、 4 0 8、 4 1 0
、 4 1 2、 4 1 4、 4 1 6、 4 1 8、 4 2 0、 4 2 2、 4 2 4、 4 2 6、 4 2 8、 4 3 0
、 4 3 2、 4 3 4、 4 3 6、 4 3 8、 4 4 0、 4 4 2、 4 4 4、 4 4 6、 4 4 8、 4 5 0
、 4 5 2、 4 5 4、 4 5 6、 4 5 8、 4 6 0、 4 6 2、 4 6 4、 4 6 6、 4 6 8、 4 7 0
、 4 7 2、 4 7 4、 4 7 6、 4 7 8、 4 8 0、 4 8 2、 4 8 4、 4 8 6、 4 8 8、 4 9 0
、 4 9 2、 4 9 4、 4 9 6、 4 9 8、 5 0 0、 5 0 2、 5 0 4、 5 0 6、 5 0 8、 5 1 0
、 5 1 2、 5 1 4、 5 1 6、 5 1 8、 5 2 0、 5 2 2、 5 2 4、 5 2 6、 5 2 8、 5 3 0
、 5 3 2、 5 3 4、 5 3 6、 5 3 8、 5 4 0、 5 4 2、 5 4 4、 5 4 6、 5 4 8、 5 5 0 20
、 5 5 2、 5 5 4、 5 5 6、 5 5 8、 5 6 0、 5 6 2、 5 6 4、 5 6 6、 5 6 8、 5 7 0
、 5 7 2、 5 7 4、 5 7 6、 5 7 8、 5 8 0、 5 8 2、 5 8 4、 5 8 6、 5 8 8、 5 9 0
、 5 9 2、 5 9 4、 5 9 6、 5 9 8、 6 0 0、 6 0 2、 6 0 4、 6 0 6、 6 0 8、 6 1 0
、 6 1 2、 6 1 4、 6 1 6、 6 1 8、 6 2 0、 6 2 2、 6 2 4、 6 2 6、 6 2 8、 6 3 0
、 6 3 2、 6 3 4、 6 3 6、 6 3 8、 6 4 0、 6 4 2、 6 4 4、 6 4 6、 6 4 8、 6 5 0
、 6 5 2、 6 5 4、 6 5 6、 6 5 8、 6 6 0、 6 6 2、 6 6 4、 6 6 6 または 6 6 8 のい
ずれかのアミノ酸配列を含んでなる単離ポリペプチドが含まれる。

【0073】

本発明のポリペプチドにはまた、本質的に前述のアミノ酸配列からなる単離ポリペプチド、および前述のアミノ酸配列からなる単離ポリペプチドも含まれる。用語「単離」は、天然の状態から人の手によって改変されることを意味する。「単離」組成物または物質が天然に存在する場合、元来の環境から変化したかまたは除去されているか、あるいは両方である。例えば、生存動物に天然に存在するポリペプチドまたはポリヌクレオチドは「単離」されていないが、その天然状態で共存する物質から分離されている、同じポリペプチドまたはポリヌクレオチドは、該用語が本明細書で使用される場合、「単離」されている。本明細書において、用語「単離」は、天然供給源から単離されており、そして／または組換え技術を用いて調製されているポリペプチド（または他の構成要素）を意図する。 30

【0074】

本発明のポリペプチド配列は、偶数番号の配列番号 2～668 の参照配列に同一、すなわち 100% 同一であることも可能であるし、または同一性%が 100% 未満であるように、参照配列に比較した際、特定の整数までのアミノ酸改変を含むことも可能である。こうした改変には、少なくとも 1 つのアミノ酸欠失、保存的および非保存的置換を含む置換、または挿入が含まれる。改変は、参照ポリペプチド配列のアミノ末端もしくはカルボキシ末端位で、またはこれらの末端位の間のどこかで、参照アミノ酸配列中のアミノ酸の間に個々に点在するか、または参照アミノ酸配列内の 1 以上の隣接基に生じることが可能である。 40

【0075】

したがって、本発明はまた、配列表に含有されるアミノ酸配列（すなわち偶数番号の配列番号 2～668）に配列同一性を有する単離ポリペプチドも提供する。特定の配列に応じて、配列同一性の度合いは、好ましくは 50% より高い（例えば 60%、70%、80% 50

、90%、95%、97%、99%以上)。これらの相同タンパク質には、突然変異体および対立遺伝子変異体が含まれる。

【0076】

「同一性」は、当該技術分野に知られるように、配列を比較することによって決定されるような、2以上のポリペプチド配列または2以上のポリヌクレオチド配列間の関係である。当該技術分野において、「同一性」はまた、場合に応じて、こうした配列のストリング間のマッチによって決定されるような、ポリペプチドまたはポリヌクレオチド配列間の配列関連性の度合いも意味する。「同一性」および「類似性」は、限定されるわけではないが、Computational Molecular Biology, Lesk,

A. M. 監修, Oxford University Press, ニューヨーク 10
 , 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W. 監修, Academic Press, ニューヨーク, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M. および Griffin, H. G. 監修, Humana Press, ニュージャージー, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; および Sequence Analysis Primer, Gribskov, M. および Devereux, J. 監修, M Stockton Press, ニューヨーク, 1991; 並びに Carillo, H., および Lipman, D., 20
 SIAM J. Applied Math., 48:1073 (1988) に記載

されるものを含む、既知の方法によって、容易に計算可能である。同一性を決定するのに好ましい方法を設計して、試験した配列間の最大マッチを生じる。同一性および類似性を決定する方法は、公的に入手可能なコンピュータプログラムにおいて体系化されている。

2つの配列間の同一性および類似性を決定するのに好ましいコンピュータプログラム法には、限定されるわけではないが、GCGプログラムパッケージ (Devereux, J. ら 1984)、BLASTP、BLASTN、および FASTA (Altschul, S. F. ら、1990) が含まれる。BLASTXプログラムは、NCBI および他の供給源から公的に入手可能である (BLAST Manual, Altschul, S. ら, NCBI NLM NIH Bethesda, Md. 20894; A 30
 ltschul, S. ら、1990)。公知のスミス・ウォーターマン・アルゴリズムもまた、同一性を決定するのに使用可能である。

【0077】

例えば、既定の同一性%に関して、偶数番号の配列番号2～668の1つのアミノ酸総数に、それぞれのパーセント同一性のパーセント数値 (100で割ったもの) を乗じて、そしてその後、偶数番号の配列番号2～668の1つのアミノ酸の前記総数からその積を減じることによって、または:

$$n_a \leq x_a - (x_a \cdot y)$$

(式中、 n_a はアミノ酸改変数であり、 x_a は配列番号2～668の1つのアミノ酸総数であり、そして y は例えば70%では0.70、80%では0.80、85%では0.8 40
 5などであり、そして式中、 x_a および y の整数でないいかなる積も、これを x_a から減じる前に、最も近い整数に四捨五入する) によって、アミノ酸改変数を決定可能である。

【0078】

本発明は、 β 溶血性連鎖球菌の株に渡って実質的に保存される単離ポリペプチドを意図する。さらに、 β 溶血性連鎖球菌の株に渡って実質的に保存される単離ポリペプチド、および感受性被験者において、 β 溶血性連鎖球菌コロニー形成または感染を防御するかまたは改善するのに有効な単離ポリペプチドもまた、本発明に意図される。本明細書において、用語「保存」は、例えば、タンパク質中のアミノ酸総数の割合として、挿入、置換および/または欠失を経験しないアミノ酸数を指す。例えば、タンパク質が55%保存され、そして例えば263アミノ酸を有するならば、タンパク質において、アミノ酸が置換を経験 50

しない144のアミノ酸位がある。同様に、タンパク質が90%保存され、そして例えば約280アミノ酸を有するならば、アミノ酸が置換を経験する可能性がある28のアミノ酸位があり、そしてアミノ酸が置換を経験しない252（すなわち280マイナス28）のアミノ酸位がある。本発明の態様にしたがって、単離ポリペプチドは、限定なしに、好ましくは、β溶血性連鎖球菌の株に渡って、少なくとも約80%保存されており、より好ましくは、株に渡って、少なくとも約85%保存されており、さらにより好ましくは、株に渡って、少なくとも約90%保存されており、そして最も好ましくは、株に渡って、少なくとも約95%保存されている。

【0079】

偶数番号の配列番号2～668のポリペプチド構造において、修飾および変化を作成し、そしてなおβ溶血性連鎖球菌および／または化膿性連鎖球菌活性および／または抗原性を有する分子を得ることが可能である。例えば、活性および／または抗原性の認識可能な損失を伴わずに、配列中の他のアミノ酸の代わりに、特定のアミノ酸を使用可能である。ポリペプチドの生物学的な機能的活性を定義するのは、ポリペプチドの相互作用能および性質であるため、特定のアミノ酸配列置換をポリペプチド配列（またはもちろん、その根底にあるDNAコード配列）中で行って、そしてにもかかわらず、同様の特性を持つポリペプチドを得ることが可能である。

【0080】

本発明には、本明細書に記載されるような望ましい反応性を提供する生物学的均等物である、いかなる単離ポリペプチドも含まれる。用語「望ましい反応性」は、本発明の目的のため有用な結果として、当業者に認識されるであろう反応性を指す。望ましい反応性の例は本明細書に記載され、これには、限定なしに、本発明の目的のために有用であると当業者に認識されるであろうような、望ましいレベルの防御、望ましい抗体力価、望ましいオプソニン食作用活性および／または望ましい交差反応性が含まれる。望ましいオプソニン食作用活性は、OPA中のコロニー形成単位（CFU）の陰性対照に対する減少によって測定されるような、殺細菌パーセントによって示される。限定なしに、望ましいオプソニン食作用活性は、好ましくは、少なくとも約15%、より好ましくは、少なくとも約20%、さらにより好ましくは、少なくとも約40%、さらにより好ましくは、少なくとも約50%、そして最も好ましくは、少なくとも約60%である。

【0081】

本発明には、配列番号2～668のアミノ酸配列を含んでなるポリペプチドの変異体であるポリペプチドが含まれる。「変異体」には、この用語が本明細書で用いられる場合、参照ポリペプチドとは異なるが、本質的な特性を保持するポリペプチドが含まれる。一般的に、参照ポリペプチドおよび変異体の配列が全体的に緊密に類似であり、そして多くの領域で同一である（すなわち生物学的に均等である）ように、相違が限定される。変異体および参照ポリペプチドは、いかなる組み合わせでもよい1以上の置換、付加、または欠失によって、アミノ酸配列が異なることも可能である。置換または挿入アミノ酸残基は、遺伝暗号にコードされるものであってもよいし、またはなくてもよい。ポリペプチドの変異体は、対立遺伝子変異体など、天然に存在することも可能であるし、または天然に存在することが知られない変異体であることも可能である。ポリペプチドの非天然存在変異体は、直接合成によって、または突然変異誘発技術によって、作成することも可能である。

【0082】

こうした変化を作成する際、アミノ酸のハイドロパシー指数を考慮することが可能である。ポリペプチドに相互作用生物学的機能を与える際にハイドロパシーアミノ酸指数が重要であることは、当該技術分野に一般的に理解されている（Kyte & Doolittle、1982）。特定のアミノ酸を、類似のハイドロパシー指数またはスコアを有する他のアミノ酸の代わりに使用して、そしてなお、類似の生物学的活性を持つポリペプチドを生じることが可能であると知られる。各アミノ酸は、疎水性および荷電特性に基づいて、ハイドロパシー指数を割り当てられてきている。これらの指数を、以下のように、各アミノ酸の後の括弧内に列挙する：イソロイシン（+4.5）；バリン（+4.2）；ロイ

10

20

30

40

50

シン (+3.8) ; フェニルアラニン (+2.8) ; システイン / システイン (+2.5) ; メチオニン (+1.9) ; アラニン (+1.8) ; グリシン (-0.4) ; スレオニン (-0.7) ; セリン (-0.8) ; トリプトファン (-0.9) ; チロシン (-1.3) ; プロリン (-1.6) ; ヒスチジン (-3.2) ; グルタミン酸 (-3.5) ; グルタミン (-3.5) ; アスパラギン酸 (-3.5) ; アスパラギン (-3.5) ; リジン (-3.9) ; およびアルギニン (-4.5) 。

【0083】

アミノ酸残基の相対ハイドロパシー特性が、生じるポリペプチドの二次および三次構造を決定し、これが次に、ポリペプチドと、酵素、基質、受容体、抗体、抗原等の他の分子との相互作用を特徴付ける。アミノ酸を、類似のハイドロパシー指数を有する別のアミノ酸で代用して、そしてなお、機能的に均等のポリペプチドを得ることが可能であることが、当該技術分野に知られる。こうした変化において、ハイドロパシー指数が+/-2以内であるアミノ酸の置換が好ましく、+/-1以内であるものが特に好ましく、そして+/-0.5以内であるものがさらにより特に好ましい。

10

【0084】

同様のアミノ酸の置換はまた、特に、免疫学的態様で使用するために、生物学的に機能的に均等なポリペプチドまたはペプチドがそれによって生成される場合、親水性に基づいて行うことも可能である。本明細書に援用される米国特許第4,554,101号は、隣接アミノ酸の親水性に支配されるような、ポリペプチドの最大局所平均親水性が、その免疫原性および抗原性、すなわちポリペプチドの生物学的特性と相関すると言及する。

20

【0085】

米国特許第4,554,101号に詳述されるように、以下の親水性値がアミノ酸残基に割り当てられている：アルギニン (+3.0) ; リジン (+3.0) ; アスパラギン酸 (+3.0 ± 1) ; グルタミン酸 (+3.0 ± 1) ; セリン (+0.3) ; アスパラギン (+0.2) ; グルタミン (+0.2) ; グリシン (0) ; プロリン (-0.5 ± 1) ; スレオニン (-0.4) ; アラニン (-0.5) ; ヒスチジン (-0.5) ; システイン (-1.0) ; メチオニン (-1.3) ; バリン (-1.5) ; ロイシン (-1.8) ; イソロイシン (-1.8) ; チロシン (-2.3) ; フェニルアラニン (-2.5) ; およびトリプトファン (-3.4) 。アミノ酸を、類似の親水性値を有する別のものの代わりに使用し、そしてなお生物学的に均等な、そして特に免疫学的に均等なポリペプチドを得ることが可能であると理解される。こうした変化において、親水性値が±2以内であるアミノ酸の置換が好ましく、±1以内のものが特に好ましく、そして±0.5以内のものがさらにより特に好ましい。

30

【0086】

したがって、上に概略するように、アミノ酸置換は、一般的に、アミノ酸側鎖置換基の相対的類似性、例えばその疎水性、親水性、荷電、大きさ等に基づく。前述の多様な特性を考慮に入れる、典型的な置換が当業者に公知であり、そしてこれらには：アルギニンおよびリジン；グルタミン酸およびアスパラギン酸；セリンおよびスレオニン；グルタミンおよびアスパラギン；並びにバリン、ロイシンおよびイソロイシンが含まれる。以下の表X I I I に示すように、適切なアミノ酸置換には、以下が含まれる：

40

表X I I I :

【0087】

【表13】

元来の残基 典型的な残基置換

Ala	Gly; Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Ala
His	Asn; Gln
Ile	Leu; Val

Leu	Ile; Val
Lys	Arg
Met	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

10

20

したがって、本発明には、1以上のアミノ酸置換を含有する配列番号2～668のポリペプチドの機能的または生物学的均等物が含まれる。

【0088】

ポリペプチドの生物学的または機能的均等物はまた、部位特異的突然変異誘発を用いて調製することも可能である。部位特異的突然変異誘発は、根底にあるDNAの特異的突然変異誘発を通じて、その配列に由来した、第二世代ポリペプチド、または生物学的に機能的に均等であるポリペプチドを調製するのに有用な技術である。上述のように、アミノ酸置換が望ましい場合、こうした変化が望ましい可能性がある。該技術は、DNAに1以上のヌクレオチド配列変化を導入することによって、例えば前述の考慮の1以上を取り込んだ配列変異体を容易に調製し、そして試験する能力をさらに提供する。部位特異的突然変異誘発は、望ましい突然変異のDNA配列をコードする特定のオリゴヌクレオチド配列と共に、十分な数の隣接ヌクレオチドの使用を通じて、十分な大きさおよび配列複雑度のプライマー配列を提供して、横切る欠失結合部の両端に安定な二重鎖を形成して、突然変異体の産生を可能にする。典型的には、長さ約17～25ヌクレオチドのプライマーが好ましく、配列結合部両端の約5～10残基が改変される。

30

【0089】

一般的に、部位特異的突然変異誘発技術が当該技術分野に公知である。認識されるであろうように、該技術は、典型的には、一本鎖型および二本鎖型両方で存在することが可能なファージベクターを使用する。典型的には、まず、配列内に、選択した化膿性連鎖球菌ポリペプチド配列のすべてまたは一部をコードするDNA配列を含む一本鎖ベクターを得ることによって、これにしたがって部位特異的突然変異誘発を行う。望ましい突然変異配列を所持するオリゴヌクレオチドプライマーを、例えば公知の技術によって（例えば合成的に）調製する。その後、このプライマーを一本鎖ベクターにアニーリングし、そして突然変異所持鎖の合成を完了するため、大腸菌（*E. coli*）ポリメラーゼIクレノウ断片などの酵素の使用によって伸長する。こうして、1つの鎖が元来の非突然変異配列をコードし、そして第二の鎖が望ましい突然変異を所持する、ヘテロ二重鎖を形成する。その後、このヘテロ二重鎖ベクターを用いて、大腸菌細胞などの適切な細胞を形質転換し、そして突然変異を所持する組換えベクターを含むクローンを選択する。商業的に入手可能なキットは、必要な試薬を提供する。

40

【0090】

50

本発明のポリペプチドおよびポリペプチド抗原は、配列番号 2～668 のいずれかのアミノ酸配列を含んでなるポリペプチドに、実質的な配列類似性、構造的類似性、および／または機能的類似性を含んでなる、いかなるポリペプチドも含むと理解される。さらに、本発明のポリペプチドまたはポリペプチド抗原は、特定の供給源に限定されない。したがって、本発明は、多様な供給源からのポリペプチドの一般的な検出および単離を提供する。

【0091】

本発明のポリペプチドは、さらなる構造的または機能的解析に用いるか、または化膿性連鎖球菌関連ポリペプチドおよび化膿性連鎖球菌特異的抗体などの試薬の生成に用いる断片に、好適に切断することが可能である。これは、本発明の精製または未精製ポリペプチドを、エンドプロテイナーゼ g l u - C (B o e h r i n g e r、インディアナ州インディアナポリス)などのペプチダーゼで処理することによって、達成可能である。C N B rでの処理は、それによって天然化膿性連鎖球菌ポリペプチドからペプチド断片を生成可能な別の方法である。組換え技術もまた、化膿性連鎖球菌ポリペプチドの特定の断片を産生するのに使用可能である。

10

【0092】

さらに、本発明者らは、ペプチド模倣体 (p e p t i d o m i m e t i c s) として当該技術分野に知られるように、特定の化膿性連鎖球菌ポリペプチド抗原に立体的に類似の化合物を構築して、ペプチド構造の重要な部分を模倣することが可能であることを意図する。模倣体は、タンパク質二次構造の要素を模倣するペプチド含有分子である。ペプチド模倣体使用の根底にある論理的根拠は、タンパク質のペプチド主鎖が、主に、受容体およびリガンドなどの、分子相互作用を容易にするような方式で、アミノ酸側鎖を方向付けするように存在することである。

20

【0093】

本発明はまた、本発明の少なくとも 1つのポリペプチドを含んでなる融合タンパク質も含む。「融合タンパク質」は、2つの、しばしば関連しない融合遺伝子またはその断片にコードされるタンパク質を指す。例えば、免疫グロブリン分子の定常領域の多様な部分を、別のヒトタンパク質またはその一部と共に含んでなる融合タンパク質が記載されている。多くの場合、融合タンパク質の一部として免疫グロブリン F c 領域を使用すると、療法および診断における使用に好適であり、例えば薬物動態特性改善を生じる (例えば E P - A 0 2 3 2 2 6 2 1を参照されたい)。一方、いくつかの使用では、融合タンパク質を発現させ、検出し、そして精製した後、F c 部分を欠失可能であることが望ましいであろう。

30

【0094】

本発明のポリペプチドは、「成熟」タンパク質の形であることも可能であるし、または融合タンパク質などのより大きいタンパク質の一部であることも可能である。例えば分泌またはリーダー配列、プロ配列、多数ヒスチジン残基などの精製を補助する配列、または組換え体産生中の安定性のためのさらなる配列を含有する、さらなるアミノ酸配列を含むことがしばしば好適である。

【0095】

化膿性連鎖球菌ポリペプチド断片もまた、本発明に含まれる。断片は、アミノ酸配列の一部と完全に同じであるが、すべてと同じではないアミノ酸配列を有するポリペプチドである。断片は、例えば、偶数番号の配列番号 2～668 のいずれかのアミノ酸配列の、少なくとも 7 以上 (例えば 8、10、12、14、16、18、20 以上) の隣接アミノ酸を含んでなることが可能である。断片は、「独立 (f r e e s t a n d i n g) 」であっても、または最も好ましくは単一の隣接領域として、一部または領域を形成する、より大きいポリペプチド内に含まれてもよい。1つの態様において、断片には、成熟ポリペプチド配列の少なくとも 1つのエピトープが含まれる。

40

【0096】

本発明のポリペプチドは、いかなる適切な方式で調製されることも可能である。こうしたポリペプチドには、天然存在ポリペプチド、組換え的に産生したポリペプチド、合成的に

50

産生したポリペプチド、およびこれらの方法の組み合わせによって産生したポリペプチドが含まれる。こうしたポリペプチドを調製する手段が当該技術分野に公知である。

【0097】

ポリヌクレオチド

本発明はまた、本発明のポリペプチドをコードするヌクレオチド配列を含んでなる単離ポリヌクレオチド、およびこれに緊密に関連するポリヌクレオチドも提供する。これらのポリヌクレオチドには：

(i) 奇数番号の配列番号 1～147 のいずれかのヌクレオチド配列を含んでなる単離ポリヌクレオチド (表 I) ；

(ii) 奇数番号の配列番号 149～181 のいずれかのヌクレオチド配列を含んでなる単離ポリヌクレオチド (表 II) ； 10

(iii) 奇数番号の配列番号 183～187 のいずれかのヌクレオチド配列を含んでなる単離ポリヌクレオチド (表 III) ；および

(iv) 奇数番号の配列番号 189～667 のいずれかのヌクレオチド配列を含んでなる単離ポリヌクレオチド (表 IV) ；

が含まれる。

【0098】

本発明のポリペプチドをコードするポリヌクレオチドは、表 I～IV に含有されるヌクレオチド配列に同一であることも可能であるし、または遺伝暗号の重複 (縮重) の結果、やはり本発明のポリペプチドをコードする変異配列を有することも可能である。 20

【0099】

さらに、本発明は、配列番号 1～667 のヌクレオチド配列に配列同一性を有する単離ポリヌクレオチドを提供する。特定の配列に応じて、配列同一性の度合いは、好ましくは 70% より大きい (例えば 80%、90%、95%、97%、99% 以上)。

【0100】

上に論じるように、「同一性」は、当該技術分野に知られるとおり、配列を比較することによって決定されるように、2 以上のポリペプチド配列または 2 以上のポリヌクレオチド配列間の関係である。「同一性」は、既知の方法によって、容易に計算可能である。例えば、本発明のポリヌクレオチド配列は、奇数番号の配列番号 1～667 の参照ヌクレオチド配列に同一、すなわち 100% 同一であることも可能であるし、または参照ヌクレオチド配列に比較した際、ヌクレオチド改変の特定の整数までを含むことも可能である。こうした改変には、少なくとも 1 つのヌクレオチド欠失、トランジションおよびトランスバージョンを含む置換、または挿入が含まれる。改変は、参照ヌクレオチド配列の 5' 末端もしくは 3' 末端位で、または末端位の間のどこかで、参照配列中のヌクレオチドの間に個々に点在するか、または参照ヌクレオチド配列内の 1 以上の隣接基に生じることが可能である。奇数番号の配列番号 1～667 の 1 つのヌクレオチド総数に、それぞれのパーセント同一性のパーセント数値 (100 で割ったもの) を乗じて、そして奇数番号の配列番号 1～667 のいずれかの参照ヌクレオチド配列のヌクレオチドの前記総数からその積を減じることによって、ヌクレオチド改変数を決定する。 30

【0101】

例えば、奇数番号の配列番号 1～667 の 1 つのヌクレオチド配列に、少なくとも 70% の同一性を有するポリヌクレオチドに関しては、該ポリヌクレオチドは、奇数番号の配列番号 1～667 の 1 つのヌクレオチド配列の全長に渡って、 n_n 核酸改変までを含むことが可能であり、ここで、 n_n は式：

$$n_n \leq x_n - (x_n \cdot y)$$

によって計算され、そして式中、 x_n は奇数番号の配列番号 1～667 の 1 つのヌクレオチド配列のヌクレオチド総数であり、 y は 0.70 の値を有し、そして式中、 x_n および y の整数でないかなる積も、こうした積を x_n から減じる前に、最も近い整数に四捨五入する。もちろん、 y はまた、80% では 0.80、85% では 0.85、90% では 0.90、95% では 0.95 などの値を有することも可能である。 40 50

【0102】

本発明にはまた、いずれかの組み合わせで、1以上のアミノ酸残基が置換され、欠失され、または付加されるが、天然ポリペプチドの生物学的活性を保持する、配列番号2～668のアミノ酸配列を含んでなるポリペプチドのポリペプチド変異体をコードするポリヌクレオチドも含まれる。「変異体」は、この用語が本明細書で用いられる場合、参照ポリヌクレオチドとは異なるが、本質的な特性を保持するポリヌクレオチドである。変異体のヌクレオチド配列変化は、参照ポリヌクレオチドにコードされるポリペプチドのアミノ酸配列を改変してもまたはしなくてもよい。ヌクレオチド変化は、参照配列にコードされるポリペプチドのアミノ酸置換、付加、欠失、融合、および一部切除 (t r u n c a t i o n) を生じることが可能である。ポリヌクレオチド変異体は、対立遺伝子変異体など、天然存在であることも可能であるし、または、天然存在であることが知られていない変異体であることも可能である。ポリヌクレオチドの非天然存在変異体は、突然変異誘発技術によって、または直接合成によって、作成可能である。

【0103】

本発明にはまた、減少したストリンジェンシー条件下で、より好ましくはストリンジェントな条件下で、そして最も好ましくは、非常にストリンジェントな条件下で、本明細書記載のポリヌクレオチドにハイブリダイズ可能なポリヌクレオチドも含まれる。ストリンジェンシー条件の例を、以下のストリンジェンシー条件表に示す：非常にストリンジェントな条件は、少なくとも、例えば条件A～Fと同程度にストリンジェントなものであり；ストリンジェントな条件は、少なくとも、例えば条件G～Lと同程度にストリンジェントなものであり；そして減少したストリンジェンシー条件は、少なくとも、例えば条件M～Rと同程度にストリンジェントなものである。

【0104】

表XIV ストリンジェンシー条件表

【0105】

【表14】

ストリンジェンシー条件	ポリヌクレオチドハイブリッド	ハイブリッド長 (bp) ^I	ハイブリダイゼーション温度および緩衝液 ^{II}	洗浄温度および緩衝液 ^{II}
A	DNA:DNA	> 50	65°C; 1xSSC または 42°C; 1xSSC, 50%ホルムアミド	65°C; 0.3xSSC
B	DNA:DNA	< 50	T _B ; 1xSSC	T _B ; 1xSSC
C	DNA:RNA	> 50	67°C; 1xSSC または 45°C; 1xSSC, 50%ホルムアミド	67°C; 0.3xSSC
D	DNA:RNA	< 50	T _D ; 1xSSC	T _D ; 1xSSC
E	RNA:RNA	> 50	70°C; 1xSSC または 50°C; 1xSSC, 50%ホルムアミド	70°C; 0.3xSSC

ストリンジェンシー条件	ポリヌクレオチドハイブリッド	ハイブリッド長 (bp) ^I	ハイブリダイゼーション温度および緩衝液 ^{II}	洗浄温度および緩衝液 ^{II}
F	RNA:RNA	< 50	T _F ; 1xSSC	T _F ; 1xSSC
G	DNA:DNA	> 50	65°C; 4xSSC または 42°C; 4xSSC, 50%ホルムアミド	65°C; 1xSSC
H	DNA:DNA	< 50	T _H ; 4xSSC	T _H ; 4xSSC
I	DNA:RNA	> 50	67°C; 4xSSC または 45°C; 4xSSC, 50%ホルムアミド	67°C; 1xSSC
J	DNA:RNA	< 50	T _J ; 4xSSC	T _J ; 4xSSC
K	RNA:RNA	> 50	70°C; 4xSSC または 50°C; 4xSSC, 50%ホルムアミド	67°C; 1xSSC
L	RNA:RNA	< 50	T _L ; 2xSSC	T _L ; 2xSSC
M	DNA:DNA	> 50	50°C; 4xSSC または 40°C; 6xSSC, 50%ホルムアミド	50°C; 2xSSC
N	DNA:DNA	< 50	T _N ; 6xSSC	T _N ; 6xSSC
O	DNA:RNA	> 50	55°C; 4xSSC または 42°C; 6xSSC, 50%ホルムアミド	55°C; 2xSSC
P	DNA:RNA	< 50	T _P ; 6xSSC	T _P ; 6xSSC
Q	RNA:RNA	> 50	60°C; 4xSSC または 45°C; 6xSSC, 50%ホルムアミド	60°C; 2xSSC
R	RNA:RNA	< 50	T _R ; 4xSSC	T _R ; 4xSSC

10

20

b p^I : ハイブリッド長は、ハイブリダイズしているポリヌクレオチドのハイブリダイズ領域（類）に関して予測されるものである。未知の配列の標的ポリヌクレオチドにポリヌクレオチドをハイブリダイズさせた際、ハイブリッド長はハイブリダイズしているポリヌクレオチドのものと仮定される。既知の配列のポリヌクレオチドがハイブリダイズした場合、ハイブリッド長は、ポリヌクレオチド配列を並列させ、そして最適配列相補性の領域または領域類を同定することによって決定可能である。

【0106】

緩衝液^H : ハイブリダイゼーション緩衝液および洗浄緩衝液中、SSPE（1xSSPEは、0.15M NaCl、10mM NaH₂PO₄、および1.25mM EDTA、pH7.4である）をSSC（1xSSCは、0.15M NaClおよび15mMクエン酸ナトリウムである）の代わりに使用可能であり；ハイブリダイゼーション完了後、洗浄を15分間行う。

【0107】

T_B から T_R : 長さ50塩基対未満であると予測されるハイブリッドのハイブリダイゼーション温度は、ハイブリッドの融点（T_m）より5～10°C低いはずであり、T_mは、以下の等式にしたがって決定する。長さ18塩基対未満のハイブリッドに関しては、T_m（°C）= 2（A+T塩基数）+ 4（G+C塩基数）。長さ18から49塩基対の間のハイブリッドに関しては、T_m（°C）= 81.5 + 16.6（log₁₀ [Na⁺]）+ 0.41（%G+C）-（600/N）、式中、Nはハイブリッド中の塩基数であり、そして[Na⁺]はハイブリダイゼーション緩衝液中のナトリウムイオン濃度である（1xSSCの[Na⁺] = 0.165M）。

30

40

50

【0108】

ポリヌクレオチドハイブリダイゼーションのストリンジェンシー条件のさらなる例が、本明細書に援用される、Sambrook, J., E. F. Fritsch, および T. Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, ニューヨーク州コールドスプリングハーバー, 第9章および第11章、並びにCurrent Protocols in Molecular Biology, 1995, F. M. Ausubelら監修, John Wiley & Sons, Inc., セクション2.10および6.3-6.4に提供される。

10

【0109】

本発明はまた、これらのポリヌクレオチドに完全に相補的なポリヌクレオチドも提供し、そしてまた、アンチセンス配列も提供する。アンチセンスオリゴヌクレオチドとも称される、本発明のアンチセンス配列には、本発明のポリペプチドをコードするポリヌクレオチドの発現を遮断する、内部で生成される配列および外部から投与される配列両方が含まれる。本発明のアンチセンス配列は、例えば、約15～20塩基対を含んでなる。例えば、上流非翻訳配列に結合するプロモーターを妨げることによって、またはリボソームが結合することを妨げることにより本発明のポリペプチドをコードする転写物の翻訳を妨げることによって、転写を阻害するように、アンチセンス配列を設計することが可能である。

【0110】

本発明のポリヌクレオチドは多くの方式で（例えば化学合成によって、DNAライブラリーから、生物自体から）調製され、そして多様な型（例えば一本鎖、二本鎖、ベクター、プローブ、プライマー）を取ることが可能である。用語「ポリヌクレオチド」には、DNAおよびRNA、並びに修飾主鎖を含有するものなどの、その類似体もまた含まれる。

20

【0111】

本発明のポリヌクレオチドを、ポリペプチドの組換え産生に用いる場合、ポリヌクレオチドは、成熟ポリペプチドまたはその断片のコード配列のみ、リーダーまたは分泌配列、プレ、プロ、またはプレプロタンパク質配列、または他の融合タンパク質部分をコードするものなどの他のコード配列と同じ読み枠にある成熟ポリペプチドまたは断片のコード配列を含むことが可能である。例えば、融合ポリペプチドの精製を促進するマーカー配列をコード配列に連結することが可能である。ポリヌクレオチドはまた、転写、非翻訳配列、スプライシングおよびポリアデニル化シグナル、リボソーム結合部位、およびmRNAを安定化する配列などの、非コード5'および3'配列も含有することが可能である。

30

【0112】

発現系およびベクター

組換え体産生には、一般的に、宿主細胞を操作して、本発明の発現系、その一部、またはポリヌクレオチドを取り込む。カルシウムリン酸トランスフェクション、DEAEデキストラン仲介トランスフェクション、トランスベクション (transfection)、マイクロインジェクション、超音波、カチオン性脂質仲介トランスフェクション、エレクトロポレーション、形質導入、スクレイプ装填 (scrape loading)、弾道導入 (ballistic introduction)、または感染など、Davisら, BASIC METHODS IN MOLECULAR BIOLOGY (1986) およびSambrookら, MOLECULAR CLONING: A LABORATORY MANUAL, 第2版, Cold Spring Harbor Laboratory Press, ニューヨーク州コールドスプリングハーバー (1989) などの多くの標準的実験マニュアルに記載される方法によって、宿主細胞へのポリヌクレオチドの導入を達成する。

40

【0113】

適切な宿主の代表的な例には、細菌細胞（例えば連鎖球菌、ブドウ球菌、大腸菌、ストレプトミセス属 (Streptomyces) および枯草菌 (Bacillus subt

50

i l i s) 細胞)、酵母細胞 (例えばピキア属 (P i c h i a)、サッカロミセス属 (S a c c h a r o m y c e s))、哺乳動物細胞 (例えばベロ、チャイニーズハムスター卵巣、ニワトリ胚線維芽細胞、BHK細胞、ヒトSW13細胞)、および昆虫細胞 (例えばS f 9、S f 21) が含まれる。

【0114】

高性能液体クロマトグラフィー、硫酸アンモニウムまたはエタノール沈殿、酸抽出、陰イオンまたは陽イオン交換クロマトグラフィー、ホスホセルロースクロマトグラフィー、疎水性相互作用クロマトグラフィー、アフィニティークロマトグラフィー、ヒドロキシアパタイトクロマトグラフィー、およびレクチンクロマトグラフィーを含む公知の方法によって、組換え細胞培養から、組換え的に産生されたポリペプチドを回収し、そして精製する 10

【0115】

非常に多様な発現系が用いられる。こうした系には、とりわけ、染色体、エピソームおよびウイルスに由来する系、例えば細菌プラスミド、サルモネラ属 (S a l m o n e l l a) などの弱毒細菌 (米国特許第4, 837, 151号) 由来、バクテリオファージ由来、トランスポゾン由来、酵母エピソーム由来、挿入要素由来、酵母染色体要素由来、ワクシニアおよび他のポックスウイルス、シンドビス、アデノウイルス、バキュロウイルス、SV40などのパポバウイルス、鶏痘ウイルス、偽狂犬病ウイルスおよびレトロウイルス、ベネズエラウマ脳炎ウイルス (米国特許第5, 643, 576号) などのアルファウイルス、小水疱性口内炎ウイルス (米国特許第6, 168, 943号) などの非分割マイナス鎖RNAウイルスなどのウイルス由来のベクター、並びにコスミドおよびファージミドなどのプラスミドおよびバクテリオファージ遺伝子要素由来のものなど、その組み合わせ由来のベクターが含まれる。発現系には、プロモーターおよび他の制御要素など (ポリアデニル化シグナルなど) の、発現を制御すると共に引き起こす調節領域が含まなければならない。一般的に、宿主において、ポリヌクレオチドを維持するか、増やすか、または発現してポリペプチドを産生するのに適した、いかなる系またはベクターも使用可能である。例えばS a m b r o o kら, M O L E C U L A R C L O N I N G, A L A B O R A T O R Y M A N U A L (上記) に示すものなどの、多様な公知のそして日常的な技術のいずれによって、適切なヌクレオチド配列を発現系に挿入することも可能である。 20

【0116】

本発明はまた、本発明のポリヌクレオチドまたはポリヌクレオチド類を含んでなるベクター (例えば発現ベクター、配列決定ベクター、クローニングベクター)、本発明のベクターで遺伝的に操作される宿主細胞、および組換え技術による本発明のポリペプチドの産生も提供する。細胞不含翻訳系もまた、本発明のDNA構築物由来のRNAを用いて、こうしたタンパク質を産生するのに使用可能である。 30

【0117】

好ましいベクターは、レンチウイルス、レトロウイルス、ヘルペスウイルス、アデノウイルス、アデノ関連ウイルス、ワクシニアウイルス、バキュロウイルス、および望ましい細胞向性を持つ他の組換えウイルスなどのウイルスベクターである。したがって、ウイルスベクターを用いて、またはDNAの直接導入を通じて、機能するまたは突然変異体タンパク質またはポリペプチド、あるいはその断片をコードする遺伝子を、i n v i v o、e x v i v o、またはi n v i t r oで導入することが可能である。標的とされる組織における発現は、ウイルスベクターまたは受容体リガンドを用いるなど、トランスジェニックベクターで特定の細胞を標的とすることによって、または組織特異的プロモーターを用いることによって、あるいはその両方などで、達成可能である。標的化遺伝子搬送は、PCT公開番号WO 95/28494号に記載される。 40

【0118】

i n v i v oまたはe x v i v o標的化および療法に一般的に用いられるウイルスベクターは、DNAに基づくベクターおよびレトロウイルスベクターである。ウイルスベクターを構築し、そして用いるための方法が当該技術分野に知られる (例えばM i l l e r 50

およびRosman, BioTechniques, 1992, 7:980-990)。好ましくは、ウイルスベクターは複製不全であり、すなわち、これらは標的細胞において、自立的に複製することが不能である。好ましくは、複製不全ウイルスは最小ウイルスであり、すなわち、ゲノムを被包してウイルス粒子を生じるのに必要なゲノム配列のみを保持する。

【0119】

DNAウイルスベクターには、弱毒または欠陥DNAウイルス、例えば限定されるわけではないが、単純疱疹ウイルス(HSV)、パピローマウイルス、エプスタイン・バーウイルス(EBV)、アデノウイルス、アデノ関連ウイルス(AAV)等が含まれる。ウイルス遺伝子を完全に、またはほぼ完全に欠く欠陥ウイルスが好ましい。欠陥ウイルスは、細胞への導入後、感染性でない。欠陥ウイルスベクターを使用すると、ベクターが他の細胞に感染する可能性を懸念することなく、特定の局所領域において、細胞に投与することが可能である。したがって、特定の組織を特異的に標的とすることが可能である。特定のベクターの例には、限定されるわけではないが、欠陥ヘルペスウイルス1(HSV1)ベクター(Kaplittrら, Molec. Cell. Neurosci., 1991, 2:320-330)、糖タンパク質L遺伝子を欠く欠陥ヘルペスウイルスベクター、または他の欠陥ヘルペスウイルスベクター(PCT公開番号WO 94/21807号および第WO 92/05263号); Stratford-Pericaudetら(J. Clin. Invest., 1992, 90:626-630; LaSalleら, Science, 1993, 259:988-990もまた参照されたい)に記載されるベクターなどの弱毒アデノウイルスベクター; および欠陥アデノ関連ウイルスベクター(Samulskiら, J. Virol., 1987, 61:3096-3101; Samulskiら, J. Virol., 1989, 63:3822-3828; Lebkowskiら, Mol. Cell. Biol., 1988, 8:3988-3996)が含まれる。

10

20

【0120】

限定されるわけではないが、Avigen, Inc. (カリフォルニア州アラメダ; AAVベクター)、Cell Genesys (カリフォルニア州フォスターシティ; レトロウイルス、アデノウイルス、AAVベクター、およびレンチウイルスベクター)、Clontech (レトロウイルスおよびバキュロウイルスベクター)、Genovo, Inc. (ペンシルバニア州シャロンヒル; アデノウイルスおよびAAVベクター)、Genvec (アデノウイルスベクター)、IntroGene (オランダ・ライデン; アデノウイルスベクター)、Molecular Medicine (レトロウイルス、アデノウイルス、AAV、およびヘルペスウイルスベクター)、Norgen (アデノウイルスベクター)、Oxford Biomedica (英国オックスフォード; レンチウイルスベクター)、およびTransgene (フランス・ストラズブール; アデノウイルス、ワクシニア、レトロウイルス、およびレンチウイルスベクター)を含む多様な企業が商業的にウイルスベクターを生産している。

30

【0121】

アデノウイルスは、本発明のヌクレオチドを、多様な細胞種に効率的に搬送するよう修飾可能な真核生物DNAウイルスである。アデノウイルスの多様な血清型が存在する。これらの血清型のうち、本発明の範囲内で、2型または5型ヒトアデノウイルス(Ad2またはAd5)あるいは動物起源のアデノウイルス(PCT公開番号WO 94/26914号)を用いるのに優先性が与えられる。本発明の範囲内で使用可能な動物起源のアデノウイルスには、イヌ、ウシ、ネズミ(例えばMav1、Beardら, Virology, 1990, 75-81)、ヒツジ、ブタ、鳥、およびサル(例えばSAV)起源のアデノウイルスが含まれる。好ましくは、動物起源のアデノウイルスはイヌアデノウイルスであり、より好ましくはCAV2アデノウイルスである(例えばマンハッタンまたはA26/61株、ATCC VR-800)。多様な複製不全アデノウイルスおよび最小アデノウイルスベクターが記載されてきている(例えばPCT公開番号WO 94/269

40

50

14号、WO 95/02697号、WO 94/28938号、WO 94/28152号、WO 94/12649号、WO 95/02697号、WO 96/22378号）。本発明にしたがった複製不全組換えアデノウイルスは、当業者に知られるいかなる技術によって調製することも可能である（例えばLevreroら、Gene, 1991, 101:195；欧州公開番号EP 185 573号；Graham, EMBO J., 1984, 3:2917；Grahamら、J. Gen. Virol., 1977, 36:59）。通常の当業者に公知の標準的分子生物学的技術を用いて、組換えアデノウイルスを回収し、そして精製する。

【0122】

アデノ関連ウイルス（AAV）は、感染した細胞のゲノムに、安定で、そして部位特異的な方式で、組み込まれることが可能な、比較的小さいDNAウイルスである。これらは、細胞増殖、形態、または分化にいかなる影響も誘導することなく、広い範囲の細胞に感染可能であり、そしてヒト病理に関与していないようである。AAVゲノムがクローニングされ、配列決定され、そして性質決定されてきている。遺伝子を*in vitro*および*in vivo*でトランスファーするためのAAV由来のベクターの使用が記載されている（PCT公開番号WO 91/18088号およびWO 93/09239号；米国特許第4,797,368号および第5,139,941号；欧州公開番号EP 488 528号を参照されたい）。本発明にしたがった複製不全組換えAAVは、2つのAAV逆方向末端反復（ITR）領域が隣接する、目的の核酸配列を含有するプラスミド、およびAAVキャプシド形成遺伝子（*rep*および*cap*遺伝子）を所持するプラスミドを、ヒトヘルパーウイルス（例えばアデノウイルス）に感染した細胞株に同時トランスフェクションすることによって調製可能である。その後、産生されるAAV組換え体を標準的技術によって精製する。

【0123】

別の態様において、例えば、米国特許第5,399,346号；Mannら、Cell, 1983, 33:153；米国特許第4,650,764号および第4,980,289号；Markowitzら、J. Virol., 1988, 62:1120；米国特許第5,124,263号；欧州公開番号EP 453 242号および第EP 178 220号；Bernsteinら、Genet. Eng., 1985, 7:235；McCormick, BioTechnology, 1985, 3:689；PCT公開番号WO 95/07358号；およびKuoら、Blood, 1993, 82:845に記載されるように、遺伝子をレトロウイルスベクターに導入可能である。レトロウイルスは、分裂細胞に感染する組込みウイルスである。レトロウイルスゲノムには、2つのLTR、キャプシド形成配列、および3つのコード領域（*gag*、*pol*および*env*）が含まれる。組換えレトロウイルスベクターにおいて、*gag*、*pol*および*env*遺伝子は、一般的に、すべてまたは部分的に欠失しており、そして目的の異種核酸配列で置換されている。これらのベクターは、HIV、MoMuLV（「ネズミモロニー白血病ウイルス」）、MSV（「ネズミモロニー肉腫ウイルス」）、HaSV（「ハーベイ肉腫ウイルス」）、SNV（「脾臓壊死ウイルス」）、RSV（「ラウス肉腫ウイルス」）、およびフレンドウイルスなどの異なる種類のレトロウイルスから構築可能である。適切なパッケージング細胞株、特に細胞株PA317（米国特許第4,861,719号）、PsiCRIP細胞株（PCT公開番号WO 90/02806号）、およびGP+*env* Am-12細胞株（PCT公開番号WO 89/07150号）が記載されている。さらに、組換えレトロウイルスベクターは、転写活性を抑制するため、LTR内に修飾を含有すると共に、*gag*遺伝子の一部を含むことが可能な広範囲のキャプシド形成配列を含有することが可能である（Benderら、J. Virol., 1987, 61:1639）。一般の当業者に知られる標準的技術によって、組換えレトロウイルスベクターを精製する。

【0124】

感染性粒子として機能するか、または単回トランスフェクションを経るように、レトロウ

イルスベクターを構築可能である。前者の場合、発癌性形質転換特性に関与するものを除いて、遺伝子すべてを保持し、そして異種遺伝子を発現するように、ウイルスを修飾する。ウイルスパッケージングシグナルを破壊するが、異種遺伝子およびパッケージングシグナルを含有するように操作した同時導入ウイルスをパッケージングするのに必要な構造遺伝子を保持するように、非感染性ウイルスベクターを操作する。したがって、生じるウイルス粒子は、さらなるウイルスを産生することが不可能である。

【0125】

レトロウイルスベクターはまた、1周期のレトロウイルス複製を可能にし、そしてトランスフェクション効率を増幅する、DNAウイルスによって導入することも可能である（PCT公開番号WO 95/22617号、WO 95/26411号、WO 96/39036号およびWO 97/19182号を参照されたい）。 10

【0126】

別の態様において、レンチウイルスベクターは、直接搬送、並びに脳、網膜、筋肉、肝臓、および血液を含む、いくつかの組織種における、導入遺伝子の発現維持のための剤として使用可能である。ベクターは、これらの組織において、分裂および非分裂細胞に効率的に形質導入し、そして目的の遺伝子の長期発現を維持することが可能である。概説には、Naldini, Curr. Opin. Biotechnol., 1998, 9:457-63を参照されたい；Zuffereyら, J. Virol., 1998, 72:9873-80もまた参照されたい。レンチウイルスパッケージング細胞株が入手可能であり、そして一般的に当該技術分野に知られる。これらは、遺伝子治療のため、高力価レンチウイルスベクターの産生を促進する。例は、少なくとも3〜4日間、106 IU/mlより高い力価でウイルス粒子を生成可能な、テトラサイクリン誘導性VSV-G偽型レンチウイルスパッケージング細胞株である（Kafriら, J. Virol., 1999, 73:576-584）。誘導性細胞株に産生されるベクターは、必要に応じてin vitroおよびin vivoで非分裂細胞に効率的に形質導入するために濃縮可能である。 20

【0127】

別の態様において、リポフェクションによって、裸のDNAとして、または他のトランスフェクション促進剤（ペプチド、ポリマーなど）と共に、in vivoでベクターを導入可能である。合成カチオン性脂質を用いて、マーカーをコードする遺伝子のin vivoトランスフェクションのためのリポソームを調製することが可能である（Felgnerら, Proc. Natl. Acad. Sci. U. S. A., 1987, 84:7413-7417；FelgnerおよびRingold, Science, 1989, 337:387-388；Mackeyら, Proc. Natl. Acad. Sci. U. S. A., 1988, 85:8027-8031；Ulmerら, Science, 1993, 259:1745-1748）。核酸トランスファーのための有用な脂質化合物および組成物は、PCT特許公開番号WO 95/18863号およびWO 96/17823号、および米国特許第5,459,127号に記載される。標的化の目的のため、脂質を他の分子に化学的にカップリングすることが可能である（Mackeyら、上記を参照されたい）。標的化したペプチド、例えばホルモンまたは神経伝達物質、および抗体などのタンパク質、または非ペプチド分子を、リポソームに化学的にカップリングすることが可能である。 30 40

【0128】

また、in vivoでベクターを裸のDNAプラスミドとして導入することも可能である。当該技術分野に知られる方法、例えばエレクトロポレーション、マイクロインジェクション、細胞融合、DEAEデキストラン、リン酸カルシウム沈殿、遺伝子銃の使用、またはDNAベクター輸送体の使用によって、遺伝子治療のための裸のDNAベクターを望ましい宿主細胞に導入することが可能である（例えばWuら, J. Biol. Chem., 1992, 267:963-967；WuおよびWu, J. Biol. Chem., 1988, 263:14621-14624；カナダ特許出願第2, 50

012, 311号; Williamsら, Proc. Natl. Acad. Sci. USA, 1991, 88:2726-2730)。受容体仲介DNA搬送アプローチもまた使用可能である(Curielら, Hum. Gene Ther., 1992, 3:147-154; WuおよびWu, J. Biol. Chem., 1987, 262:4429-4432)。米国特許第5, 580, 859号および第5, 589, 466号は、哺乳動物において、トランスフェクション促進剤を含まない、外因性DNA配列の搬送を開示する。最近、エレクトロトランスファーと称する、比較的低い電圧で高い効率のin vivo DNAトランスファー技術が記載されてきている(Mirら, C. P. Acad. Sci., 1988, 321:893; PCT公開番号WO 99/01157号; WO 99/01158号; WO 99/01175号)。

【0129】

カチオン性オリゴペプチド(例えばPCT特許公開番号WO 95/21931号)、DNA結合タンパク質由来ペプチド(例えばPCT特許公開番号WO 96/25508号)、またはカチオン性ポリマー(例えばPCT特許公開番号WO 95/21931号)またはブピバカイン(米国特許第5, 593, 972号)などの他の分子もまた、in vivoで核酸トランスフェクションを促進するのに有用である。

【0130】

本発明の単離ポリペプチドは、外来ポリペプチドとしてポリペプチドまたは免疫原断片の発現に必要な遺伝子材料を含有する、生存ベクターを用いて、特に生存組換え細菌、ウイルス、または他の生存病原体を用いて、哺乳動物に搬送することが可能である。特に、サルモネラ属、赤痢菌属(Shigella)、エルシニア属(Yersinia)、ビブリオ属(Vibrio)、エシェリキア属(Escherichia)およびBCGなどの、胃腸管にコロニー形成する細菌がワクチンベクターとして開発されてきており、そしてこれらおよび他の例がHolmgrenら(1992)およびMcGheeら(1992)に論じられる。

【0131】

以下は、1以上の免疫原候補タンパク質を挿入可能な、RNAベクターリストの一部として使用可能である。

モノネガウイルス目(Mononegavirales)の非分割マイナス鎖RNAウイルスの分類

パラミクソウイルス科

パラミクソウイルス亜科

パラミクソウイルス属

センダイウイルス(マウスパラインフルエンザウイルス1型)

ヒトパラインフルエンザウイルス(PIV)1型および3型

ウシパラインフルエンザウイルス(BPV)3型

ルブラウイルス属

サルウイルス5(SV)(イヌパラインフルエンザウイルス2型)

ムンプスウイルス

ニューキャッスル病ウイルス(NDV)(鳥パラミクソウイルス1)

ヒトパラインフルエンザウイルス(PIV-2、4aおよび4b型)

麻疹ウイルス属

麻疹ウイルス(MV)

イルカ麻疹ウイルス

イヌジステンパーウイルス(CDV)

小反芻獣疫ウイルス

アザラシ(phocine)ジステンパーウイルス

牛疫ウイルス

未分類

10

20

30

40

50

ヘンドラウイルス
 ニパウイルス
 ニューモウイルス亜科
 ニューモウイルス属
 ヒト呼吸器合胞体ウイルス (RSV)
 ウシ呼吸器合胞体ウイルス
 マウス肺炎ウイルス
 メタニューモウイルス属
 ヒトメタニューモウイルス
 鳥ニューモウイルス (以前の七面鳥鼻気管炎ウイルス)

10

ラブドウイルス科
 狂犬病ウイルス属
 狂犬病ウイルス
 ベシキュロウイルス属 (Vesiculovirus)
 水泡性口内炎ウイルス (VSV)
 エフェメロウイルス属 (Ephemerovirus)
 ウシ流行熱ウイルス

フィロウイルス科

フィロウイルス属
 マールブルグウイルス

20

RNAウイルスベクターは、基本的に、モノネガウイルス目の非分割マイナスセンス一本鎖RNAウイルスの少なくとも1つのゲノムまたはアンチゲノムをコードする配列を含んでなる単離核酸分子である。単離核酸分子は、ゲノム、アンチゲノム、またはその修飾型をコードするポリヌクレオチド配列を含んでなる可能性がある。1つの態様において、ポリヌクレオチドは、機能可能であるように連結されたプロモーター、望ましいゲノムまたはアンチゲノム、および転写ターミネーターをコードする。

【0132】

本発明の好ましい態様において、ポリヌクレオチドは、ヌクレオチド挿入、再配置、欠失、または置換によって、野生型RNAウイルスから修飾されているゲノムまたはアンチゲノムをコードする。ゲノムまたはアンチゲノム配列は、ヒトまたは非ヒトウイルス由来であることが可能である。ポリヌクレオチド配列はまた、2以上の供給源由来のゲノムまたはアンチゲノムを組換え的に連結することから形成するキメラゲノムをコードすることも可能である。例えば、RSVのA群由来の1以上の遺伝子を、RSVのB群の対応する遺伝子の代わりに挿入する；またはウシPIV (BPIV)、PIV-1またはPIV-2由来の1以上の遺伝子を、PIV-3の対応する遺伝子の代わりに挿入する；またはRSVをPIVの遺伝子と置換することが可能であるなどである。さらなる態様において、ポリヌクレオチドは、ヒト、ウシまたはネズミウイルスである、モノネガウイルス目のRNAウイルスのゲノムまたはアンチゲノムをコードする。本発明の方法によって形成される組換えウイルスは、療法または予防目的で使用されるため、ポリヌクレオチドはまた、選択したRNAウイルスの弱毒型または感染型をコードすることも可能である。多くの態様において、ポリヌクレオチドは、RNAウイルスの弱毒感染型をコードする。特に好ましい態様において、ポリヌクレオチドは、本明細書に援用される、公開国際特許出願WO 98/13501に記載されるように、3'ゲノムプロモーター領域に少なくとも1つの弱毒化突然変異を有し、そしてRNAポリメラーゼ遺伝子中に少なくとも1つの弱毒化突然変異を有する、モノネガウイルス目の非分割マイナス鎖一本鎖RNAウイルスのゲノムまたはアンチゲノムをコードする。

30

40

【0133】

ベクターとして、上述のような望ましいゲノムおよびアンチゲノムの修飾型をコードするポリヌクレオチド配列はまた、本発明の免疫原タンパク質の1以上の遺伝子またはヌクレオチド配列もコードする。さらに、1以上の異種遺伝子もまた、望ましいように、望まし

50

い免疫原組成物／ベクターを形成する際に含むことが可能である。望ましい組換えウイルスの適用に応じて、異種遺伝子は、補因子、サイトカイン（インターロイキンなど）、Tヘルパーエпитープ、制限マーカー、アジュバント、または異なる微生物病原体（例えばウイルス、細菌、または真菌）のタンパク質、特に防御免疫応答を誘発可能なタンパク質をコードすることが可能である。異種遺伝子はまた、遺伝子治療に用いる剤を提供するのにも使用可能である。好ましい態様において、異種遺伝子は、組換えウイルスの予防または療法特性を改善するよう選択される、インターロイキン-12などのサイトカインをコードする。

【0134】

抗体

偶数番号の配列番号2～668のアミノ酸配列、その断片、およびその類似体を含む、本発明のポリペプチド、またはこれらを発現する細胞はまた、本発明のポリペプチドに免疫特異的な抗体を産生する免疫原としても使用可能である。本発明には、β溶血性連鎖球菌および化膿性連鎖球菌ポリペプチドに免疫特異的な抗体、並びに細胞、細胞または組織抽出液、あるいは生物学的液体中のβ溶血性連鎖球菌および化膿性連鎖球菌ポリペプチドの存在を検出するか、あるいはその量または濃度を測定するためのこうした抗体の使用が含まれる。

【0135】

本発明の抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、および抗イディオタイプ抗体が含まれる。ポリクローナル抗体は、抗原で免疫した動物の血清由来の抗体分子の不均一集団である。モノクローナル抗体は、特定の抗原に対する、実質的に均質の抗体集団である。当業者に知られる方法、例えばKohlerおよびMilstein, 1975, Nature 256:495-497および米国特許第4,376,110号によって、モノクローナル抗体を得ることが可能である。こうした抗体は、IgG、IgM、IgE、IgA、GILDおよびそのサブクラスいずれかを含む、いかなる免疫グロブリンクラスであることも可能である。

【0136】

キメラ抗体は、その異なる部分が異なる動物種に由来する分子であり、例えば、ネズミモノクローナル抗体由来の可変領域およびヒト免疫グロブリン定常領域を有するものである。キメラ抗体およびその産生法は当該技術分野に公知である（Cabilllyら, 1984, Proc. Natl. Acad. Sci. USA 81:3273-3277; Morrisonら, 1984, Proc. Natl. Acad. Sci. USA 81:6851-6855; Boulianneら, 1984, Nature 312:643-646; Cabilllyら、欧州特許出願125023（1984年11月14日公開）; Taniguchiら、欧州特許出願171496（1985年2月19日公開）; Morrisonら、欧州特許出願173494（1986年3月5日公開）; Neubergerら、PCT出願WO 86/01533（1986年3月13日公開）; Kudoら、欧州特許出願184187（1986年6月11日公開）; Morrisonら、欧州特許出願173494（1986年3月5日公開）; Sahaganら, 1986, J. Immunol. 137:1066-1074; Robinsonら、PCT/US 86/02269（1987年5月7日公開）; Liuら, 1987, Proc. Natl. Acad. Sci. USA 84:3439-3443; Sunら, 1987, Proc. Natl. Acad. Sci. USA 84:214-218; Betterら, 1988, Science 240:1041-1043）。これらの参考文献は本明細書に援用される。

【0137】

抗イディオタイプ（抗Id）抗体は、抗体の抗原結合部位と一般的に会合する特有の決定基を認識する抗体である。抗Idを調製しようとするモノクローナル抗体で、モノクローナル抗体の供給源と同一の種および遺伝子型（例えばマウス系統）の動物を免疫することによって、抗Id抗体を調製する。免疫動物は、免疫する抗体のイディオタイプ決定基を

10

20

30

40

50

認識し、そしてこれらのアイソタイプ決定基に対する抗体（抗 I d 抗体）を産生することによって応答するであろう。

【0138】

したがって、本発明のポリペプチドに対して生成されるモノクローナル抗体を用いて、適切な動物において、抗 I d 抗体を誘導することが可能である。こうした免疫マウス由来の脾臓細胞を用いて、抗 I d モノクローナル抗体を分泌する抗 I d ハイブリドーマを産生することが可能である。さらに、抗 I d 抗体を、キーホールリンペット（keyhole limpet）ヘモシアニン（KLH）などのキャリアーとカップリングし、そしてさらなる BALB/c マウスを免疫するのに使用可能である。これらのマウス由来の血清は、R-PTPアーゼエピトープに特異的な最終 mAb の結合特性を有する、抗-抗 I d 抗体を含有するであろう。したがって、抗 I d 抗体は、化膿性連鎖球菌ポリペプチドなどの、評価しようとするエピトープに構造的に類似のイディオタイプエピトープまたは「イディオトープ」を有する。 10

【0139】

用語「抗体」はまた、損なわれていない（intact）分子と共に抗原に結合可能な Fab などの断片両方を含むことを意味する。Fab 断片は、損なわれていない抗体の Fc 断片を欠き、循環からより迅速に一掃され、そして損なわれていない抗体より少ない非特異的組織結合を有する可能性がある（Wahlら，1983，J. Nucl. Med. 24:316-325）。本発明に有用な抗体の Fab および他の断片は、損なわれていない分子のための方法にしたがって、化膿性連鎖球菌ポリペプチドの検出および定量化に使用可能であると認識されるであろう。 20

【0140】

抗 I d 抗体はまた、さらに別の動物において、免疫応答を誘導して、いわゆる抗-抗 I d 抗体を産生する「免疫原」としても使用可能である。抗-抗 I d は、抗 I d を誘導した元来の mAb とエピトープ的に同一である可能性がある。したがって、mAb のイディオタイプ決定基に対する抗体を用いることによって、同一の特異性の抗体を発現する他のクローンを同定することが可能である。

【0141】

抗体は多様な方式で用いられ、例えばタンパク質が発現されることの確認に、またはタンパク質が発現される場所の確認に用いられる。例えば、標識抗体（例えば FACS 用に蛍光標識したもの）を損なわれていない細菌と共にインキュベーションすることが可能であり、そして細菌表面上の標識の存在がタンパク質の位置を確認する。 30

【0142】

日常的なプロトコルを用いて、ポリペプチドまたはエピトープ所持断片、類似体、または細胞を動物に投与することによって、本発明のポリペプチドに対して生成する抗体を得ることが可能である。モノクローナル抗体を調製するため、連続細胞株培養から生じる抗体を提供する技術いずれかを用いる。

【0143】

免疫原組成物

やはり提供するのはい免疫原組成物である。本発明の免疫原組成物は、ヒト（好ましくは）および非ヒト動物などの哺乳動物において、連鎖球菌感染の治療に、使用可能である。例えば、動物は、ウシ、イヌ、ウマ、ネコ、およびブタであることが可能である。配列番号 415（ORF 1021）は、S. エクイ（S. equi）にも見られるタンパク質に対応することが注目される。したがって、この配列は、ウマ感染を治療する免疫原組成物において、それと共に他の動物またはヒトにおいても使用可能である。特定の適用には、限定されるわけではないが、ウマ科の鼻咽頭および排出リンパ節の非常に伝染性の疾患である腺疫の治療、並びに、ウシ、ウマ、およびブタの呼吸器感染および乳腺炎の治療が含まれる。 40

【0144】

本発明の免疫原組成物は、予防的（すなわち感染を防御するか、または感染開始を減少さ 50

せる)または療法的(すなわち感染後、感染によって引き起こされる疾患または副作用を治療する)かいずれであることも可能である。

【0145】

免疫原組成物は、本発明のポリペプチドを含んでなることが可能である。そうするために、1以上のポリペプチドを適切な濃度に調整し、そして適切なアジュバント、希釈剤、キャリアーいずれか、またはその組み合わせいずれかと配合することが可能である。生理学的に許容しうる培地をキャリアーおよび/または希釈剤として使用可能である。これらには、限定されるわけではないが、水、適切な等張培地、グリセロール、エタノールおよび他の慣用的溶媒、リン酸緩衝生理食塩水などが含まれる。

【0146】

本明細書において、「アジュバント」は、ポリペプチドであれポリヌクレオチドであれ、抗原の免疫原性を増進させるよう作用する物質である。したがって、アジュバントはしばしば、免疫応答を促進するために与えられ、そして当業者に公知である。適切なアジュバントには、限定されるわけではないが、リン酸アルミニウムおよび水酸化アルミニウムなどのアルミニウム塩(ミョウバン)、結核菌(*Mycobacterium tuberculosis*)、百日咳菌(*Bordetella pertussis*)、細菌リポ多糖、アミノアルキルグルコサミンリン酸化合物(AGP)、あるいはその誘導体または類似体が含まれ、これらはCorixa(モンタナ州ハミルトン)から入手可能であり、そして本明細書に援用される米国特許第6,113,918号に記載される。1つのこうしたAGPは、529(以前はRC529として知られた)としても知られる、2-エチル 2-デオキシ-4-O-ホスホノ-3-O-2-b-D-グルコピラノシドである。この529アジュバントは水性型として、または安定なエマルジョンとして配合される。他のアジュバントは、米国特許第4,912,094号に記載されるMPL(登録商標)(3-O-デアシル化モノホスホリル脂質A)(Corixa)、CpGモチーフを含有するオリゴヌクレオチドなどの合成ポリヌクレオチド(米国特許第6,207,646号)、米国特許第5,057,540号に記載されるQuil AまたはSTIMULON(登録商標)QS-21(Antigenics、マサチューセッツ州フラミンガム)などのサポニン類、百日咳毒素(PT)、または大腸菌熱不安定性毒素(LT)、特にLT-K63、LT-R72、CT-S109、PT-K9/G129;例えば国際特許公開番号WO 93/13302号およびWO 92/19265号を参照されたい、コレラ毒素(野生型または突然変異型いずれか、例えば公開国際特許出願WO 00/18434号にしたがって、アミノ酸29位のグルタミン酸が別のアミノ酸、好ましくはヒスチジンに置換されたものである)。

【0147】

多様なサイトカインおよびリンホカインがアジュバントとして使用するのに適している。1つのこうしたアジュバントは、本明細書に援用される、米国特許第5,078,996号に記載されるようなヌクレオチド配列を有する、顆粒球・マクロファージコロニー刺激因子(GM-CSF)である。GM-CSF cDNAを含有するプラスミドが大腸菌に形質転換され、そして寄託番号第39900号で、American Type Culture Collection(ATCC), 10801 University Boulevard, Manassas, VA 20110-2209に寄託されている。サイトカイン、インターロイキン-12(IL-12)は、本明細書に援用される米国特許第5,723,127号に記載される、別のアジュバントである。限定されるわけではないが、インターロイキン1アルファ、1ベータ、2、4、5、6、7、8、10、13、14、15、16、17および18、インターフェロナーアルファ、ベータおよびガンマ、顆粒球コロニー刺激因子、並びに腫瘍壊死因子アルファおよびベータを含む、他のサイトカインまたはリンホカインが免疫変調活性を有することが示されてきており、そしてアジュバントとして使用するのに適している。

【0148】

ポリペプチドにはまた、場合によって、ペプチド、ポリペプチド、またはタンパク質、あ

10

20

30

40

50

るいは多糖にコンジュゲート化するかまたは連結した、ポリペプチドの少なくとも部分が含まれることも可能である。

【0149】

本発明の免疫原組成物には、さらに、本明細書に援用される米国特許第4,673,574号、第4,902,506号、第5,097,020号、および第5,360,897号（ロチェスター大学に譲渡）に開示されるような、免疫原コンジュゲートが含まれることが可能である。これらの特許は、還元末端を有し、そして細菌病原体の細菌莢膜ポリマー、および細菌毒素またはトキシイド由来の免疫原莢膜ポリマー断片の還元アミノ化産物である、免疫原コンジュゲートを解説する。本発明にはまた、ヒトにおいて、有効レベルの抗莢膜ポリマー抗体を誘発するこれらのコンジュゲートを含有する免疫原組成物も含まれる。 10

【0150】

本発明のポリペプチドを2以上含むことによって、またそれと共に、限定されるわけではないが、C5aペプチダーゼ、Mタンパク質、アドヘシン等を含む、1以上の既知の化膿性連鎖球菌ポリペプチドと1以上の本発明のポリペプチドを組み合わせることによって、免疫原組成物の組み合わせを提供する。

【0151】

本発明の免疫原組成物はまた、遺伝子発現を調節する制御配列と機能可能であるように関連する、本発明のポリヌクレオチド配列も含んでなる。DNAの発現を促進するであろう制御要素、すなわちプロモーターおよび／またはエンハンサー要素の調節下、プラスミドなどの発現ベクター中に、目的のポリヌクレオチド配列を設計する。好ましい態様において、ヒトサイトメガロウイルス極初期プロモーター／エンハンサーを用いる（米国特許第5,168,062号）。プロモーターは細胞特異的であり、そしてあらかじめ決定された細胞でのみ、ポリヌクレオチドの実質的な転写を可能にすることが可能である。 20

【0152】

ポリヌクレオチドは、「裸の」DNA（米国特許第5,580,859号）として、宿主に直接導入されるか、またはブピバカインおよび他の局所麻酔薬（米国特許第5,593,972号）およびカチオン性ポリアミン（米国特許第6,127,170号）などの、免疫を促進する剤と複合して配合される。

【0153】

このポリヌクレオチド免疫法において、本発明のポリペプチドは、*in vivo*で一過性の方式で発現され；遺伝子材料は宿主染色体にまったく挿入されないか、または組み込まれない。この方法は、その目的が染色体に目的の遺伝子材料を挿入するかまたは組み込むことである、遺伝子治療とは区別すべきである。アッセイを用いて、免疫によって投与されるポリヌクレオチドが宿主において形質転換表現型を生じさせないことを確認する（米国特許第6,168,918号）。 30

【0154】

ひとたび配合したら、本発明の免疫原組成物は、被験者に直接投与するか、組換えタンパク質発現のため、被験者由来の細胞に *ex vivo* または *in vitro* で搬送することが可能である。被験者に直接搬送するため、投与は、鼻内、非経口、経口、腹腔内、静脈内、皮下、あるいは鼻内、口、目、肺、膣、または直腸表面などの粘膜表面いずれかに、エアロゾルスプレーによるなどで、局所適用するなど、慣用的な形式いずれによることも可能である。 40

【0155】

被験者は哺乳動物または鳥類であることが可能である。被験者はまた、ヒトであることも可能である。免疫学的に有効な量の免疫原組成物を適切な用量数で被験者に投与して、免疫応答を引き出す。免疫学的に有効な量は、本明細書において、少なくとも、処置した個体の免疫系が細菌感染の臨床的影響を減少させる応答を生成するようにするのに十分な、単回用量、または一連の用量の一部としていずれかでの、哺乳動物宿主（好ましくはヒト）へのその量の投与を意味する。防御は、免疫原組成物の単回用量によって与えられるこ 50

とも可能であるし、または防御を維持するため、後の追加免疫用量に加えて、いくつかの用量の投与が必要である可能性もある。これは、細菌負荷の最小限の減少から感染の防御までの範囲にあることが可能である。理想的には、処置個体は、 β 溶血性連鎖球菌感染のより深刻な臨床的徴候を示さないであろう。投薬量は、年齢および体重など、個体の特定の条件に応じて変化する可能性がある。この量は、当業者に知られる手段によって、日常的試験で決定可能である。

【0156】

多様な試験を用いて、本発明のポリペプチドの *in vitro* 免疫原性を評価する。例えば、ポリペプチドは、組換え的に発現するかまたは化学的に合成して、そして免疫プロットによって、被験者血清をスクリーニングするのに使用可能である。被験者および被験者血清間の陽性反応は、被験者が以前、問題のポリペプチドに免疫応答を生じ、すなわちポリペプチドが免疫原であることを示す。この方法はまた、免疫優性ポリペプチドを同定するのに使用可能である。

10

【0157】

ELISA アッセイはまた、*in vitro* 免疫原性を評価するのに用い、ここで、目的のポリペプチド抗原を96ウェルプレートなどのプレート上にコーティングし、そしてワクチン接種したかまたは天然に曝露されたかいずれかの動物（例えばヒト）由来の試験血清を、コーティングした抗原と反応させる。試験ポリペプチド抗原に特異的な抗体いずれかが存在する場合、当業者に知られる標準法によって、検出することが可能である。

【0158】

あるいは、同じ血清を、全化膿性連鎖球菌細胞と反応させることが可能である。その後、コロイド金コンジュゲート化抗体を用いて、血清に存在する反応性抗体を検出し、そしてLV-SEMによって視覚化することが可能である。

20

【0159】

2つの動物攻撃 (*challenge*) アッセイモデルを用いて、ワクチン抗原の有効性を試験することが可能である。第一のものは粘膜免疫を検討する。確立された方法にしたがって、ワクチン候補で、非経口的にまたは粘膜的に、マウスを能動的に免疫する。その後、鼻内投与によって、野生型化膿性連鎖球菌でマウスを攻撃する。その後、マウスの鼻腔／咽頭腔における化膿性連鎖球菌持続性を測定可能である。有効性は、動物の喉からの細菌のクリアランス増進によって反映される。

30

【0160】

あるいは、能動的非経口免疫に続いて、化膿性連鎖球菌細胞の皮下注射によって、全身感染に対する防御を評価することが可能である。有効性は、死亡減少および／または注射部位の組織病理の減少によって測定される。

【0161】

試料中の検出

やはり提供されるのは、生物学的試料において、 β 溶血性連鎖球菌および化膿性連鎖球菌を検出し、そして同定する方法である。1つの態様において、該方法は、(a) 本発明のポリヌクレオチドと生物学的試料を、相補的塩基対のハイブリダイゼーションを可能にする条件下で接触させ、そして(b) 試料中のハイブリダイゼーション複合体の存在を検出する工程を含んでなる。別の態様において、該方法は、(a) 本発明の抗体と生物学的試料を、免疫複合体の形成に適した条件下で接触させ、そして(b) 試料中の免疫複合体の存在を検出する工程を含んでなる。さらに別の態様において、該方法は、(a) 本発明のポリペプチドと生物学的試料を、免疫複合体の形成に適した条件下で接触させ、そして(b) 試料中の免疫複合体の存在を検出する工程を含んでなる。

40

【0162】

本発明の抗原、またはその抗原性断片を免疫アッセイで用いて抗体レベルを検出するか、または逆に、抗化膿性連鎖球菌抗体を用いて抗原レベルを検出する。よく定義された組換え抗原に基づく免疫アッセイを開発して、侵襲性診断法の代わりに使用することが可能である。例えば血液または血清試料を含む生物学的試料内で、本発明のポリペプチドに対す

50

る抗体を検出可能である。免疫アッセイのためのプロトコルは、例えば、競合、または直接反応、またはサンドイッチ型アッセイに基づくことが可能である。プロトコルはまた、例えば固体支持体を用いることも可能であるし、または免疫沈降によることも可能である。本発明のポリペプチドはまた、受容体ーリガンド研究にも有用である可能性がある。

【0163】

以下の実施例は例示であり、そして本発明はこれに限定されることを意図されない。

【実施例1】

【0164】

実施例1

細菌、培地、および試薬

適切な抗生物質を含有するSOB（0.5%酵母エキス、2.0% Tryp、10mM塩化ナトリウム、2.5mM塩化カリウム、10mM塩化マグネシウム、10mM硫酸マグネシウム）中で、大腸菌を培養し、そして維持した。アンピシリンは100μg/mlの濃度で、クロラムフェニコールは30μg/mlの濃度で、そしてカナマイシンは50μg/mlの濃度で用いた。30g/l トッド・ヒューイット、5g/l 酵母エキス（THY）ブロス中で、化膿性連鎖球菌株SF370（ATCC寄託番号700294）を培養した。

【0165】

バイオインフォマティクス／遺伝子マイニング

化膿性連鎖球菌M1株のゲノム未注釈配列をオクラホマ大学のウェブサイトからダウンロードし、そして解析してオープンリーディングフレーム（ORF）を同定した。このゲノム配列は、GenBankに提出され、そして寄託番号AE004092を割り当てられたと報告され、そして株M1 GASは、ATCCに提出され、そして寄託番号ATCC 700294を与えられたと報告された。

【0166】

ORFは、3つの潜在的な開始部位コドン、ATG、GTG、またはTTGのいずれか1つ、および3つの潜在的な停止コドン、TAA、TAG、TGAのいずれか1つを有すると定義された。3つのORFファインダーアルゴリズム：GLIMMER（59）；GeneMark（34）；および発明者の譲受人に開発された第三のアルゴリズムの特有のセットを用いて、すべてのORFを決定する有効性を増進した。

【0167】

決定したORFの正確さを評価するため、離散コサイン変換（DiCTion）として当該技術分野に知られる、離散数学的コサイン関数を使用して、各ORFにスコアを割り当てた。DiCTionスコア>1.5のORFは、タンパク質産物をコードする高い確率を有するとみなす。3つのORF発見アルゴリズムに予測されるORFの最低の長さを、74アミノ酸のタンパク質をコードするであろう225ヌクレオチド（停止コドンを含む）に設定した。

【0168】

ORFの残りの最終検索として、tBLASTnを用いて、公的タンパク質データベース（以下に記載）に対して、>75ヌクレオチドの非コード領域すべてを検索した。これは、フレームシフトを含有する遺伝子領域（42）または抗原変異を引き起こすのに役割を有する可能性がある遺伝子断片（21）を同定するのを補助した。ここで見出した残りのORFをいずれも化膿性連鎖球菌のORFデータベースに加えた。発明者らのグループ内部の図式的解析プログラムを用いて、ゲノム配列に比較して、予測されるORFの6つの読み枠および位置をすべて示した。これは、他のORFと長い重複を有するORFを除去するのを補助したが、他のORFに完全に埋まるORFの既知の例がある（25、33）。

【0169】

BLAST v2.0ギャップ化検索アルゴリズム、BLASTpを用いて、この化膿性連鎖球菌ORFの最初の注釈付けを行って、相同配列を同定した。 $< e^{-10}$ のいずれか

10

20

30

40

50

のカットオフ「e」値を有意とみなした。FASTAおよびPSI-BLASTを含む他の検索アルゴリズムもまた用いた。相同検索に用いた非重複タンパク質配列データベースは、毎日アップデートされる、GenBank、SWISS-PROT、PIR、およびTREMBLデータベース配列からなった。 $>e^{-10}$ のBLASTp結果を持つORFを、化膿性連鎖球菌に特有であるとみなした。

【0170】

既知のまたは推測されるワクチン標的遺伝子と共に、タンパク質の位置または機能を同定するワードを用いて、全Blast結果のキーワード検索を行った。さらに、最初のBlast結果に関連するMEDLINE参考文献すべてのキーワード検索を行って、ORFに関するさらなる情報を採した。

10

【0171】

DNA解析のため、各遺伝子内の%G+C含量を同定した。ORFの%G+C含量は、ORF内のコドンすべての第三のヌクレオチド位の(G+C)含量として計算した。報告される値は、該生物で見られるすべてのORFに関して得られるこうした値の算術平均からのこの値の相違であった。これらのORFは、多くの他の病原性アイランドと調和するパターン(22)の、ピロリ菌(*H. pylori*)由来のcag病原性アイランドの場合に見られているような(2)、水平移行から生じる可能性があるため、さらなる解析には8以上の絶対値が重要とみなされた。

【0172】

いくつかのパラメーターを用いて、予測されるタンパク質の分配を決定した。細胞膜を渡る転位置を運命付けるタンパク質は、N末端で陽性荷電残基が隣接する中央疎水性領域で構成されるリーダーシグナル(シグナル配列としても知られる)をコードする(56)。プログラムSignalPを用いて、シグナルペプチドおよびその切断部位を同定した(46)。細菌中のタンパク質局在を予測するため、ソフトウェアPSORTを用いた(44)。このプログラムは神経網アルゴリズムを用いて、グラム陽性細菌では、細胞質、周辺質、および細胞膜、それと共にグラム陰性細菌では外膜へのタンパク質の局在を予測する。ソフトウェアプログラムTopPred2(10)を用いて、タンパク質の膜貫通(TM)ドメインを解析した。このプログラムは、潜在的に膜の脂質二重層に渡る可能性がある、疎水性であるタンパク質領域を予測する。外膜タンパク質は、典型的には、 α らせんTMドメインを持たない。

20

30

【0173】

タンパク質ドメインまたは保存されるタンパク質領域の多数並列の隠れマルコフモデル(HMM) Pfamデータベース(61)を用いて、現存するタンパク質ファミリーに属する可能性がある化膿性連鎖球菌タンパク質を同定した。このアウトプットのキーワード検索を用いて、Blast検索規準では見逃した可能性がある、表面に局在する化膿性連鎖球菌タンパク質を同定するのを補助した。本発明者の譲受人もまた、HMMモデルを開発した。コンピュータアルゴリズム、HMM Lipopを開発して、132の生物学的に性質決定された、30を超える生物由来の非化膿性連鎖球菌細菌リポタンパク質を用いて、リポタンパク質を予測した。このトレーニングセットは、実験的に立証された原核生物リポタンパク質から生成された。タンパク質の開始からシステインアミノ酸までに、次の2つのさらなるアミノ酸を加えたタンパク質配列を用いて、HMMを生成した。LPXTG細胞壁局所化シグナルを含有する、約70の既知の原核生物タンパク質を用いて、HMM(15)を開発して、ペプチドグリカン層に係留される細胞壁タンパク質を予測した(38、45)。このモデルはLPXTG配列を用いるだけでなく、下流配列の2つの特徴、疎水性膜貫通ドメインおよび陽性荷電カルボキシ末端も含んだ。非共有的に、ペプチドグリカン層と相互作用し、そして上述のLPXTGタンパク質種と異なるいくつかのタンパク質もある。これらのタンパク質は、カルボキシ末端にコンセンサス配列を有するようである(32)。この領域のHMMを開発し、そしてこの種に属する化膿性連鎖球菌タンパク質を同定した。

40

【0174】

50

また、化膿性連鎖球菌の同定されたORFにコードされるタンパク質を他の特性に関しても評価した。タンデム反復ファインダー（5）は、MSCRAMM（20）および髄膜炎菌の相可変表面タンパク質（51）に見出されるものなどの反復DNA配列を含有するORFを同定した。Arg-Gly-Asp（RGD）接着モチーフを含有するタンパク質は、その受容体として作用するインテグリンと共に、細胞接着の主要認識系を構成する。RGD認識は、微生物が真核生物組織への進入を獲得するのに用いる1つの機構である（29、63）。しかし、すべてのRGD含有タンパク質が細胞接着を仲介するわけではない。カルボキシ端にプロリンを持つRGD含有ペプチド（RGDP）が細胞接着アッセイで不活性であり（52）、そしてしたがって、排除されることが示されている。Geantfammerソフトウェアを用いて、相同ファミリーにタンパク質をクラスター形成した（50）。ファミリー種の予備的解析は、ワクチン候補クラスター内の新規ORFを提供すると共に、潜在的タンパク質機能を定義した。

10

【0175】

化膿性連鎖球菌のトリプシン消化

化膿性連鎖球菌のスターター培養をTHY中、37℃、5%CO₂中、または大気O₂中で一晚増殖させた。その後、200ml新鮮THY中で各スターター培養を1:25に希釈し、そしてそれぞれ、CO₂または大気O₂中いずれかで、1~1.3のOD₄₉₀に増殖させた。その後、4,000xg、15分間の遠心分離によって、細胞を採取し、そして10mlの20mM Tris、pH8.0、150mM NaCl緩衝液で3回洗浄した。最後の洗浄後、0.8Mスクロースを含有する、2mlの同一緩衝液に各ペレットを再懸濁し、そして2つの試験管に等しく分配した。各増殖条件の一方の試験管に、40μgのトリプシンを添加し；もう一方の試験管を陰性消化対照として用いた。細胞懸濁物を37℃で4時間震盪した。生存細胞計数および低電圧走査型電子顕微鏡観察（LV-SEM）での視覚化用に各懸濁物の試料を採取した。その後、懸濁物を遠心分離し、そして上清を収集し、そして低タンパク質結合2μMフィルターを通じてろ過した。

20

【0176】

マイクロキャピラリーHPLCインターフェース

自動化マイクロエレクトロスプレー逆相HPLC上で、ペプチド抽出物を解析した。マイクロエレクトロスプレーインターフェースは、長さ10cmまで10μm C18逆相ビーズ（YMC、ノースカロライナ州ウィルミントン）を充填したPico frit融合シリカスプレー針、長さ50cm、75μm ID、8μm開口部直径（New Objective、マサチューセッツ州ケンブリッジ）からなった。質量分析計検出計の前面に配置したベース上に保持した光ファイバーホルダー（Melles Griot、カリフォルニア州アービン）に、Pico frit針を乗せた。チタンユニオンを通じてカラムの後方を垂直にして、エレクトロスプレーインターフェースの電気的連結を供給した。ユニオンをある長さの融合シリカキャピラリー（FSC）チュービングでFAMOS自動試料採取装置（LC-Packings、カリフォルニア州サンフランシスコ）に連結し、この装置をHPLC溶媒ポンプ（ABI 140C、Perkin-Elmer、コネチカット州ノーウォーク）に連結した。HPLC溶媒ポンプは、50μl/分の流量を搬送し、PEEKマイクロタイト分割T（Upchurch Scientific、ワシントン州オークハーバー）を用いて、これを250nl/分に減少し、そしてその後、FSCトランスファーラインを用いて、自動試料採取装置に搬送した。内部ユーザープログラムを用いて、LCポンプおよび自動試料採取装置を各々制御した。試料をプラスチック自動試料採集バイアルに入れ、密封し、そして5μl試料ループを用いて注入した。

30

40

【0177】

マイクロキャピラリーHPLC-質量分析

Savant Speed Vac濃縮装置（ThermoQuest、ニューヨーク州ホールドブルック）を用いて、表面消化物由来の抽出ペプチドを10倍に濃縮し、そしてその後、0~50%溶媒B（A:0.1M HOAc、B:90% MeCN/0.1M HOAc）の50分間勾配を用いて、マイクロエレクトロスプレーHPLC系によって

50

分離した。スプレー電圧1.5 kVで操作し、そして加熱キャピラリー温度125℃を用いて、Finnigan LCQ-DECAイオントラップ質量分析計(ThermoQuest、カリフォルニア州サンノゼ)上で、ペプチド解析を行った。装置と共に提供されるデータ獲得ソフトウェアを用いて、自動化MS/MSモードで、データを獲得した。獲得法には、1回のMSスキャン(375~600 m/z)、その後、そのMSスキャン中の2つの最も豊富なイオンのMS/MSスキャンが含まれた。その後、装置は、第二のMSスキャン(600~1000 m/z)、その後、そのスキャン中の2つの最も豊富なイオンのMS/MSスキャンを行った。動的排除および同位体排除機能を使用して、解析するペプチドイオン数を増加させた(設定: 3 amu = 排除幅、3分間 = 排除期間、30秒間 = 排除前期間、3 amu = 同位体排除幅)。

10

【0178】

データ解析

化膿性連鎖球菌の完全ゲノム由来のタンパク質データベースを用い、Finnigan Bioworksデータ解析パッケージ(ThermoQuest、カリフォルニア州サンノゼ)に取り込んだSEQUESTコンピュータアルゴリズム(17)を用いて、MS/MSデータの自動解析を行った。

【0179】

クローニングおよびタンパク質発現

予測される成熟タンパク質の開始部分に順方向5'プライマーがアニーリングするであろうように、望ましいORFのPCR増幅用に、プライマーセットを設計した。リポタンパク質に関しては、ジスルフィド架橋を最小限にするため、成熟タンパク質のシステイン残基をコードするコドンの直後にアニーリングするように、5'順方向プライマーを設計した。相対する逆方向3'プライマーの設計は、予測されるタンパク質の種類に応じた。LPXTGを含有するタンパク質に関しては、細胞壁アンカー領域の開始部分(5'端)にアニーリングするように、プライマーを設計した。すべての他の予測されるタンパク質に関しては、これらがORFの3'端にアニーリングするように設計した。さらに、5'順方向プライマーをまず、相対する3'逆方向プライマーと共に、チオレドキシンへのインフレーム融合を可能にするように設計して、下流hisパッチおよびV5エピトープを含むようなリードスルーを可能にする(pBAD/チオ-TOPPO(登録商標)、Invitrogen、カリフォルニア州カールスバッド)。pBADベクターは、アラビノース誘導性プロモーターを用いる。並行して、これらの同じPCR産物をまた、pCRT7 TOPPO(登録商標)(Invitrogen、カリフォルニア州カールスバッド)にもクローニングした。これは、精製のためのXpressエピトープおよびhisタグへのN末端融合を可能にした。

20

30

【0180】

すべてのPCR反応は、化膿性連鎖球菌M1株、SF370(ATCC寄託番号700294)をテンプレートとして用いた。PCR産物を大腸菌宿主TOP10に形質転換し、そして100 µg/mlアンピシリンを含有するSOB上に蒔いた。ベクターに特異的な5'プライマーおよび遺伝子挿入物にアニーリングする特異的な3'逆方向プライマーを用いたPCR増幅によって、コロニーをスクリーニングした。50 µlの50%グリセロールを含有する96ウェルマイクロタイタープレートのウェルにコロニーを植え付けた。遺伝子あたり10~12コロニーをプレートの一行に植え付けた。第二の96ウェルPCRプレートでは、目的の遺伝子に特異的な50 µl反応をセットアップした。グリセロールに懸濁した細胞1 µlをPCR反応のテンプレートとして用いた。期待される大きさのバンドを生じる反応を、さらに解析した。50%グリセロール中に植え付けた細胞は添加されたSOB培地を有し、そしてこれを37℃で5~8時間インキュベーションして、そして-70℃で凍結した。

40

【0181】

一晚増殖させるため、PCR陽性コロニーを2 ml培養に接種した。培養の一部を用いて、プラスミドDNAを調製し、これを制限消化によって解析して挿入物を確認する一方、

50

別の一部を用いて、発現のため、10 ml 発現培養 (pBAD プラスミド用) に植え付けた。0.5% L-アラビノースを用いて、対数中期培養を2時間誘導した。スクリーニング前に、T7/NT プラスミドを発現株 BLR (DE3) pLysS に形質転換した。1 mM IPTG の添加によって T7/NT 培養を誘導し、そして2時間インキュベーションした。誘導培養の全細胞溶解物を、SDS-PAGE 上、2つ組で泳動した。一方のゲルはクーマシーで染色し、そしてもう一方はニトロセルロースに移し、そして相当するエピトープタグに対する抗体で探査した (probed)。

【0182】

1~2 l 体積中で陽性クローンを増殖させ、そして大規模精製のため、誘導した。細胞の凍結融解、その後、DNアーゼ/RNアーゼ消化および RC5B 冷蔵遠心分離装置 (sorbol (登録商標)、Dupont、デラウェア州ウィルミントン) 中、9,000 x g、15 分間の遠心分離によって、組換えタンパク質の可溶性および発現レベルを評価した。不溶性材料から可溶性分画を取り除き、そして両方を分離し、そして SDS-PAGE によってタンパク質局在および発現に関して評価した。Ni-NTA (Qiagen Inc.、カリフォルニア州バレンシア) 樹脂上に溶解細胞の可溶性分画を通過させ、そしてイミダゾールで結合タンパク質を溶出することによって、可溶性融合タンパク質を精製した。PD-10 カラム (Amersham Pharmacia Biotech、ニュージャージー州ピスカタウェイ) 上で、溶出タンパク質を緩衝液交換した。

10

【0183】

不溶性組換えタンパク質を洗浄し、そして PBS、0.1% TRITON-X100 中で3回遠心分離した。その後、封入体を PBS 4M 尿素中で可溶化し、そして PD-10 カラム (Amersham Pharmacia、ニュージャージー州ピスカタウェイ) を通して、PBS、0.01% TRITON-X100、0.5M NaCl に緩衝液交換した。ローリーアッセイによってタンパク質を定量化し、そして SDS-PAGE によって純度および濃度に関してチェックした。

20

【0184】

ポリクローナル抗血清の生成

第0週、第3週、および第5週に、上述のように調製した 5 µg 精製タンパク質、100 µg AlPO₄、および 50 µg MPL (登録商標) でスイス・ウェブスター・マウス (群あたり5匹) を免疫し、そしてその後、第8週に採血した。

30

【0185】

化膿性連鎖球菌のイムノゴールド標識および LV-SEM

先に記載されるように、細菌細胞を標識した (49)。簡潔には、対数後期の細菌培養を2回洗浄し、そして 10 mM リン酸緩衝生理食塩水 (PBS) (pH 7.4) に 1×10^8 細胞/ml の濃度に再懸濁し、そしてポリ-L-リジンでコーティングしたカバーガラス上に置いた。穏やかに洗浄してカバーガラスから過剰な細菌を除き、そして未標識試料を固定液 (7.5% スクロースを含有する 0.1 M カコジル酸ナトリウム緩衝液中の、2.0% グルタルアルデヒド) に30分間入れた。コロイド金で標識しようとする細菌を、0.5% ウシ血清アルブミンを含有する PBS で洗浄し、そして上述のように調製した免疫前または過免疫 (hyper-immune) マウスポリクローナル抗体を室温で1時間適用した。その後、細菌を穏やかに洗浄し、そして 18 nm コロイド金粒子 (Jackson ImmunoResearch Laboratories, Inc.、ペンシルバニア州ウェストグローブ) にコンジュゲート化したヤギ抗マウス抗体の 1:6 希釈を室温で10分間適用した。最後に、すべての試料を PBS で穏やかに洗浄し、そして上述の固定液に入れた。0.1 M カコジル酸ナトリウム緩衝液で、2回、10分間洗浄して、固定液を試料から除き、そして 1% 四酸化オスミウムを含有する 0.1 M カコジル酸ナトリウム中、30分間、後固定した。その後、0.1 M カコジル酸ナトリウムで試料を2回洗浄し、エタノールで脱水し、Samdri-780A (Tousimis、メリーランド州ロックビル) を用いて、アンダーソンの CO₂ 法によって臨界点乾燥して、そして白金の 1~2 nm 断続層でコーティングした。慣用的なトポグラフィック画像化のための第

40

50

二の電子検出装置および原子数コントラストによってコロイド金の視覚化を増進する高分解能ロビンソン後方散乱検出装置を用いて、低加速電圧（1～4.5 keV）で操作するLEO 1550電界放射走査型電子顕微鏡を用いて、化膿性連鎖球菌細胞を視覚化した。

【実施例2】

【0186】

実施例2

免疫および攻撃

マウスの非経口免疫

第0週、第4週および第6週に、用量あたり、50 μg MPL（登録商標）（Corixa、モンタナ州ハミルトン）および100 μg AlPO₄と混合した目的のタンパク質5 μgを、生理食塩水で最終容量200 μlにし、そしてその後、マウスに皮下注射（s.c.）して、6週齢のメスCD1（Charles River Breeding Laboratories, Inc.、マサチューセッツ州ウィルミントン）またはスイス・ウェブスター（Taconic Farms Inc.、ニューヨーク州ジャーマンタウン）マウスを免疫する。対照マウスに、同じアジュバントと混合した5 μg破傷風トキソイドを注射する。最後の追加免疫7日後、すべてのマウスから採血し；その後、血清を単離し、そして-20℃で保存する。

【0187】

マウス鼻内攻撃モデル

最後の免疫10日後、20%正常ウサギ血清を含有するトッド・ヒューイット／酵母ブロス中で増殖させ、そして10 mlのPBSに再懸濁した、攻撃化膿性連鎖球菌株（ $1 \times 10^8 \sim 9 \times 10^8$ コロニー形成単位（CFU））の16時間培養を、25 gのメスCD1（Charles River Breeding Laboratories, Inc.、マサチューセッツ州ウィルミントン）またはスイス・ウェブスター（Taconic Farms Inc.、ニューヨーク州ジャーマンタウン）マウスに鼻内投与する。血液寒天プレート上に培養の希釈を蒔くことによって、生存数を決定する。

【0188】

i.p.注射によって、1.2 mgのケタミンHCl（Fort Dodge Animal Health、アイオワ州フォートドッジ）で、各マウスを麻酔する。麻酔したマウスの鼻孔に細菌懸濁物を接種する（マウスあたり10 μl）。攻撃16時間後、マウスを屠殺し、鼻を取り除き、そして組織ホモジナイザー（Ultra-Turax T25、Janke & Kunkel Ika-Labortechnik、ドイツ・シュタウフェン）を用いて3 ml無菌生理食塩水中でホモジナイズする。ホモジネートを生理食塩水中で10倍連続希釈し、そして1 mlあたり200 mgのストレプトマイシンを含有する血液寒天プレート上に蒔く。37℃で一晩インキュベーションした後、プレート上のβ溶血性コロニーを計数する。正常フロラに生き残っている可能性があるβ溶血性細菌と区別するため、ストレプトマイシン耐性によって、すべての攻撃株に印を付ける。

【0189】

皮下マウス攻撃モデル

5週齢（20～30 g）の非近交系免疫適格無毛オスマウス（Crl:SKH1-hrBR株）（Charles River、マサチューセッツ州ウィルミントン）を皮下注射に用いる。人道的な安楽死後、組織試料を収集する。

【0190】

実施例1に記載するように増殖させた化膿性連鎖球菌細胞を採取し、そして無菌氷冷発熱物質不含リン酸緩衝生理食塩水（PBS）で1度洗浄する。600 nmでの光学密度（OD₆₀₀）を調整して、必要な接種材料を生じる。0.1 mlに含有される化膿性連鎖球菌（ 1×10^8 CFU）を、ツベルクリンシリンジを用いて、各動物の右のわき腹に皮下注射する。同体積のPBSで対照マウスを処理する。5%ヒツジ血液（Becton Dickinson、メリーランド州コッキースビル）を含有するトリプトース寒天プレー

ト上のコロニー計数によって、マウスあたりに接種したC F U数を各実験に関して検証する。攻撃後、マウスを21日間観察する。心臓穿刺によって、死んだ各動物から血液を収集し、そして血液寒天プレート上で培養する。

【0191】

組織収集および組織像

接種前に、乱数生成装置を用いて、動物を群に割り当て、そして血液試料を抜き取って、ベースライン血液学的データを確立する。接種24時間、48時間、および72時間後、血液および組織試料を収集する。血液および組織収集に用いる方法は、すべての時点で同一である。

【0192】

動物の後眼窩洞 (retro-orbital sinus) から血液試料を得て、そして種特異的ソフトウェアと共にTechnicon H* 1 (ニューヨーク州タリータウン) 血液学解析装置を用いて、完全血算解析を行う。膿瘍または注射部位周囲の広い周縁切除によって、皮膚試料を収集する。これらの試料には、常に、注射部位由来の組織および比較用の連続する肉眼的に正常な組織が含まれる。試料の解剖学的方向を保持するように注意を払う。組織試料はまた、心臓、肝臓、脾臓、および肺からも得る。

【0193】

塩化亜鉛を補充した10%中性緩衝ホルマリン (Antech, Ltd., ミシガン州バトルクリーク) 中ですべての組織を固定する。全肺にまず、ホルマリンを注入し、そしてその後、他の臓器と共に、浸すことによって固定する。試料を18~24時間ホルマリンに入れ、そしてその後、プロセッシング前に、70%エチルアルコールに移す。上昇勾配のエチルアルコール中で脱水し、キシレンで清浄化し、そしてパラフィン浸潤する、標準的組織学的方法を使用する。パラフィンブロックを回転ミクロトームでプロセッシングして、4 μm切片を得る。組織学的切片をヘマトキシリンおよびエオジンで染色して、そしてマウントする。選択した組織の切片を作成し、そして組織グラム染色で染色する。

【0194】

マウス測定

GAS接種直前にマウスの重量を測定する。接種12時間後、およびその後、最初の週は毎日、動物重量および膿瘍サイズを測定する。その後、全部で21日間、週間隔で動物を観察する。膿瘍の寸法をカリパスで測定する；長さ (L) および幅 (W) 値を用い、球形楕円体 (spherical ellipsoid) 用の等式を使用して、膿瘍体積 $[V = 4/3 \pi (L/2)^2 \times (W/2)]$ および面積 $[A = \pi (L/2) \times (W/2)]$ を計算した。

【実施例3】

【0195】

実施例3

「ウェット化学」による性質決定のため、まず、77のORFを選択した。これらの研究の側面には、以下が含まれた：1) 全細胞ELISAによって測定されるような、各精製タンパク質に対して生成した特異的マウスポリクローナル血清が細菌表面に反応する能力、2) LV-SEMによって測定されるような、これらの同じ血清が、対数期または静止期増殖中の細菌細胞表面に反応する能力、3) 化膿性連鎖球菌の株 (M血清型) と共に、C群およびG群を含む、他の種の連鎖球菌に渡る遺伝子の遺伝子保存、4) ドットブロットによって測定されるような、これらの株による特定のタンパク質の表現型的発現；5) 定量的PCR (qPCR) による転写レベルでの目的の遺伝子の発現、および6) in vitro オプソニン食作用アッセイにおいて、ヒト抗体がこれらのタンパク質に対してオプソニン性である能力。

【0196】

77のORFをクローニングして、大腸菌で発現させ、そして発現タンパク質のうち62を精製した。全細胞ELISAおよびLV-SEMによる解析が完了した特異的抗体を生成するため、これらの精製タンパク質をマウスに注射した。さらに、化膿性連鎖球菌株お

10

20

30

40

50

よび連鎖球菌種に渡る遺伝子保存に関して、24のORFを評価した；*in vitro* および *in vivo* のqPCRによって、転写レベルでの発現に関して、いくつかを評価した。最後に、化膿性連鎖球菌タンパク質に特異的なヒト抗体を精製し、そしてオプソニン食作用アッセイで評価した。

【0197】

全細胞酵素結合免疫吸着アッセイ (ELISA)

化膿性連鎖球菌株SF-370を用いて、0.5%酵母エキスを含有するトッド・ヒューイットブロス (THY) に接種し、そして37℃で一晩培養した。遠心分離によって細胞を採取し、そしてリン酸緩衝生理食塩水 (PBS) で2回洗浄した。0.2のOD₆₀₀まで、細菌をPBSに再懸濁し、そして96ウェルポリスチレンマイクロタイタープレートの各ウェルを100μlの細菌懸濁物でコーティングした。その後、プレートを室温で空気乾燥し、マイラープレートシーラーで密封し、そして逆さにして4℃で3ヶ月まで保管した。アッセイのための調製において、Tris緩衝生理食塩水 (TBS) / 0.1% Brij-35でプレートを3回洗浄し、100μl/ウェルのORF特異的抗血清を各ウェルに添加して、そして37℃で2時間インキュベーションした。その後、TBS / 0.1% Brij-35でプレートを3回洗浄し、100μl/ウェルの二次抗体コンジュゲートを各ウェルに添加して、そして室温で1時間インキュベーションした。最後に、PBSで3回洗浄した後、100μl/ウェルの基質を各ウェルに添加し、そして室温で60分間発色 (develop) させた。その後、50μl/ウェルの3N NaOHを添加することによって反応を停止した。ELISAプレーリーダーを用いて、吸光度値 (OD₄₀₅) を測定した。

10

20

【0198】

遺伝子保存のポリメラーゼ連鎖反応 (PCR) 解析

試験した細菌株には、化膿性連鎖球菌由来の10株、SF370 (M1)、90-226 (M1)、80-003 (M1)、CS210 (M2)、CS194 (M4)、83-112 (M5)、CS204 (OF+, M11, T11)、CS24 (M12)、95-0061 (M28)、CS101 (M49)、および第四のM1血清型SpeB+、2つのS. zooepidemicus (S. zooepidemicus) 株、CS258およびGB21、並びに3つのG群連鎖球菌株、CS241、CS140、およびCS242が含まれた。THY中、5mlの一晩培養を増殖させた。2.5mlの各培養を遠心分離し、そして480μlの50mM EDTA、120μlの10mg/mlリゾチームおよび2μlの2500単位/mlミュータノリシン (mutanolysin) に再懸濁した。試料を1時間、37℃でインキュベーションした。残りのゲノム精製のため、PromegaのWizardゲノムDNA精製キットにしたがった。全長遺伝子のためのプライマーセットを、そして次にqPCR (以下を参照されたい) のために設計したプライマーをアッセイに用いた。PCR周期条件は、以下のとおりである：94℃で1分間維持、94℃15秒間および58℃10分間の16周期、先の94℃15秒間および58℃10分間を各々15秒間延ばして12周期、72℃で10分間維持、そして最後に4℃維持。アガロースゲル中の移動度によって、PCR産物を検証した。適切な大きさの強いバンドを含有する増幅を、いずれも陽性結果とみなした。

30

40

【0199】

定量的PCR (qPCR)

上記細菌培養から、または感染させホモジナイズしたマウス組織からRNAを単離した。試料を2mlのRNA later (Ambion、米国テキサス州オースティン) に懸濁し、そしてドライアイス/エタノールを用いて迅速凍結し、そして使用するまで-70℃に保管した。全部で3回の凍結融解周期のため、試料を室温に融解して、そしてその後、上述の方法を用いて再び凍結した。100μlの10mg/mlリゾチームおよび10μlの2500単位/mlミュータノリシンいずれかで試料を処理して、そして37℃で1時間インキュベーションするか、または等体積の0.1mmガラスビーズと試料を混合して、そして4800rpmで1分間、ビーズビーターに入れて、細胞を溶解した。上清を

50

ビーズから回収し、そしてさらに400 μ lのRNA laterをビーズに添加して、そして上述のように混合した。ビーズから回収した上清または消化溶液を、等体積のRNA queous溶解／結合溶液(Ambion)と混合し、そして勢いよくボルテックスした。微量遠心分離装置の最高速度で、試料を2分間回転させて、いかなる残存組織もペレットにした。上清を等体積の64%エタノールと混合し、そして一度に700 μ lずつ、フィルターカートリッジを通過させた。RNA queousマニュアルに記載されるように、フィルターカートリッジを洗浄した。2 x 25 μ lの95℃溶出溶液を用いて試料を溶出した。DNA-free(Ambion)を用いて、2回の1.5 μ l DNアーゼ処理を、各々、37℃で1時間行って、いかなるゲノム混入も取り除いた。cDNAを生成するRETROscript(Ambion)プロトコルに記載されるように、熱変性を伴う40 μ l最終体積のRT反応において、20 μ lの精製RNAを用いた。試料を85℃で変性させ、そして42℃で1時間インキュベーションし、その後、92℃で10分間インキュベーションすることによって、逆転写した。

10

【0200】

Primer Expressソフトウェア(Applied Biosystems、米国カリフォルニア州フォスターシティ)を用いて設計した、各ORFに特異的なプライマーおよびプローブを用いて、定量的PCRを行った。2 x TaqmanユニバーサルPCRマスターミックス(Applied Biosystems)、300 nM順方向プライマー、300 nM逆方向プライマー、200 nM FAM/TAMRAプローブ、およびcDNAテンプレートを用いて、25 μ l反応をセットアップした。PCR反応は、以下のとおりであった：50℃2分間、95℃10分間、95℃15秒間および60℃1分間の40周期。リボソーム16S RNAを内部対照として用い、すべての結果は16S Ct値に対して規準化する。標準曲線からの結果に基づいて、これらのウェルに添加するcDNAを100倍に希釈して、目的のORFに類似のCt値を生じた。

20

【0201】

ヒト多核白血球(PMN)の精製

Percol gradientを用いて、4人のドナー由来のヒト全血のプールから、PMNを精製した。Percolをハンクス平衡化塩溶液(HBSS)で希釈することによって、3層勾配を調製した。それぞれ、再高密度相は2.7:1で、中間は1.079:1、そして上相は1.07:1のPercol:HBSSであった。10 ml体積の全血を勾配に重層し、そして20℃、2600 RPMで、20分間遠心分離した。上層を取り除き、グルコースを含むPBSで洗浄してPercolを取り除き、遠心分離し、そして滅菌水に再懸濁して、赤血球を溶解した。正常生理食塩水の20倍濃縮溶液を添加して平衡化し、再遠心分離して溶解細胞を取り除き、PMNを再懸濁して、そして計数した。カルシウムおよびマグネシウムを含有するPBSに細胞を希釈し、そして使用前に37℃にした。

30

【0202】

ヒト血清由来のORF特異的抗体のブロット解析

2 μ gのタンパク質をニトロセルロース上にコーティングし、そして15分間風乾させた。BLOTTO中、室温で30分間ブロットをインキュベーションし、そしてその後、5 mlのプールしたヒト血清血漿と4℃で16時間インキュベーションした。0.2% Tween 20を含むPBSでニトロセルロースをリンスし、そしてアルカリホスファターゼにコンジュゲート化したヤギ抗ヒトIgGと室温で2時間インキュベーションした。このブロットを再洗浄し、そしてNBT/BCIP基質中で発色させた。

40

【0203】

ヒト抗体のアフィニティー精製

各100 μ gの化膿性連鎖球菌精製タンパク質を、ニトロセルロース片に接着させ、5% BLOTTOで15分間ブロッキングし、そしてその後、PBSでリンスした。血清を4℃で一晩吸着させた後、PBSでニトロセルロース片を洗浄し、そしてpH 3.0の100 mMグリシンでリンスして結合した抗体を溶出させた。溶出した抗体を1 M Tri

50

s pH 8.8で中和し、そしてPBS中で透析した。SF-370株に対するOPAのため、これらの抗体をPMNおよびヒト全血と共に試験した。

【0204】

オプソニン食作用アッセイ (OPA)

化膿性連鎖球菌株SF-370を用いて、THYブロスに接種し、そして静置状態で一晚増殖させた。一晚培養を新鮮な培地で希釈し、そして0.5~0.7のOD₆₅₀までさらに培養した。細胞を遠心分離し、PBSで1回洗浄し、そして氷冷PBSに再懸濁して0.5のOD₆₅₀にした。細胞をPBS中で1:5,000に希釈し、そして試験抗体または抗血清と4℃で30分間混合した。標的細胞あたり100および200エフェクター細胞の比で、あらかじめ温めたPMNを細菌および抗体に添加した。反応を37℃で1時間、震盪装置上でインキュベーションし、そして最後に氷冷PBSで反応を停止し、そして二つ組でBHI寒天上に蒔いた。 10

【0205】

全ヒト血液を用いたOPA

個々のヘパリン処理ヒト血液を得て、そして使用まで37℃で15~30分間インキュベーションした。記載するように細菌を調製し、そして50μl試験抗体と4℃で15分間インキュベーションし、その後、430μlの全血を添加した。反応を37℃で1.5時間、震盪装置上でインキュベーションし、そして二つ組でBHI寒天上に蒔いた。各実験は、個体の全血試料に相当し、プールではない。

【0206】

結果

全細胞ELISA

ORF特異的抗体が全細胞の表面に反応する能力を、ELISAによって試験した。先に記載するように、マウスで抗体を産生した。反応性は、化膿性連鎖球菌細胞表面上に発現するタンパク質の量および/または抗体が結合するのを可能にする方式でのタンパク質の曝露の相違を実証する。ELISA力価を表XVに示し、そしてこの力価は、発現したタンパク質の量、または抗体反応性を可能にするよう曝露されたエピトープ数のいずれかの相違を反映する、反応性の範囲を示す。免疫前バックグラウンド力価よりはるかに優れた値を太字で示す。

【0207】

表XV. 化膿性連鎖球菌ORFに対する全細胞ELISA力価

【0208】

【表15】

10

20

30

ORF番号	ELISA力価	ORF番号	ELISA力価	
68	1,635	1358	6,201	
73	1,702	1487	4,007	
145	2,105	1659	3,240	
218	1,139	1664	5,355	
232	1,277	1698	2,032	
309	1,456	1723	1,273	10
347	2,766	1788	3,324	
433	1,431	1789	1,475	
554	22,873	1818	40,271	
661	1,727	1820	2,498	
668	1,869	1878	895	
678	2,144	1983	1,179	
685	3,094	2015	1,800	20
704	1,716	2019	24,669	
721	680	2064	1,486	
729	1,381	2258	4,962	
747	11,733	2379	19,220	
850	4,861	2417	4,225	
967	4,823	2450	4,255	
1157	1,827	2452	2,256	
1191	1,248	2459	2,166	30
1202b	1,194	2477	5,412	
1218	220,289	2497	666	
1224	21,170	2593	8,602	
1284	1,374	2601	2,000	
1316	6,407			

遺伝子保存

いくつかの連鎖球菌株のPCR解析を行って、多様なORFの保存の度合いを決定した。この解析からの結果を図11に見ることが可能である。すべてのPCR産物をゲル電気泳動によって解析し、そしてバンドサイズを予測値と比較した。陽性と示すすべてのORFは、予測された大きさに移動するPCR産物を示した。データは、高い度合いのゲノム保存を示し、試験した24のORFのうち21が、化膿性連鎖球菌の11株すべてに渡って保存されていた。さらに、18がC群およびG群間で保存され；最低量の保存は、B群連鎖球菌の株で観察された。

【0209】

選択された化膿性連鎖球菌ORFの定量的PCR

定量的PCRを行って、化膿性連鎖球菌ゲノムに含有されるいくつかのORFの転写を検

証した。さらに、シミュレーション感染モデルにおいて、遺伝子発現を *in vivo* で検証する手段として、この方法を用いた。2つの既知の転写制御因子、*rofA* および *Mga*、並びに他の1つのハウスキーピング遺伝子、*gyrA* がさらなる対照として含まれた。試験したすべての遺伝子が発現され、そして条件に応じて、いくつかは転写レベルに変動を示した。増幅がバックグラウンドを超えて検出可能であるPCR周期を示す、*Ct* 数でその値を表す。したがって、より低い *Ct* 値は、より多量の mRNA が出発材料に存在したことを示す。1の *Ct* 相違は、検出される mRNA 量の2倍の相違に相関する。図12は、この解析結果を示す。すべてのORFは、テンプレート不含対照より、有意に低い *Ct* 値を示した。ORF2019は、肺または *in vitro* 培養いずれかで観察されるより、大腿で、155倍低い発現を示した。一方、ORF2477は、感染8時間後に肺から抽出した際、大腿または *in vitro* 培養に比較して、mRNA レベルに49倍の増加を示した。これらのデータは、試験したORFすべてが、*in vitro* および *in vivo* で転写され、そして細菌が曝露される条件によって影響を受けたことを示す。

【0210】

化膿性連鎖球菌タンパク質に対するヒト血清の反応性

ヒト血清から抗体を精製して、PMNが化膿性連鎖球菌を飲み込んで、そして殺す能力を、ORF特異的抗体が増進する能力を試験した。図は、ドットプロットによる、いくつかの化膿性連鎖球菌タンパク質に対するヒト血清の反応性を示し、この血清が、オプソニン食作用研究用抗体の供給源として適していることを示す。表XVIは、これらのプロットの結果を要約する。プロットの結果は、24のORFタンパク質のうち14がヒト血清との反応性に関して陽性の試験結果になったことを示す。類似の実験において、単一のヒト血清をタンパク質に対して試験し、そして結果は表XVIに示すものと同一であった。その反応性および入手可能な材料の量に基づいて、アフィニティー精製抗体研究で使用するため、いくつかのタンパク質を選択した。

【0211】

表XVI. 反応性タンパク質に関するORF同定

【0212】

【表16】

	A	B	C	D	E	F	G	H
1	ScpA	145	232	554	668	721	1224	1284
2	2452	1659	1698	1788	1818	1820	2379	2459
3	2477	2593	2601	1218	433	1358	2019	1664

注：太字＝陽性

精製PMNを用いた、アフィニティー精製ヒト抗ORF抗体のオプソニン食作用活性

4つのヒト血液試料のプールからPMNを精製し、そして化膿性連鎖球菌SF-370の増殖は上述のとおりであった。細菌、PBS希釈剤およびPMNは、陰性対照として働いた。試験抗体を含有する反応から回収したCFUを、陰性対照のものを含有する反応から回収したCFUで割ることによって、殺菌パーセントを計算した。表XVIIに要約するこれらの研究の結果は、精製PMNとインキュベーションした際に、アフィニティー精製抗体が、SF-370に対するオプソニン活性を有することを示す。特に、ScpAおよびORF1224に対する抗体は、これらを試験した3回すべてで、OPAで測定されるように、陰性対照に対して、50%殺菌より高い殺菌を生じた。

【0213】

表XVII. エフェクター細胞として精製PMNを用いた、化膿性連鎖球菌タンパク質に対するアフィニティー精製ヒト抗体のオプソニン食作用活性

10

20

30

40

50

【0214】

【表17】

	ORF抗体のオプソニン食作用殺菌（パーセント） ¹					
	ScpA	1224	1218	145	2459	1698
実験1	60	64	63	ND	ND	ND
実験2	65	53	59	ND	ND	ND
実験3	62	85	45	71	31	61
平均	62.3	67.3	55.7	71	31	61

10

¹ 陰性対照と比較した際のオプソニン食作用活性。PMN対細菌の比は100：1であった。アフィニティー精製抗体は、反応混合物の10%であった（1：10希釈）。

【0215】

ND＝データなし

全血を用いた、アフィニティー精製ヒト抗体のオプソニン食作用活性

化膿性連鎖球菌を用いた伝統的なOPAは、エフェクター細胞供給源として全血を利用してきた。全血の存在下で、アフィニティー精製抗体がオプソニン活性を有するかどうかを決定するため、実験を行った。結果を表XVIIに要約し、そしてこれは、PMNの供給源として血液を用いた個体に応じて多様な結果を示す。しかし、ORF1224および125に対する抗体は、試験した個々の血液試料7つすべてで、一貫してより大きいOPA力価を生じた。対照的に、ScpAに対する抗体は、7つの血液試料すべてで、一貫して劣ったOPA力価を生じた。ScpAに対する抗体をPMNと試験した際、3回のアッセイのうち3回で、50%より高い殺菌があったため、これは予期されないことであった。5つの他のタンパク質に対する抗体は、相同株に対し、化膿性連鎖球菌SF-370に対して、より一貫しないOPAを有した。ORF1284に対する抗体が、7回の実験のうち4回で、50%より高い殺菌を生じたことに注目すべきである。

20

【0216】

表XVII. エフェクター細胞の供給源として全血を用いたOPA

【0217】

30

【表18】

被験者	ORF抗体のオプソニン食作用殺菌（パーセント） ¹							
	S cpA	145	1224	1284	1698	1818	2459	1218
1	16	77	86	60	56	45	82	56
2	36	50	79	86	68	72	64	28
3	16	47	56	53	39	42	66	33
4	14	48	54	41	25	63	62	33
5	19	69	56	35	63	42	19	42
6	7	57	68	54	62	54	65	36
7	5	64	59	42	33	38	19	16
平均	14	58	64	51	32	50	47	33
標準偏差	10	12	13	17	20	13	25	12

40

50

¹ 全血、細菌およびPBSを含有する反応に比較した際の、オプソニン食作用活性

【実施例4】

【0218】

実施例4

化膿性連鎖球菌外毒素1の生物学的活性

生物学的活性に関して、SPEIを性質決定する研究を行った。データは、SPEIがスーパー抗原活性を有し、そしてT細胞受容体Vβ領域(TCR Vβ)6、7、9、および21、3を示すT細胞の増殖を非特異的に誘導することを示す。

【0219】

SPEI

等電点電気泳動およびアフィニティークロマトグラフィーの組み合わせによって、SPEIを精製した。ドデシル硫酸ナトリウムポリアクリルアミドゲル電気泳動によって、精製毒素が均質であることが示された。

【0220】

スーパー抗原性アッセイ

ウェルあたり 2×10^5 細胞の濃度で、96ウェルマイクロタイタープレートのウェルに、ウサギ脾臓細胞を植え付けた。1.0 μg/ウェルから始めて、 10^{-8} μg/ウェルまで、毒素の10倍希釈を4つ組でウェルに添加した。陰性対照としてPBSのみの存在下で、そして陽性対照として他のSPEの存在下でインキュベーションした細胞に、これらの希釈を比較した。脾臓細胞を37℃で3日間増殖させ、そして1 μCi ³H-チミジンで一晩パルス処理した。細胞を翌日採取し、そしてDNAへの³H-チミジン取り込みによって測定されるような細胞増殖をシンチレーションカウンター(Beckman Instruments、カリフォルニア州フラートン)中で測定した。

【0221】

T細胞レパトリのフローサイトメトリー解析

Ficoll-Hypaque(Histopaque、Sigma)上での密度勾配沈降によって、3人の正常ヒトドナーから得た末梢血単核細胞(PBMC)を、ヘパリン処理静脈血から単離した。その後、細胞をハンクス平衡化塩溶液(HBSS)(Mediatech Cellgro、バージニア州ハーンドン)で3回洗浄し、そして細胞培養のため、培地に再懸濁した。10%熱不活性化ウシ胎児血清(FCS)(Gemini Bioproducts、カリフォルニア州ウッドランド)、20mM HEPES緩衝液(Mediatech Cellgro)、100U/mlペニシリン(Mediatech Cellgro)、100 μg/mlストレプトマイシン(Mediatech Cellgro)、および2mM Lグルタミン(Mediatech Cellgro)を補ったRPMI 1640(Mediatech Cellgro)中で、PBMC(1×10^6 細胞/ml)を培養した。抗CD3(20 ng/ml)、またはSPEI(100 ng/ml)いずれかの存在下で、細胞を3日間培養し、洗浄し、そして先に記載されるように、T細胞レパトリの免疫蛍光解析のため、洗浄し、そして染色する前に、インターロイキン2(50 U/ml)の存在下で、さらに1日増殖させた。

【0222】

フローサイトメトリー研究のため、PBMCをHBSSで洗浄し、そして染色溶液[5% FCS(Gemini Bioproducts)、1%免疫グロブリン(Alpha Therapeutic Corp.、カリフォルニア州ロサンゼルス)、0.02%アジ化ナトリウム(Sigma)を含むPBS]中で、 10×10^6 細胞/mlに再懸濁した。ヒトTCR Vβ2、3、5.1、5.2、7、8、11、12、13.1、13.2、14、16、17、20、21.3、22(Immunotech、メイン州ウェストブルック)、TCR Vβ9、23(Pharmingen、カリフォルニア州サンディエゴ)およびTCR Vβ6.7フルオレセインイソチオシアネート(FITC)(Endogen、マサチューセッツ州ウォバーン)に対する、一団のビオチン化モノクローナル抗体で、96ウェル丸底プレート中で細胞を染色し、その後、暗所中、37℃で3

10

20

30

40

50

0 分間インキュベーションした。インキュベーション期間後、300 x g、4℃で5分間遠心分離することによって、洗浄緩衝液 [PBS、2% FCS (Gemini Bio products)、0.02% アジ化ナトリウム (Sigma)] で細胞を2回洗浄した。細胞ペレットを染色溶液に再懸濁し、そして抗CD3 アロフィコシアニン (APC)、抗CD4 フィコエリトリン (PE) (Becton Dickinson、カリフォルニア州サンノゼ)、抗CD8 (FITC) (Becton Dickinson) およびストレプトアビジン・ペリディニン・クロロフィルタンパク質 (PerCP) コンjugate (Becton Dickinson) と、4℃で30分間インキュベーションした。300 x g、4℃で5分間遠心分離することによって、染色した細胞を、再び洗浄緩衝液で2回、そしてPBS中の0.02% アジ化ナトリウム (Sigma) で1回洗浄した。最後に、PBS中の1% (v/v) ホルムアルデヒド (Polysciences、ペンシルバニア州ワーリングトン) 200 µl で細胞を固定した。先に記載されるように、4色フローサイトメトリー (FACS Calibur、Becton Dickinson) を用いて解析を行った。血球計算器セットアップおよびデータ獲得法もまた、先に記載されている。Cellquest プログラム (Becton Dickinson) を用いて、リストモードマルチパラメータデータファイル (前方散乱、側方散乱、および4蛍光パラメータを含む各ファイル) を解析した。T細胞芽細胞集団に対して設定した光散乱ゲートで、活性化集団の解析を行った。陰性対照試薬を用いて、実験抗体の染色特異性を検証した。

10

20

【0223】

ミニ浸透圧ポンプ

6匹のアメリカン・ダッチ・ベルテッド (American Dutch belted) ウサギ3群の左わき腹に、500 µg のSPE I または200 µg のTSST-1 を含有する、皮下ミニ浸透圧ポンプを移植した。15日間の期間に渡って、毒素の致死性を評価した。

【0224】

結果

DNAへの3Hチミジン取り込みによって測定されるように、4日間のアッセイで、ウサギ脾臓細胞増殖を誘導する能力に関して、SPE I を評価した (図14)。SPE I は、やはり図に含まれる対照SPE毒素に匹敵して分裂促進性であった。SPE I の分裂促進活性の完全な減退は、他の毒素で観察されるのと同様に、 10^{-6} から 10^{-7} µg/ウェルの間であった。

30

【0225】

SPE I は、抗CD3抗体で刺激した細胞に比較して、TCR Vβ6、7、9、および21、3を所持するヒトT細胞を有意に刺激し (図15)、これはSPE I がスーパー抗原であることと一致した。いくつかのT細胞集団、例えばTCR Vβ14または17を持つT細胞は、抗CD3抗体で刺激した細胞に比較して、有意に減少した。

【0226】

大部分の発熱性毒素スーパー抗原は、皮下移植ミニ浸透圧ポンプ中、200から500 µgの毒素濃度でウサギに投与した際、致死性である。SPE I は500 µg用量でこの特性を示さなかった (3/3生存)。対照的に200 µgのTSST-1は完全に致死性であった (3/3死亡 (succumbed))。

40

【0227】

考察

発熱性毒素スーパー抗原は、Tリンパ球増殖を非特異的に誘導する能力によって定義されるが、T細胞受容体のβ鎖の可変部分の組成に依存する (6)。したがって、例えば、TSST-1は、応答するT細胞の抗原特異性に関わらず、TCR Vβ2を所持するヒトT細胞いずれかの増殖を刺激するであろう。この高レベルの刺激は、T細胞およびマクロファージ両方からのサイトカインの大規模放出を導く。特に重要なのは、TSSTと関連する低血圧およびショックを引き起こす腫瘍壊死因子αおよびβの放出である。

50

【0228】

データは、S P E I がスーパー抗原としてT細胞を刺激することを示す。したがって、S P E I は、T C R V β 6、7、9、および21、3を含有するヒト末梢血単核細胞を増殖させる。これらの選択されたT細胞集団のこの上昇は、同時に起こる非刺激T細胞の相対的な減少と共に、S P E I の特徴的シグナルであり、そしてV β 歪曲 (s k e w i n g) と称される。

【0229】

さらに、多くの発熱性毒素スーパー抗原は、T S S のモデルのように(8)、皮下移植したミニ浸透圧ポンプ中でウサギに投与した際、致死性である。これらのポンプは、7日間の期間に渡って、一定量の毒素を放出するように設計される。しかし、実験は15日間続け、これは、ウサギが最大その期間まで、投与された毒素に屈服する可能性があるためである。S P E I は、T S S のこのモデルでは致死性ではなかった。多くの発熱性毒素スーパー抗原がこのアッセイで致死性であるが、注目に値する例外がある。例えば、新たに同定されたブドウ球菌内毒素LおよびQはこのモデルでは致死性でないが、これらの2つの毒素は、ファミリーの期待されるすべての他の活性を共有する(スーパー抗原性を含む)。これらの後者の毒素に関しては、これらが7日間の毒素放出期間全体で、ミニ浸透圧ポンプ中で安定でないか、またはポンプ中に沈殿するかいずれかであると示唆されている。したがって、S P E I は、発熱性毒素スーパー抗原の決定的なスーパー抗原特性を共有する。

【0230】

特定の態様に関して上に例示し、そして記載するが、にもかかわらず、本発明は、示される詳細に限定されることを意図しない。むしろ、請求項の均等物の適用範囲および範囲内で、そして本発明の精神から逸脱することなく、細部に多様な修飾を行うことが可能である。

【0231】

参考文献

【0232】

【化1-1】

10

20

1. 1997. Case definitions for Infectious Conditions Under Public Health Surveillance. CDC.
2. Alm, R. A., L. S. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori* [published erratum appears in Nature 1999 Feb 25;397(6721):719]. Nature. 397:176-80. 10
3. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-402.
4. Anderson, T. F. 1951. Techniques for the preservation of three-dimensional structure in preparing specimens for the electron microscope. Trans N Y Acad Sci. 13:130-134. 20
5. Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27:573-80.
6. Chen, C. C., and P. P. Cleary. 1989. Cloning and expression of the streptococcal C5a peptidase gene in *Escherichia coli*: linkage to the type 12 M protein gene. Infect. Immun. 57:1740-1745.
7. Chmouryguina, I., A. Suvorov, P. Ferrieri, and P. P. Cleary. 1996. Conservation of the C5a peptidase genes in group A and B streptococci. Infect. Immun. 64:2387-2390. 30
8. Cockerill, F. R., 3rd, R. L. Thompson, J. M. Musser, P. M. Schlievert, J. Talbot, K. E. Holley, W. S. Harmsen, D. M. Ilstrup, P. C. Kohner, M. H. Kim, B. Frankfort, J. M. Manahan, J. M. Steckelberg, F. Roberson, and W. R. Wilson. 1998. Molecular, serological, and clinical features of 16 consecutive cases of invasive streptococcal disease. Southeastern Minnesota Streptococcal Working Group. Clin Infect Dis. 26:1448-58.
9. Courtney, H. S., Y. Li, J. B. Dale, and D. L. Hasty. 1994. Cloning, sequencing, and expression of a fibronectin/fibrinogen-binding protein from group A streptococci. Infect Immun. 62:3937-46. 40

10. Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Engineering*. 10:673-6.
11. Cunningham, M. W., and A. Quinn. 1997. Immunological crossreactivity between the class I epitope of streptococcal M protein and myosin. *Adv Exp Med Biol*. 418:887-92.
12. Dale, J. B., R. W. Baird, H. S. Courtney, D. L. Hasty, and M. S. Bronze. 1994. Passive protection of mice against group A streptococcal pharyngeal infection by lipoteichoic acid. *J Infect Dis*. 169:319-23. 10
13. Dale, J. B., M. Simmons, E. C. Chiang, and E. Y. Chiang. 1996. Recombinant, octavalent group A streptococcal M protein vaccine. *Vaccine*. 14:944-8.
14. Dale, J. B., R. G. Washburn, M. B. Marques, and M. R. Wessels. 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. *Infect Immun*. 64:1495-501. 20
15. Eddy, S. R. 1996. Hidden Markov models. *Cur Opin Struct Bio*. 6:361-5.
16. Ellen, R. P., and R. J. Gibbons. 1972. M protein-associated adherence of *Streptococcus pyogenes* to epithelial surfaces: prerequisite for virulence. *Infect Immun*. 5:826-830.
17. Eng, J. K., A. L. McCormack, and J. R. Yates, 3rd. 1994. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *Am Soc Mass Spectrometry*. 5:976-89. 30
18. Fischetti, V. A., V. Pancholi, and O. Schneewind. 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. *Mol Microbiol*. 4:1603-5.
19. Fogg, G. C., and M. G. Caparon. 1997. Constitutive expression of fibronectin binding in *Streptococcus pyogenes* as a result of anaerobic activation of *rofA*. *J Bacteriol*. 179:6172-80.
20. Foster, T. J., and M. Hook. 1998. Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol*. 6:484-8. 40

21. Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, R. Lathigra, O. White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J. F. Tomb, R. D. Fleischmann, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. van Vugt, N. Palmer, M. D. Adams, J. Gocayne, J. C. Venter, and et al. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi* [see comments]. *Nature*. 390:580-6.
22. Hacker, J., G. Blum-Oehler, I. Muhldorfer, and H. Tschape. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol*. 23:1089-97. 10
23. Hanski, E., and M. Caparon. 1992. Protein F, a fibronectin-binding protein, is an adhesion of the group A streptococcus *Streptococcus pyogenes*. *Proc Natl Acad Sci., USA*. 89:6172-76.
24. Hanski, E., P. A. Horwitz, and M. G. Caparon. 1992. Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. *Infect Immun*. 60:5119-5125. 20
25. Hernandez-Sanchez, J., J. G. Valadez, J. V. Herrera, C. Ontiveros, and G. Guarneros. 1998. lambda bar minigene-mediated inhibition of protein synthesis involves accumulation of peptidyl-tRNA and starvation for tRNA. *EMBO Journal*. 17:3758-65.
26. Huang, T. T., H. Malke, and J. J. Ferretti. 1989. The streptokinase gene of group A streptococci: cloning, expression in *Escherichia coli*, and sequence analysis. *Mol Microbiol*. 3:197-205. 30
27. Hynes, W. L., A. R. Dixon, S. L. Walton, and L. J. Aridgides. 2000. The extracellular hyaluronidase gene (*hylA*) of *Streptococcus pyogenes*. *FEMS Microbiol Lett*. 184:109-12.
28. Hynes, W. L., L. Hancock, and J. J. Ferretti. 1995. Analysis of a second bacteriophage hyaluronidase gene from *Streptococcus pyogenes*: evidence for a third hyaluronidase involved in extracellular enzymatic activity. *Infect Immun*. 63:3015-20.
29. Isberg, R. R., and G. Tran Van Nhieu. 1994. Binding and internalization of microorganisms by integrin receptors. *Trends Microbio*. 2:10-4. 40

30. Jones, K. F., and V. A. Fischetti. 1988. The importance of the location of antibody binding on the M6 protein for opsonization and phagocytosis of group A M6 streptococci. *J Exp Med.* 167:1114-23.
31. Kihlberg, B. M., M. Collin, A. Olsen, and L. Bjorck. 1999. Protein H, an antiphagocytic surface protein in *Streptococcus pyogenes*. *Infect Immun.* 67:1708-14.
32. Koebnik, R. 1995. Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins [letter; comment]. *Molecular Microbiology.* 16:1269-70. 10
33. Loessner, M. J., S. Gaeng, and S. Scherer. 1999. Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of *Staphylococcus aureus* bacteriophage 187. *J Bacteriol.* 181:4452-60.
34. Lukashin, A. V., and M. Borodovsky. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107-15.
35. Lukomski, S., C. A. Montgomery, J. Rurangirwa, R. S. Geske, J. P. Barrish, G. J. Adams, and J. M. Musser. 1999. Extracellular cysteine protease produced by *Streptococcus pyogenes* participates in the pathogenesis of invasive skin infection and dissemination in mice. *Infect Immun.* 67:1779-88. 20
36. Madore, D. V. 1998. Characterization of immune response as an indicator of *Haemophilus influenzae* type b vaccine efficacy. *Pediatr Infect Dis J.* 17:S207-10.
37. Matsuka, Y. V., S. Pillai, S. Gubba, J. M. Musser, and S. B. Olmsted. 1999. Fibrinogen cleavage by the *Streptococcus pyogenes* extracellular cysteine protease and generation of antibodies that inhibit enzyme proteolytic activity. *Infect Immun.* 67:4326-33. 30
38. Mazmanian, S. K., G. Liu, H. Ton-That, and O. Schneewind. 1999. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science.* 285:760-3.
39. McAtee, C. P., K. E. Fry, and D. E. Berg. 1998. Identification of potential diagnostic and vaccine candidates of *Helicobacter pylori* by "proteome" technologies. *Helicobacter.* 3:163-9. 40
40. McAtee, C. P., M. Y. Lim, K. Fung, M. Velligan, K. Fry, T. Chow, and D. E. Berg. 1998. Identification of potential diagnostic and vaccine candidates of *Helicobacter*

pylori by two-dimensional gel electrophoresis, sequence analysis, and serum profiling. Clin Diagn Lab Immunol. 5:537-42.

41. McAtee, C. P., M. Y. Lim, K. Fung, M. Velligan, K. Fry, T. P. Chow, and D. E. Berg. 1998. Characterization of a *Helicobacter pylori* vaccine candidate by proteome techniques. J Chromatogr B Biomed Sci Appl. 714:325-33.

42. Mejlhede, N., J. F. Atkins, and J. Neuhard. 1999. Ribosomal -1 frameshifting during decoding of *Bacillus subtilis* cdd occurs at the sequence CGA AAG. J. Bacteriol. 181:2930-7.

10

43. Molinari, G., S. R. Talay, P. Valentin-Weigand, M. Rohde, and G. S. Chhatwal. 1997. The fibronectin-binding protein of *Streptococcus pyogenes*, SfbI, is involved in the internalization of group A streptococci by epithelial cells. Infect Immun. 65:1357-63.

44. Nakai, K., and M. Kanehisa. 1991. Expert system for predicting protein localization sites in gram-negative bacteria. Proteins. 11:95-110.

45. Navarre, W. W., and O. Schneewind. 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol Mol Biol Rev. 63:174-229.

20

46. Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering. 10:1-6.

47. Nizet, V., B. Beall, D. J. Bast, V. Datta, L. Kilburn, D. E. Low, and J. C. De Azavedo. 2000. Genetic locus for streptolysin S production by group A streptococcus. Infect Immun. 68:4245-54.

30

48. Nordstrand, A., W. M. McShan, J. J. Ferretti, S. E. Holm, and M. Norgren. 2000. Allele substitution of the streptokinase gene reduces the nephritogenic capacity of group A streptococcal strain NZ131. Infect Immun. 68:1019-25.

49. Olmsted, S. B., S. L. Erlandsen, G. M. Dunny, and C. L. Wells. 1993. High-resolution visualization by field emission scanning electron microscopy of *Enterococcus faecalis* surface proteins encoded by the pheromone-inducible conjugative plasmid pCF10. J Bacteriol. 175:6229-37.

40

50. Park, J., and S. A. Teichmann. 1998. DIVCLUS: an automatic method in the GEANFAMMER package that finds homologous domains in single- and multi-domain proteins. *Bioinformatics*. 14:144-50.

51. Parkhill, J., M. Achtman, K. D. James, S. D. Bentley, C. Churcher, S. R. Klee, G. Morelli, D. Basham, D. Brown, T. Chillingworth, R. M. Davies, P. Davis, K. Devlin, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Leather, S. Moule, K. Mungall, M. A. Quail, M. A. Rajandream, K. M. Rutherford, M. Simmonds, J. Skelton, S. Whitehead, B. G. Spratt, and B. G. Barrell. 2000. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491 [see comments]. *Nature*. 404:502-6.

10

52. Pierschbacher, M. D., and E. Ruoslahti. 1987. Influence of stereochemistry of the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J Biol Chem*. 262:17294-8.

53. Pizza, M., V. Scarlato, V. Maignani, M. M. Giuliani, B. Arico, M. Comanducci, G. T. Jennings, L. Baldi, E. Bartolini, B. Capecchi, C. L. Galeotti, E. Luzzi, R. Manetti, E. Marchetti, M. Mora, S. Nuti, G. Ratti, L. Santini, S. Savino, M. Scarselli, E. Storni, P. Zuo, M. Broecker, E. Hundt, B. Knapp, E. Blair, T. Mason, H. Tettelin, D. W. Hood, A. C. Jeffries, N. J. Saunders, D. M. Granoff, J. C. Venter, E. R. Moxon, G. Grandi, and R. Rappuoli. 2000. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing [see comments]. *Science*. 287:1816-20.

20

54. Podbielski, A., A. Flosdorff, and J. Weber-Heynemann. 1995. The group A streptococcal virR49 gene controls expression of four structural vir regulon genes. *Infect Immun*. 63:9-20.

30

55. Proft, T., S. Louise Moffatt, C. J. Berkahn, and J. D. Fraser. 1999. Identification and Characterization of Novel Superantigens from *Streptococcus pyogenes*. *J Exp Med*. 189:89-102.

56. Pugsley, A. P. 1993. The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev*. 57:50-108.

57. Quinn, A., K. Ward, V. A. Fischetti, M. Hemric, and M. W. Cunningham. 1998. Immunological relationship between the class I epitope of streptococcal M protein and myosin. *Infect Immun*. 66:4418-24.

40

58. Reda, K. B., V. Kapur, D. Goela, J. G. Lamphear, J. M. Musser, and R. R. Rich. 1996. Phylogenetic distribution of streptococcal superantigen SSA allelic variants provides evidence for horizontal transfer of ssa within *Streptococcus pyogenes*. *Infect Immun.* 64:1161-5.
 59. Salzberg, S. L., A. L. Delcher, S. Kasif, and O. White. 1998. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* 26:544-8.
 60. Sonnenberg, M. G., and J. T. Belisle. 1997. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. *Infect Immun.* 65:4515-24. 10
 61. Bateman, A. T., R. Birney, S. P. Durbin, K. L. Howe and E. L. L. Sonnhammer. 2000. The Pfam protein families database. *Nuc. Acids. Res.* 28:263-6.
 62. Stevens, D. L. 1995. Streptococcal toxic-shock syndrome: spectrum of disease, pathogenesis, and new concepts in treatment. *Emerg Infect Dis.* 1:69-78. 20
 63. Stockbauer, K. E., L. Magoun, M. Liu, E. H. Burns, Jr., S. Gubba, S. Renish, X. Pan, S. C. Bodary, E. Baker, J. Coburn, J. M. Leong, and J. M. Musser. 1999. A natural variant of the cysteine protease virulence factor of group A streptococcus with an arginine-glycine-aspartic acid (RGD) motif preferentially binds human integrins α 5 β 3 and α IIb β 3 [In Process Citation]. *Proc Natl Acad Sci., USA.* 96:242-7.
 64. Ton-That, H., G. Liu, S. K. Mazmanian, K. F. Faull, and O. Schneewind. 1999. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc Natl Acad Sci U S A.* 96:12424-12429. 30
 65. Weldingh, K., I. Rosenkrands, S. Jacobsen, P. B. Rasmussen, M. J. Elhay, and P. Andersen. 1998. Two-dimensional electrophoresis for analysis of *Mycobacterium tuberculosis* culture filtrate and purification and characterization of six novel proteins. *Infect Immun.* 66:3492-500.
 66. Yutsudo, T., K. Okumura, M. Iwasaki, A. Hara, S. Kamitani, W. Minamide, H. Igarashi, and Y. Hinuma. 1994. The gene encoding a new mitogenic factor in a *Streptococcus pyogenes* strain is distributed only in group A streptococci. *Infection and Immunity.* 62:4000-4004. 40
- 【化 1 - 8 】
67. Published International Patent Application Number WO99/27944.
 68. U.S Patent Number 4,666,829.

【0233】

【図1】図1は、オープンリーディングフレーム（ORF）同定の図式的表示を示す。

【図2】図2は、トリプシン消化後の化膿性連鎖球菌の低電圧走査型電子顕微鏡写真（LV-SEM）を示し、この中で、細胞完全性が維持され、そして単層さえ存在している。バーは1 μm に等しい。

【図3】図3は、トリプシンでの消化前および消化後の化膿性連鎖球菌のLV-SEMを示す。パネルA（左のパネル）は、トリプシン消化前の細胞を示し、細胞はより大きく、そして表面物質を示す。パネルB（右のパネル）は、消化後の細胞を示し、細胞はより小さく、そしていかなる表面タンパク質も欠けて見える。バーは1 μm に等しい。

【図4】図4は、ORF 218にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。 10

【図5】図5は、ORF 554にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。

【図6】図6は、ORF 1191にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。

【図7】図7は、ORF 2064にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。

【図8】図8は、ORF 2601にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。

【図9】図9は、ORF 1316にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。 20

【図10】図10は、ORF 1224にコードされるタンパク質を発現する化膿性連鎖球菌のLV-SEMを示す。

【図11】図11は、いくつかの化膿性連鎖球菌株のPCR解析を示し、株に渡る遺伝子保存を例示する。

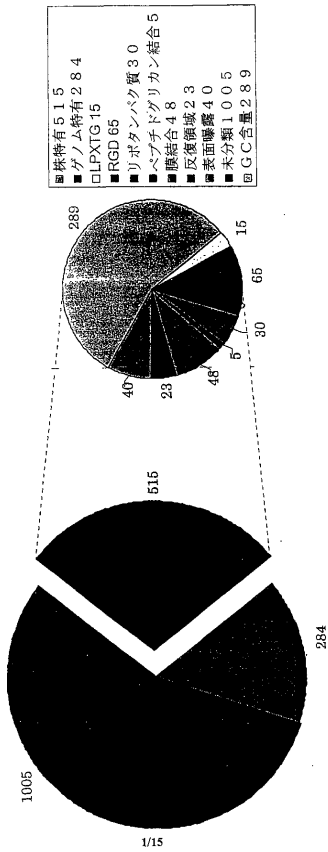
【図12】図12は、選択した化膿性連鎖球菌ORFの定量的PCR解析を示し、試験したすべてのORFが*in vitro*および*in vivo*で転写されていることを立証する。

【図13】図13は、ORF遺伝子産物とヒト血清の反応性を示すドットブロットを示す。 30

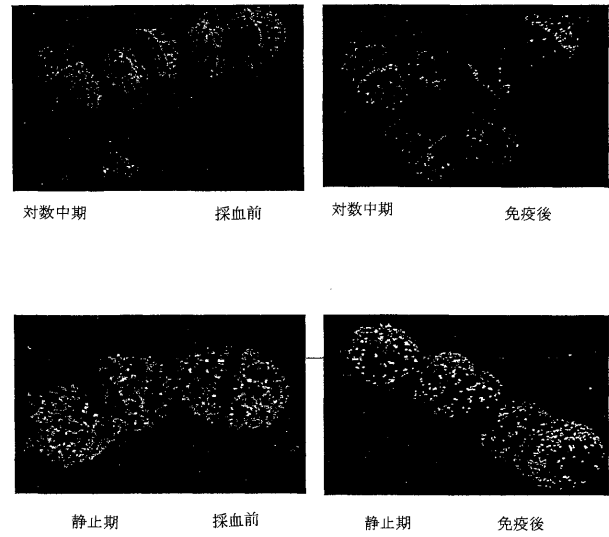
【図14】図14は、他のSPE類に比較した、ウサギ脾臓細胞増殖を誘導するSPE Iの能力を示す。

【図15】図15は、抗CD3抗体による刺激（白いバー）に比較した、SPE Iに誘導されるヒトT細胞受容体刺激プロフィール（黒いバー）を示す。

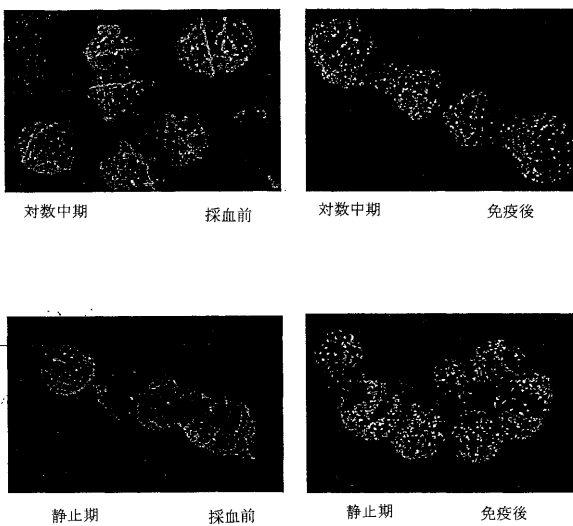
【図 1】



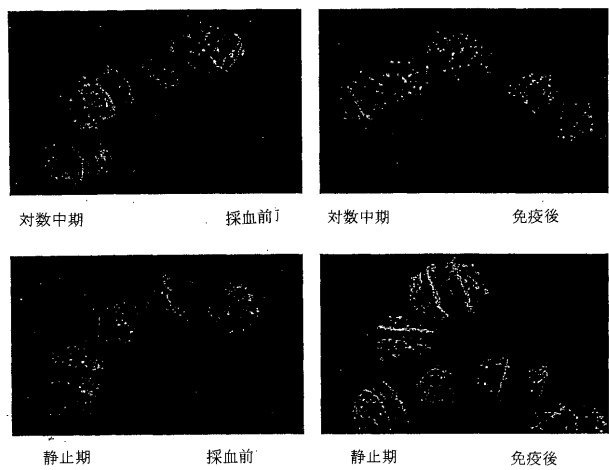
【図 4】



【図 5】



【図 6】



【図 7】

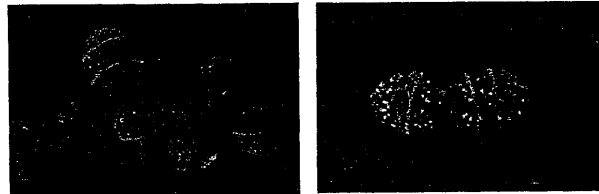


対数中期

採血前

対数中期

免疫後



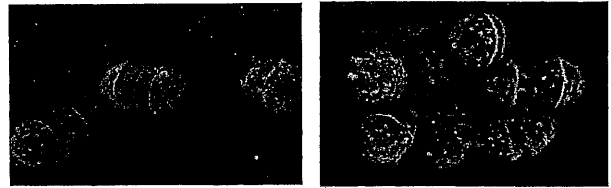
静止期

採血前

静止期

免疫後

【図 8】

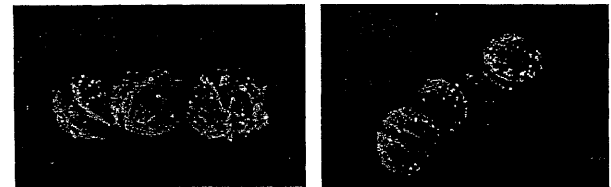


対数中期

採血前

対数中期

免疫後



静止期

採血前

静止期

免疫後

【図 9】

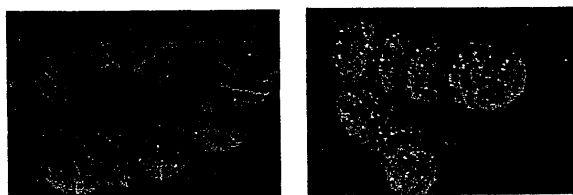


対数中期

採血前

対数中期

免疫後



静止期

採血前

静止期

免疫後

【図 10】



対数中期

採血前

対数中期

免疫後



静止期

採血前

静止期

免疫後

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
24 October 2002 (24.10.2002)

PCT

(10) International Publication Number
WO 02/083859 A2

(51) International Patent Classification: C12N

(21) International Application Number: PCT/US02/11610

(22) International Filing Date: 12 April 2002 (12.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/283,358 13 April 2001 (13.04.2001) US(71) Applicant (for all designated States except US): WYETH
[US/US]; Five Giralda Farms, Madison, NJ 07940 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EL, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PI, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

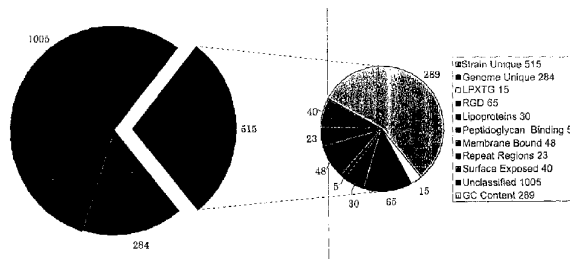
(75) Inventors/Applicants (for US only): OLMSTEAD, Stephen, Bruce [US/US]; 395 Kneag Road, Pittsford, NY 14534 (US); ZAGURSKY, Robert, John [US/US]; 569 Fox Hill Drive, Victor, NY 14564 (US); NICKBARG, Elliott, Bruce [US/US]; 32 Audrey Road, Belmont, MA 02140 (US); WINTER, Laurie, Anne [US/US]; 2915 Colby Street, Brockport, NY 14420 (US).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(74) Agents: SCHULMAN, Robert, M. et al.; Intellectual Property Department, Hinton & Williams, 1900 K Street, N.W., Suite 1200, Washington, DC 20006-1109 (US).

(54) Title: SURFACE PROTEINS OF *STREPTOCOCCUS PYOGENES*(57) Abstract: β -hemolytic streptococci polynucleotides, polypeptides, particularly *Streptococcus pyogenes* polypeptides and polynucleotides, and antibodies of these polypeptides are described. The polynucleotides, polypeptides, and antibodies of the invention can be formulated for use as immunogenic compositions. Also disclosed are methods for immunizing against and reducing β -hemolytic streptococcal infection, and for detecting β -hemolytic streptococci in a biological sample.

WO 02/083859 A2

WO 02/083859

PCT/US02/11610

SURFACE PROTEINS OF *STREPTOCOCCUS PYOGENES*

FIELD OF THE INVENTION

This invention relates generally to β -hemolytic streptococcal polypeptides and polynucleotides, particularly *Streptococcus pyogenes* polypeptides and polynucleotides.

- 5 More specifically, the invention relates to polypeptides of *Streptococcus pyogenes* which are surface localized, and antibodies of these polypeptides. The invention also relates to nucleotide sequences encoding polypeptides of *Streptococcus pyogenes*, and expression vectors including these nucleotide sequences. The invention further relates to immunogenic compositions, and methods for immunizing against and reducing β -hemolytic streptococcal
- 10 infection. The invention also relates to methods of detecting these nucleotides and polypeptides and for detecting β -hemolytic streptococci and *Streptococcus pyogenes* in a biological sample.

BACKGROUND OF THE INVENTION

- Traditional phenotypic criteria for classification of streptococci include both
- 15 hemolytic reactions and Lancefield serological groupings. However, with taxonomic advances, it is now known that unrelated species of β -hemolytic (defined as the complete lysis of sheep erythrocytes in agar plates) streptococci may produce identical Lancefield antigens and that strains genetically related at the species level may have heterogeneous Lancefield antigens. In spite of these exceptions to the traditional rules of streptococcal
- 20 taxonomy, hemolytic reactions and Lancefield serological tests can still be used to divide streptococci into broad categories as a first step in identification of clinical isolates. Ruoff, K.L., R.A. Whaley, and D. Beighton. 1999. *Streptococcus*. In P.R. Murray, E.J. Baron, M.A. Pfaller, F.C. Tenover, and R.H. Tenover (eds.), *Manual of Clinical Microbiology*. American Society of Microbiology Press, Washington D.C.

- 25 β -hemolytic isolates with Lancefield group A, C, or G antigen can be subdivided into two groups: large-colony (>0.5 mm in diameter) and small-colony (<0.5 mm in diameter) formers. Large-colony-forming group A (*Streptococcus pyogenes*), C, and G strains are "pyogenic" streptococci replete with a variety of effective virulence mechanisms. *Streptococcus agalactiae* (group B) is still identified reliably by its production of Lancefield
- 30 group B antigen or other phenotypic traits.

WO 02/083859

PCT/US02/11610

A need exists to develop compositions and methods to ameliorate and prevent infections caused by β -hemolytic streptococci, including groups A, B, C and G. Similarity between these species includes not only virulence factors, but also disease manifestations. Included in the latter are pneumonia, arthritis, abscesses, rhinopharyngitis, metritis, puerperal sepsis, neonatal septicemia, wound infections, meningitis, peritonitis, cellulitis, pyoderma, necrotizing fasciitis, toxic shock syndrome, septicemia, infective endocarditis, pericarditis, glomerulonephritis, and osteomyelitis.

Streptococcus pyogenes are gram-positive diplococci that colonize the pharynx and skin of humans, sites that then serve as the primary reservoir for this organism. An obligate parasite, this bacterium is transmitted by either direct contact of respiratory secretions or by hand-to-mouth. The majority of *Streptococcus pyogenes* infections are relatively mild illnesses, such as pharyngitis or impetigo. Currently, there are anywhere from twenty million to thirty-five million cases of pharyngitis alone in the U.S., costing about \$2 billion for physician visits and other related expenses. Additionally, nonsuppurative sequelae such as rheumatic fever, scarlet fever, and glomerulonephritis result from *Streptococcus pyogenes* infections. Globally, acute rheumatic fever (ARF) is the most common cause of pediatric heart disease (Bibliography entry 1).

From the initial portals of entry, pharynx, and skin, *Streptococcus pyogenes* can disseminate to other parts of the body where bacteria are not usually found, such as the blood, deep muscle and fat tissue, or the lungs, and can cause invasive infections. Two of the most severe but least common forms of invasive *Streptococcus pyogenes* disease are necrotizing fasciitis and streptococcal toxic shock syndrome (STSS). Necrotizing fasciitis (described in the media as "flesh-eating bacteria") is a destructive infection of muscle and fat tissue. STSS is a rapidly progressing infection causing shock and injury to internal organs such as the kidneys, liver, and lungs. Much of this damage is due to a toxemia rather than localized damage due to bacterial growth.

In 1995, invasive *Streptococcus pyogenes* infections and STSS became mandated reportable diseases. In contrast to the millions of individuals that acquire pharyngitis and impetigo, the U.S. Centers for Disease Control and Prevention (CDC) mandated case reporting indicates that in 1997 there were from 15,000 to 20,000 cases of invasive *Streptococcus pyogenes* disease in the United States, resulting in over 2,000 deaths (1). Other reports estimate invasive disease to be as high as 10-20 cases per 100,000 individuals

WO 02/083859

PCT/US02/11610

per year (62). More specifically, of the 15,000 to 20,000 cases of invasive disease, 1,100 to 1,500 are cases of necrotizing fasciitis and 1,000 to 1,400 are cases of STSS, with a 20% and 60% mortality rate, respectively. Also included in serious invasive disease are cases of myositis, which carries a fatality rate of 80% to 100%. An additional 10% to 15% of individuals with other forms of invasive group A streptococcal disease die. These numbers have increased since case reporting was initiated in 1995 and reflect a general trend that has occurred over the past decade or two. Additionally, it is commonly agreed that the stringency of the case definitions results in lower and, thus, misleading numbers, in that many cases are successfully resolved due to early diagnosis and treatment before the definition has been met.

While *Streptococcus pyogenes* remains exquisitely sensitive to penicillin and its derivatives, treatment does not necessarily eradicate the organism. Approximately 5% to 20% of the human population remain carriers depending on the season (62), despite antibiotic therapy. The reasons for this are not totally clear and may involve a variety of mechanisms. In cases of serious invasive infections, treatment often requires aggressive surgical intervention. For those cases involving STSS or related disease, clindamycin (a protein synthesis inhibitor) is the preferred antibiotic as it penetrates tissues well and prevents exotoxin production. There are reports of some resistance to tetracycline, sulfa, and most recently, erythromycin. Clearly, there remains a need for compositions to prevent and treat β -hemolytic infection.

Numerous virulence factors have been identified for *Streptococcus pyogenes*, some secreted and some surface localized. Although it is encapsulated, the capsule is composed of hyaluronic acid and is not suitable as a candidate antigen for inclusion in immunogenic compositions, since it is commonly expressed by mammalian cells and is nonimmunogenic (14). The T antigen and Group Carbohydrate are other candidates, but may also elicit cross-reactive antibodies to heart tissue. Lipoteichoic acid is present on the surface of *Streptococcus pyogenes*, but raises safety concerns similar to LPS.

The most abundant surface proteins fall into a family of proteins referred to as M or "M-like" proteins because of their structural similarity. While members of this class have similar biological roles in inhibiting phagocytosis, they each have unique substrate binding properties. The best characterized protein of this family is the helical M protein. Antibodies directed to homologous M strains have been shown to be opsonic and protective (12, 13, 16). Complicating the use of M protein as a candidate antigen is the fact that there have been

WO 02/083859

PCT/US02/11610

approximately 100 different serotypes of M protein identified with several more untyped. Typically, the Class I M serotypes, exemplified by serotypes M1, M3, M6, M12, and M18, are associated with pharyngitis, scarlet fever, and rheumatic fever and do not express immunoglobulin binding proteins. Class II M serotypes, such as M2 and M49, are associated
5 with the more common localized skin infections and the sequelae glomerulonephritis, and do express immunoglobulin binding proteins (54). It is important to note that there is little, if any, heterologous cross-reactivity of antibodies to M serotypes. Equally important is the role these antibodies play in rheumatic fever. Specific regions of M protein elicit antibodies that cross react with host heart tissue, causing or at least correlating with cellular damage (11, 57).

10 M and M-like proteins belong to a large family of surface localized proteins that are defined by the sortase-targeted LPXTG motif (38, 64). This motif, located near the carboxy-terminus of the protein, is first cleaved by sortase between the threonine and glycine residues of the LPXTG motif. Once cleaved, the protein is covalently attached via the carboxyl of threonine to a free amide group of the amino acid cross-bridge in the peptidoglycan, thus
15 permanently attaching the protein to the surface of the bacterial cell. Included in this family of sortase-targeted proteins are the C5a peptidase (6, 7), adhesins for fibronectin (9, 19, 23, 24), vitronectin, and type IV collagen, and other M-like proteins that bind plasminogen, IgA, IgG, and albumin (31).

Numerous secreted proteins have been described, several of which are considered to
20 be toxins. Most *Streptococcus pyogenes* isolates from cases of serious invasive disease and streptococcal toxic shock syndrome (STSS) produce streptococcal pyrogenic exotoxins (SPE) A and C (8). Other pyrogenic exotoxins have also been identified in the genomic *Streptococcus pyogenes* sequence completed at the University of Oklahoma, submitted to GenBank and assigned accession number AE004092, and have been characterized (55).

25 Other toxins such as Toxic Shock Like Syndrome toxin, Streptococcal Superantigen (58), and Mitogenic Factor (66) play lesser-defined roles in disease. Streptolysin O could also be considered a possible candidate antigen, because it causes the release of IL- β release. In addition, a variety of secreted enzymes have also been identified that include the Cysteine protease (35, 37), Streptokinase (26, 48), and Hyaluronidase (27, 28).

30 Given the number of known virulence factors produced by *Streptococcus pyogenes*, it is clear that an important characteristic for a successful β -hemolytic streptococcal immunogenic composition would be its ability to stimulate a response that would prevent or

WO 02/083859

PCT/US02/11610

limit colonization early in the infection process. This protective response would either block adherence and/or enhance the clearance of cells through opsonophagocytosis. Antibodies to M protein have been shown to be opsonic and provide a mechanism to overcome the anti-phagocytic properties of the protein (30) in much the same way that anti-serotype B capsular antibodies have demonstrated protection from disease caused by *Haemophilus influenzae* B (36). In addition, antibodies specific to Protein F have been shown to block adherence and internalization by tissue culture cells (43).

There remains a need to further identify immunogenic compositions, and methods for the prevention or amelioration of β -hemolytic streptococcal colonization or infection. There also remains a need to further identify surface proteins of *Streptococcus pyogenes* and polynucleotides that encode *Streptococcus pyogenes* polypeptides. Also, there remains a need for methods of detecting β -hemolytic streptococci and *Streptococcus pyogenes* colonization or infection.

SUMMARY OF THE INVENTION

To meet these and other needs, and in view of its purposes, the present invention provides compositions and methods for the prevention or amelioration of β -hemolytic streptococcal colonization or infection. The invention also provides *Streptococcus pyogenes* polypeptides and polynucleotides, recombinant materials, and methods for their production. Another aspect of the invention relates to methods for using such *Streptococcus pyogenes* polypeptides and polynucleotides.

The polypeptides of the invention include isolated polypeptides comprising at least one of an amino acid sequence of any of even numbered SEQ ID NOS: 2-668. The invention also includes amino acid sequences that have at least 70% identity to any of an amino acid sequence of even numbered SEQ ID NOS: 2-668, and mature polypeptides of the amino acid sequences any of even numbered SEQ ID NOS: 2-668. The invention further includes immunogenic fragments and biological equivalents of these polypeptides. Also provided are antibodies that immunospecifically bind to the polypeptides of the invention.

The polynucleotides of the invention include isolated polynucleotides that comprise nucleotide sequences that encode a polypeptide of the invention. These polynucleotides include isolated polynucleotides comprising at least one of a nucleotide sequence of any of odd numbered SEQ ID NOS: 1-667, and also include other nucleotide sequences that, as a

WO 02/083859

PCT/US02/11610

result of the degeneracy of the genetic code, also encode a polypeptide of the invention. The invention also includes isolated polynucleotides comprising a nucleotide sequence that has at least 70% identity to a nucleotide sequence that encodes a polypeptide of the invention, and isolated polynucleotides comprising a nucleotide sequences that has at least 70% identity to a nucleotide sequence any of odd numbered SEQ ID NOS: 1-667. In addition, the isolated polynucleotides of the invention include nucleotide sequences that hybridize under stringent hybridization conditions to a nucleotide sequence that encodes a polypeptide of the invention, nucleotide sequences that hybridize under stringent hybridization conditions to a nucleotide sequence of any of odd numbered SEQ ID NOS: 1-667, and nucleotide sequences that are fully complementary to these polynucleotides. Furthermore, the invention includes expression vectors and host cells comprising these polynucleotides.

The invention further provides methods for producing the polypeptides of the invention. In one embodiment, the method comprises the steps of (a) culturing a recombinant host cell of the invention under conditions suitable to produce a polypeptide of the invention and (b) recovering the polypeptide from the culture.

The invention also provides immunogenic compositions. In one embodiment, the immunogenic compositions comprise an immunogenic amount of at least one component which comprises a polypeptide of the invention in an amount effective to prevent or ameliorate a β -hemolytic streptococcal colonization or infection in a susceptible mammal. The component may comprise the polypeptide itself, or may comprise the polypeptide and any other substance (e.g., one or more chemical agents, proteins, etc.) that can aid in the prevention and/or amelioration of β -hemolytic streptococcal colonization or infection. These immunogenic compositions can further comprise at least a portion of the polypeptide, optionally conjugated or linked to a peptide, polypeptide, or protein, or to a polysaccharide. In another embodiment, the immunogenic compositions comprise an immunogenic amount of a component which comprises a polynucleotide of the invention, the component being in an amount effective to prevent or ameliorate a β -hemolytic streptococcal colonization or infection in a susceptible mammal. The component may comprise the polynucleotide itself, or may comprise the polynucleotide and any other substance (e.g., one or more chemical agents, proteins, etc.) that can aid in the prevention and/or amelioration of β -hemolytic streptococcal colonization or infection. In yet another embodiment, the immunogenic compositions comprise a vector that comprises a polynucleotide of the invention. The

WO 02/083859

PCT/US02/11610

immunogenic compositions of the invention can also include an effective amount of an adjuvant.

The invention also includes methods of protecting a susceptible mammal against β -hemolytic streptococcal colonization or infection. In one embodiment, the method comprises administering to a mammal an effective amount of an immunogenic composition comprising an immunogenic amount of a polypeptide of the invention, which amount is effective to prevent or ameliorate β -hemolytic streptococcal colonization or infection in the susceptible mammal. In another embodiment, the method comprises administering to the mammal an effective amount of an immunogenic composition comprising a polynucleotide of the invention, which amount is effective to prevent or ameliorate β -hemolytic streptococcal colonization or infection in the susceptible mammal. The immunogenic compositions of the invention can be administered by any conventional route, for example, by subcutaneous or intramuscular injection, oral ingestion, or intranasally.

The invention further includes compositions and methods for reducing at least one of the number and the growth of β -hemolytic streptococci in a mammal having a β -hemolytic streptococcal colonization or infection. In one embodiment, the composition comprises an antibody of the invention. In another embodiment, the composition comprises an antisense oligonucleotide capable of blocking expression of a nucleotide sequence encoding a polypeptide of the invention.

Also provided are methods for reducing side effects caused by β -hemolytic streptococcal infection in a mammal. In one embodiment, the method comprises administering to the mammal an effective amount of a composition comprising an antibody of the invention, which amount is effective to reduce at least one of the number of and the growth of β -hemolytic streptococci in the mammal. In another embodiment, the method comprises administering to the mammal an effective amount of a composition comprising an antisense oligonucleotide capable of blocking expression of a nucleotide sequence encoding a polypeptide of the invention, which amount is effective to reduce at least one of the number of and the growth of β -hemolytic streptococci in the mammal.

Also provided are methods for detecting and/or identifying β -hemolytic streptococci in a biological sample. In one embodiment, the method comprises (a) contacting the biological sample with a polynucleotide of the invention under conditions that permit

WO 02/083859

PCT/US02/11610

hybridization of complementary base pairs and (b) detecting the presence of hybridization complexes in the sample, wherein the detection of hybridization complexes indicates the presence of β -hemolytic streptococci in the biological sample. In another embodiment, the method comprises (a) contacting the biological sample with an antibody of the invention

5 under conditions suitable for the formation of immune complexes and (b) detecting the presence of immune complexes in the sample, wherein the detection of immune complexes indicates the presence of β -hemolytic streptococci in the biological sample. In yet another embodiment, the method comprises (a) contacting the biological sample with a polypeptide of the invention under conditions suitable for the formation of immune complexes and (b)
10 detecting the presence of immune complexes in the sample, wherein the detection of immune complexes indicates the presence of antibodies to β -hemolytic streptococci in the biological sample.

The invention further provides immunogenic compositions. In one embodiment, the immunogenic composition comprises at least one polypeptide of the invention. In another
15 embodiment, the immunogenic composition comprises at least one polynucleotide of the invention. In yet another embodiment, the immunogenic composition comprises at least one antibody of the invention.

Also provided is an isolated polynucleotide comprising a nucleotide sequence that has at least 70% identity to a nucleotide sequence that encodes a polypeptide of the invention, the
20 polynucleotide being identified by the steps comprising (a) obtaining a first and second PCR primer derived from a nucleotide that encodes a mature polypeptide of any of SEQ ID NOS: 2-668, wherein the first and second primers are capable of initiating nucleic acid synthesis in an outward manner under PCR conditions, and wherein the first primer is capable of being extended in an antisense direction and the second primer is capable of being extended in a
25 sense direction and (b) combining the first and second PCR primer with a cDNA library that contains the polynucleotide under PCR conditions suitable for synthesizing the nucleotide sequence from the first and second primers.

Also provided is a method for extending a polynucleotide of the invention using polymerase chain reaction (PCR), the method comprising the steps of (a) obtaining a first and
30 second PCR primer derived from the polynucleotide, wherein the first and second PCR primers are capable of initiating nucleic acid synthesis in an outward manner under PCR conditions, and wherein the first PCR primer is capable of being extended in an antisense

WO 02/083859

PCT/US02/11610

direction and the second PCR primer is capable of being extended in a sense direction and (b) combining the first and second PCR primers with the polynucleotide contained in a cDNA library under PCR conditions suitable for synthesizing nucleotide sequences from the first and second PCR primers, thereby extending the polynucleotide.

- 5 It is to be understood that the foregoing general description and the following detailed description are exemplary, but are not restrictive, of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 depicts a graphical representation of open reading frame (ORF) identification.

Fig. 2 depicts a low-voltage scanning electron micrograph (LV-SEM) of

- 10 *Streptococcus pyogenes* after digestion with trypsin, wherein cell integrity is maintained and an even monolayer is present. The bar equals 1 μm .

Fig. 3 depicts a LV-SEM of *Streptococcus pyogenes* before and after digestion with trypsin. Panel A (the left panel) shows cells before tryptic digestion, wherein the cells are larger and display surface material. Panel B (the right panel) shows cells after digestion, 15 wherein the cells are smaller and appear devoid of any surface proteins. The bars equal 1 μm .

Fig. 4 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by ORF 218.

- Fig. 5 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by 20 ORF 554.

Fig. 6 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by ORF 1191.

Fig. 7 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by ORF 2064.

- 25 Fig. 8 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by ORF 2601.

WO 02/083859

PCT/US02/11610

Fig. 9 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by ORF 1316.

Fig. 10 depicts a LV-SEM of *Streptococcus pyogenes* expressing protein encoded by ORF 1224.

5 Fig. 11 depicts PCR analysis of several *Streptococcus pyogenes* strains to illustrate gene conservation across the strains.

Fig. 12 depicts quantitative PCR analysis of selected *Streptococcus pyogenes* ORFs to demonstrate that all ORFs tested are transcribed *in vitro* and *in vivo*.

10 Fig. 13 depicts a dot blot showing reactivity of human serum with the ORF gene products.

Fig. 14 depicts ability of SPE I to induce rabbit splenocyte proliferation compared to other SPEs.

Fig. 15 depicts human T cell receptor stimulation profile induced by SPE I (black bars) compared to stimulation by anti CD3 antibodies (open bars).

15 DETAILED DESCRIPTION OF THE INVENTION

The present invention provides compositions and methods to ameliorate and prevent infections caused by all β -hemolytic streptococci, including groups A, B, C and G. To identify polynucleotides and polypeptides useful for the amelioration and prevention of infections caused by β -hemolytic streptococci, two strategies, a genomic approach and a
20 proteomic approach, were used to identify surface localized, *Streptococcus pyogenes* proteins.

The genomic approach included an extensive genomic analysis in silico of the *Streptococcus pyogenes* genome using several algorithms designed to identify and characterize genes that would encode surface localized proteins. The proteomic approach
25 was undertaken to identify proteins present on the surface of *Streptococcus pyogenes*. Reliance on both approaches was important to overcome the deficiencies of each approach. Genomic mining provides the genetic capabilities, but gives little information as to the actual phenotypic expression. Conversely, proteomic analysis identifies actual proteins localized to

WO 02/083859

PCT/US02/11610

the surface of the cell, but protein expression may be regulated and the specific conditions under which the bacterial cells are cultured may influence the set of proteins identified.

The results of the genomic and proteomic approaches were combined and the ORFs of interest were categorized into one of four groups: (i) ORFs encoding surface localized proteins identified by proteomics (Table I, odd numbered SEQ ID NOS: 1-147); (ii) ORFs encoding putative lipoproteins (Table II, odd numbered SEQ ID NOS: 149-181, 669); (iii) ORFs encoding putative polypeptides containing a LPXTG motif (Table III, odd numbered SEQ ID NOS: 183-187); and (iv) ORFs encoding other putative surface localized polypeptides (Table IV, odd numbered SEQ ID NOS: 189-667). The ORFs contained in Tables I-IV are non-redundant, i.e., the ORFs listed in Tables I-IV each appear once though many ORFs possess characteristics that match another table. Thus, for example, there are ORFs listed in Table I (ORFs encoding surface localized proteins identified by proteomics) that could also be classified in one or more of Tables II-IV, but are not included in those tables.

Table I. Open Reading Frames (ORFs) encoding surface localized proteins identified by proteomics

SEQ ID NO: 1 (ORF 66)	SEQ ID NO: 51 (ORF 1237)	SEQ ID NO: 101 (ORF 1975)
SEQ ID NO: 3 (ORF 102)	SEQ ID NO: 53 (ORF 1238)	SEQ ID NO: 103 (ORF 2019)
SEQ ID NO: 5 (ORF 145)	SEQ ID NO: 55 (ORF 1253)	SEQ ID NO: 105 (ORF 2064)
SEQ ID NO: 7 (ORF 232)	SEQ ID NO: 57 (ORF 1284)	SEQ ID NO: 107 (ORF 2086)
SEQ ID NO: 9 (ORF 238)	SEQ ID NO: 59 (ORF 1316)	SEQ ID NO: 109 (ORF 2106)
SEQ ID NO: 11 (ORF 436)	SEQ ID NO: 61 (ORF 1330)	SEQ ID NO: 111 (ORF 2116)
SEQ ID NO: 13 (ORF 516)	SEQ ID NO: 63 (ORF 1358)	SEQ ID NO: 113 (ORF 2120)
SEQ ID NO: 15 (ORF 554)	SEQ ID NO: 65 (ORF 1487)	SEQ ID NO: 115 (ORF 2123)
SEQ ID NO: 17 (ORF 589)	SEQ ID NO: 67 (ORF 1495)	SEQ ID NO: 117 (ORF 2202)
SEQ ID NO: 19 (ORF 661)	SEQ ID NO: 69 (ORF 1557)	SEQ ID NO: 119 (ORF 2214)
SEQ ID NO: 21 (ORF 668)	SEQ ID NO: 71 (ORF 1638)	SEQ ID NO: 121 (ORF 2330)
SEQ ID NO: 23 (ORF 678)	SEQ ID NO: 73 (ORF 1650)	SEQ ID NO: 123 (ORF 2354)
SEQ ID NO: 25 (ORF 704)	SEQ ID NO: 75 (ORF 1654)	SEQ ID NO: 125 (ORF 2377)
SEQ ID NO: 27 (ORF 743)	SEQ ID NO: 77 (ORF 1659)	SEQ ID NO: 127 (ORF 2379)
SEQ ID NO: 29 (ORF 825)	SEQ ID NO: 79 (ORF 1698)	SEQ ID NO: 129 (ORF 2387)
SEQ ID NO: 31 (ORF 850)	SEQ ID NO: 81 (ORF 1788)	SEQ ID NO: 131 (ORF 2417)
SEQ ID NO: 33 (ORF 934)	SEQ ID NO: 83 (ORF 1794)	SEQ ID NO: 133 (ORF 2420)
SEQ ID NO: 35 (ORF 993)	SEQ ID NO: 85 (ORF 1816)	SEQ ID NO: 135 (ORF 2422)
SEQ ID NO: 37 (ORF 1036)	SEQ ID NO: 87 (ORF 1818)	SEQ ID NO: 137 (ORF 2450)
SEQ ID NO: 39 (ORF 1140)	SEQ ID NO: 89 (ORF 1819)	SEQ ID NO: 139 (ORF 2459)
SEQ ID NO: 41 (ORF 1157)	SEQ ID NO: 91 (ORF 1850)	SEQ ID NO: 141 (ORF 2477)
SEQ ID NO: 43 (ORF 1191)	SEQ ID NO: 93 (ORF 1854)	SEQ ID NO: 143 (ORF 2586)
SEQ ID NO: 45 (ORF 1218)	SEQ ID NO: 95 (ORF 1878)	SEQ ID NO: 145 (ORF 2593)
SEQ ID NO: 47 (ORF 1224)	SEQ ID NO: 97 (ORF 1902)	SEQ ID NO: 147 (ORF 2601)
SEQ ID NO: 49 (ORF 1234)	SEQ ID NO: 99 (ORF 1943)	

WO 02/083859

PCT/US02/11610

Table II. Open Reading Frames (ORFs) encoding putative lipoproteins

SEQ ID NO: 149 (ORF 68)	SEQ ID NO: 161 (ORF 685)	SEQ ID NO: 173 (ORF 1789)
SEQ ID NO: 151 (ORF 309)	SEQ ID NO: 163 (ORF 729)	SEQ ID NO: 175 (ORF 1882)
SEQ ID NO: 153 (ORF 347)	SEQ ID NO: 165 (ORF 747)	SEQ ID NO: 177 (ORF 1918)
SEQ ID NO: 155 (ORF 540)	SEQ ID NO: 167 (ORF 1202)	SEQ ID NO: 179 (ORF 1983)
SEQ ID NO: 157 (ORF 601)	SEQ ID NO: 169 (ORF 1723)	SEQ ID NO: 181 (ORF 2452)
SEQ ID NO: 159 (ORF 664)	SEQ ID NO: 171 (ORF 1755)	SEQ ID NO: 669 (ORF 1664)

Table III. Open Reading Frames (ORFs) encoding putative polypeptides containing a
5 LPXTG motif

SEQ ID NO: 183 (ORF 433)	SEQ ID NO: 185 (ORF 967)	SEQ ID NO: 187 (ORF 2497)
--------------------------	--------------------------	---------------------------

Table IV. Open Reading Frames (ORFs) encoding other putative surface localized
polypeptides

SEQ ID NO: 189 (ORF 4)	SEQ ID NO: 349 (ORF 741)	SEQ ID NO: 509 (ORF 1682)
SEQ ID NO: 191 (ORF 5)	SEQ ID NO: 351 (ORF 754)	SEQ ID NO: 511 (ORF 1683)
SEQ ID NO: 193 (ORF 11)	SEQ ID NO: 353 (ORF 774)	SEQ ID NO: 513 (ORF 1720)
SEQ ID NO: 195 (ORF 17)	SEQ ID NO: 355 (ORF 783)	SEQ ID NO: 515 (ORF 1725)
SEQ ID NO: 197 (ORF 18)	SEQ ID NO: 357 (ORF 788)	SEQ ID NO: 517 (ORF 1726)
SEQ ID NO: 199 (ORF 20)	SEQ ID NO: 359 (ORF 805)	SEQ ID NO: 519 (ORF 1732)
SEQ ID NO: 201 (ORF 25)	SEQ ID NO: 361 (ORF 814)	SEQ ID NO: 521 (ORF 1736)
SEQ ID NO: 203 (ORF 49)	SEQ ID NO: 363 (ORF 818)	SEQ ID NO: 523 (ORF 1771)
SEQ ID NO: 205 (ORF 64)	SEQ ID NO: 365 (ORF 844)	SEQ ID NO: 525 (ORF 1772)
SEQ ID NO: 207 (ORF 65)	SEQ ID NO: 367 (ORF 848)	SEQ ID NO: 527 (ORF 1775)
SEQ ID NO: 209 (ORF 67)	SEQ ID NO: 369 (ORF 858)	SEQ ID NO: 529 (ORF 1776)
SEQ ID NO: 211 (ORF 69)	SEQ ID NO: 371 (ORF 859)	SEQ ID NO: 531 (ORF 1777)
SEQ ID NO: 213 (ORF 72)	SEQ ID NO: 373 (ORF 860)	SEQ ID NO: 533 (ORF 1783)
SEQ ID NO: 215 (ORF 73)	SEQ ID NO: 375 (ORF 871)	SEQ ID NO: 535 (ORF 1785)
SEQ ID NO: 217 (ORF 75)	SEQ ID NO: 377 (ORF 877)	SEQ ID NO: 537 (ORF 1786)
SEQ ID NO: 219 (ORF 98)	SEQ ID NO: 379 (ORF 896)	SEQ ID NO: 539 (ORF 1814)
SEQ ID NO: 221 (ORF 99)	SEQ ID NO: 381 (ORF 908)	SEQ ID NO: 541 (ORF 1820)
SEQ ID NO: 223 (ORF 130)	SEQ ID NO: 383 (ORF 909)	SEQ ID NO: 543 (ORF 1828)
SEQ ID NO: 225 (ORF 133)	SEQ ID NO: 385 (ORF 910)	SEQ ID NO: 545 (ORF 1833)
SEQ ID NO: 227 (ORF 141)	SEQ ID NO: 387 (ORF 920)	SEQ ID NO: 547 (ORF 1834)
SEQ ID NO: 229 (ORF 151)	SEQ ID NO: 389 (ORF 921)	SEQ ID NO: 549 (ORF 1839)
SEQ ID NO: 231 (ORF 165)	SEQ ID NO: 391 (ORF 926)	SEQ ID NO: 551 (ORF 1873)
SEQ ID NO: 233 (ORF 172)	SEQ ID NO: 393 (ORF 928)	SEQ ID NO: 553 (ORF 1875)
SEQ ID NO: 235 (ORF 184)	SEQ ID NO: 395 (ORF 929)	SEQ ID NO: 555 (ORF 1876)
SEQ ID NO: 237 (ORF 189)	SEQ ID NO: 397 (ORF 933)	SEQ ID NO: 557 (ORF 1888)
SEQ ID NO: 239 (ORF 199)	SEQ ID NO: 399 (ORF 952)	SEQ ID NO: 559 (ORF 1909)
SEQ ID NO: 241 (ORF 209)	SEQ ID NO: 401 (ORF 961)	SEQ ID NO: 561 (ORF 1917)
SEQ ID NO: 243 (ORF 218)	SEQ ID NO: 403 (ORF 975)	SEQ ID NO: 563 (ORF 1931)
SEQ ID NO: 245 (ORF 220)	SEQ ID NO: 405 (ORF 983)	SEQ ID NO: 565 (ORF 1970)
SEQ ID NO: 247 (ORF 223)	SEQ ID NO: 407 (ORF 991)	SEQ ID NO: 567 (ORF 1972)
SEQ ID NO: 249 (ORF 227)	SEQ ID NO: 409 (ORF 1015)	SEQ ID NO: 569 (ORF 1979)

WO 02/083859

PCT/US02/11610

SEQ ID NO: 251 (ORF 241)	SEQ ID NO: 411 (ORF 1018)	SEQ ID NO: 571 (ORF 1987)
SEQ ID NO: 253 (ORF 252)	SEQ ID NO: 413 (ORF 1020)	SEQ ID NO: 573 (ORF 1993)
SEQ ID NO: 255 (ORF 264)	SEQ ID NO: 415 (ORF 1021)	SEQ ID NO: 575 (ORF 2013)
SEQ ID NO: 257 (ORF 265)	SEQ ID NO: 417 (ORF 1026)	SEQ ID NO: 577 (ORF 2014)
SEQ ID NO: 259 (ORF 291)	SEQ ID NO: 419 (ORF 1058)	SEQ ID NO: 579 (ORF 2015)
SEQ ID NO: 261 (ORF 292)	SEQ ID NO: 421 (ORF 1110)	SEQ ID NO: 581 (ORF 2020)
SEQ ID NO: 263 (ORF 306)	SEQ ID NO: 423 (ORF 1132)	SEQ ID NO: 583 (ORF 2023)
SEQ ID NO: 265 (ORF 307)	SEQ ID NO: 425 (ORF 1152)	SEQ ID NO: 585 (ORF 2046)
SEQ ID NO: 267 (ORF 313)	SEQ ID NO: 427 (ORF 1156)	SEQ ID NO: 587 (ORF 2048)
SEQ ID NO: 269 (ORF 350)	SEQ ID NO: 429 (ORF 1188)	SEQ ID NO: 589 (ORF 2050)
SEQ ID NO: 271 (ORF 352)	SEQ ID NO: 431 (ORF 1200)	SEQ ID NO: 591 (ORF 2069)
SEQ ID NO: 283 (ORF 353)	SEQ ID NO: 433 (ORF 1203)	SEQ ID NO: 593 (ORF 2070)
SEQ ID NO: 275 (ORF 368)	SEQ ID NO: 435 (ORF 1205)	SEQ ID NO: 595 (ORF 2091)
SEQ ID NO: 277 (ORF 401)	SEQ ID NO: 437 (ORF 1210)	SEQ ID NO: 597 (ORF 2148)
SEQ ID NO: 279 (ORF 405)	SEQ ID NO: 439 (ORF 1216)	SEQ ID NO: 599 (ORF 2170)
SEQ ID NO: 281 (ORF 421)	SEQ ID NO: 441 (ORF 1228)	SEQ ID NO: 601 (ORF 2201)
SEQ ID NO: 283 (ORF 491)	SEQ ID NO: 443 (ORF 1231)	SEQ ID NO: 603 (ORF 2222)
SEQ ID NO: 285 (ORF 510)	SEQ ID NO: 445 (ORF 1265)	SEQ ID NO: 605 (ORF 2231)
SEQ ID NO: 287 (ORF 511)	SEQ ID NO: 447 (ORF 1267)	SEQ ID NO: 607 (ORF 2236)
SEQ ID NO: 289 (ORF 519)	SEQ ID NO: 449 (ORF 1269)	SEQ ID NO: 609 (ORF 2240)
SEQ ID NO: 291 (ORF 523)	SEQ ID NO: 451 (ORF 1272)	SEQ ID NO: 611 (ORF 2245)
SEQ ID NO: 293 (ORF 535)	SEQ ID NO: 453 (ORF 1275)	SEQ ID NO: 613 (ORF 2247)
SEQ ID NO: 295 (ORF 551)	SEQ ID NO: 455 (ORF 1292)	SEQ ID NO: 615 (ORF 2250)
SEQ ID NO: 297 (ORF 567)	SEQ ID NO: 457 (ORF 1300)	SEQ ID NO: 617 (ORF 2258)
SEQ ID NO: 299 (ORF 570)	SEQ ID NO: 459 (ORF 1310)	SEQ ID NO: 619 (ORF 2266)
SEQ ID NO: 301 (ORF 594)	SEQ ID NO: 461 (ORF 1311)	SEQ ID NO: 621 (ORF 2273)
SEQ ID NO: 303 (ORF 597)	SEQ ID NO: 463 (ORF 1318)	SEQ ID NO: 623 (ORF 2289)
SEQ ID NO: 305 (ORF 602)	SEQ ID NO: 465 (ORF 1321)	SEQ ID NO: 625 (ORF 2291)
SEQ ID NO: 307 (ORF 613)	SEQ ID NO: 467 (ORF 1362)	SEQ ID NO: 627 (ORF 2300)
SEQ ID NO: 309 (ORF 627)	SEQ ID NO: 469 (ORF 1395)	SEQ ID NO: 629 (ORF 2319)
SEQ ID NO: 311 (ORF 639)	SEQ ID NO: 471 (ORF 1497)	SEQ ID NO: 631 (ORF 2342)
SEQ ID NO: 313 (ORF 644)	SEQ ID NO: 473 (ORF 1500)	SEQ ID NO: 633 (ORF 2391)
SEQ ID NO: 315 (ORF 650)	SEQ ID NO: 475 (ORF 1512)	SEQ ID NO: 635 (ORF 2398)
SEQ ID NO: 317 (ORF 653)	SEQ ID NO: 477 (ORF 1513)	SEQ ID NO: 637 (ORF 2399)
SEQ ID NO: 319 (ORF 665)	SEQ ID NO: 479 (ORF 1525)	SEQ ID NO: 639 (ORF 2411)
SEQ ID NO: 321 (ORF 670)	SEQ ID NO: 481 (ORF 1527)	SEQ ID NO: 641 (ORF 2414)
SEQ ID NO: 323 (ORF 671)	SEQ ID NO: 483 (ORF 1548)	SEQ ID NO: 643 (ORF 2428)
SEQ ID NO: 325 (ORF 672)	SEQ ID NO: 485 (ORF 1573)	SEQ ID NO: 645 (ORF 2429)
SEQ ID NO: 327 (ORF 674)	SEQ ID NO: 487 (ORF 1585)	SEQ ID NO: 647 (ORF 2437)
SEQ ID NO: 329 (ORF 676)	SEQ ID NO: 489 (ORF 1586)	SEQ ID NO: 649 (ORF 2457)
SEQ ID NO: 331 (ORF 688)	SEQ ID NO: 491 (ORF 1593)	SEQ ID NO: 651 (ORF 2458)
SEQ ID NO: 333 (ORF 699)	SEQ ID NO: 493 (ORF 1608)	SEQ ID NO: 653 (ORF 2473)
SEQ ID NO: 335 (ORF 702)	SEQ ID NO: 495 (ORF 1661)	SEQ ID NO: 655 (ORF 2482)
SEQ ID NO: 337 (ORF 705)	SEQ ID NO: 497 (ORF 1667)	SEQ ID NO: 657 (ORF 2488)
SEQ ID NO: 339 (ORF 706)	SEQ ID NO: 499 (ORF 1671)	SEQ ID NO: 659 (ORF 2508)
SEQ ID NO: 341 (ORF 721)	SEQ ID NO: 501 (ORF 1672)	SEQ ID NO: 661 (ORF 2521)
SEQ ID NO: 343 (ORF 731)	SEQ ID NO: 503 (ORF 1678)	SEQ ID NO: 663 (ORF 2534)
SEQ ID NO: 345 (ORF 733)	SEQ ID NO: 505 (ORF 1680)	SEQ ID NO: 665 (ORF 2562)
SEQ ID NO: 347 (ORF 737)	SEQ ID NO: 507 (ORF 1681)	SEQ ID NO: 667 (ORF 2583)

WO 02/083859

PCT/US02/11610

Genomic Approach

The availability of complete bacterial genome sequences is currently playing an important role in the identification of immunogenic composition candidates through genomics, transcriptional profiling, and proteomics, coupled with the information processing capabilities of bioinformatics (39-41, 53, 60, 65).

The genomic approach began by identifying open reading frames (ORFs) in an unannotated sequence of *Streptococcus pyogenes* downloaded from the website of the University of Oklahoma. This genomic sequence was reported as being submitted to GenBank and assigned accession number AE004092. Strain M1 GAS was reported as being submitted to the ATCC and given accession number ATCC 700294.

An ORF is defined herein as having one of three potential start site codons, ATG, GTG, or TTG, and one of three potential stop codons, TAA, TAG, or TGA. Using this definition of an ORF, the *Streptococcus pyogenes* genome was analyzed to identify ORFs using three ORF finder algorithms, GLIMMER (59), GeneMark (34), and an algorithm developed by inventor's assignee. There were 736 ORFs commonly identified by all three algorithms. The difference in results between the different ORF finders is primarily due to the particular start codons used by each program, however, Glimmer also incorporates some evaluation for a Shine-Dalgarno box. All ORFs with common stop codons were given the same ORF designation and were treated as if they were the same ORF.

In order to evaluate the accuracy of the ORFs determined, a discrete mathematical cosine function, known in the art as a discrete cosine transformation (DiCTion), was employed to assign a score for each ORF. An ORF with a DiCTion score >1.5 was considered to have a high probability of encoding a protein product. The minimum length of an ORF predicted by the three ORF finding algorithms was set to 225 nucleotides (including stop codon) which would encode a protein of 74 amino acids.

As a final search for remnants of ORFs, all noncoding regions >75 nucleotides were searched against public protein databases using tBLASTn to identify regions of genes that contained frameshifts (42) or fragments of genes that might have a role in causing antigenic variation (21). These remnant ORFs were added to the ORF hits.

WO 02/083859

PCT/US02/11610

A graphical analysis program developed by inventor's assignee was used to show all six reading frames and the location of the predicted ORFs relative to the genomic sequence. This helped to eliminate ORFs that had large overlaps with other ORFs, although there are known cases of ORFs being totally embedded within other ORFs (25, 33).

5 The initial annotation of these *Streptococcus pyogenes* ORFs was performed using the BLAST v. 2.0 Gapped search algorithm, BLASTp, to identify homologous sequences. A cutoff "e" value of anything $< e^{-10}$ was considered significant. Other search algorithms, including FASTA and PSI-BLAST, were also used. The non-redundant protein sequence databases used for the homology searches included GenBank, SWISS-PROT, PIR, and
10 TREMBL database sequences updated daily. ORFs with a BLASTp result of $> e^{-10}$ were considered to be unique to *Streptococcus pyogenes*.

Currently, about 60% of all ORFs within a bacterial genome have some match with a protein whose function has been determined. That leaves about 40% of genomic ORFs still uncharacterized. A keyword search of the entire Blast results was carried out using known or
15 suspected candidate target genes as well as words that identified the location of a protein or function. In addition, a keyword search was performed of all MEDLINE references associated with the initial Blast results to look for additional information regarding the ORFs. The keyword search included, for example, the following search terms: adhesin(ion); fibronectin; fibrinogen; collagen; transporter; exporter; extracellular; transferase; surface; and
20 binding. Blast analysis of the ORFs resulted in 1005 ORFs listed as unclassified, 284 ORFs appeared to be specific to *Streptococcus pyogenes* since they produced Blast similarity only with proteins from this organism, and 676 ORFs were associated with a Medline reference.

For DNA analysis, the %G+C content within each gene was identified. The %G+C content of an ORF was calculated as the (G+C) content of the third nucleotide position of all
25 the codons within an ORF. The value reported was the difference of this value from the arithmetic mean of such values obtained for all ORFs found in the organism. An absolute value ≥ 8 was considered important for further analysis, as these ORFs may have arisen from horizontal transfer as has been shown in the case of *cag* pathogenicity island from *H. pylori* (2), a pattern in keeping with many other pathogenicity islands (22). ORFs that were
30 significantly different in their G+C content totaled 289. These ORFs were further examined for similarity to virulence factors acquired from another organism by horizontal transfer.

WO 02/083859

PCT/US02/11610

Several parameters were used to determine partitioning of the predicted proteins. Proteins destined for translocation across the cytoplasmic membrane encode a leader signal (also known as a signal sequence) composed of a central hydrophobic region flanked at the N-terminus by positively charged residues (56). The program SignalP was used to identify signal peptides and their cleavage sites (46). During expression, the signal peptide is cleaved to produce a mature peptide. In addition, to predict protein localization in bacteria, the software PSORT was used (44). PSORT uses a neural net algorithm to predict localization of proteins to the cytoplasm, periplasm, and/or cytoplasmic membrane for Gram-positive bacteria as well as outer membrane for Gram-negative bacteria. PSORT identified 40 ORFs predicted to be surface exposed (Table V).

Table V. Open Reading Frames (ORFs) encoding putative extracellular proteins

68	705	1202	1664	1723	2020	2385
165	729	1310	1667	1777	2046	2414
252	788	1358	1678	1909	2170	2437
510	1058	1362	1680	1972	2236	2601
601	1132	1573	1681	1975	2250	
668	1200	1638	1683	2014	2300	

In addition, transmembrane (TM) domains of proteins were analyzed using the software program TopPred2 (10). This program predicts regions of a protein that are hydrophobic that may potentially span the lipid bilayer of the membrane. Analysis by TopPred2 for hydrophobic regions of a protein that may potentially span the lipid bilayer of the membrane identified 48 ORFs that encoded putative proteins with three or more transmembrane spanning domains (Table VI) and are thus considered to be membrane bound.

Table VI. Open Reading Frames (ORFs) encoding putative proteins with three or greater transmembrane regions

8	307	594	752	1222	1598	2069
73	312	613	844	1266	1657	2091
80	395	650	925	1317	1708	2227
95	508	672	975	1488	1726	2283
141	551	706	1018	1496	1779	2424
265	567	708	1152	1513	1999	2562
306	593	731	1156	1596	2002	

The Hidden Markov Model (HMM) Pfam database of multiple alignments of protein domains or conserved protein regions (61) was used to identify *Streptococcus pyogenes* proteins that may belong to an existing protein family. Keyword searching of this output was used to identify proteins that might have been missed by the Blast search criteria. HMM

WO 02/083859

PCT/US02/11610

models were also developed by inventor's assignee. A computer algorithm, HMM Lipo, was developed to predict lipoproteins using 132 biologically characterized non-*Streptococcus pyogenes* bacterial lipoproteins from over 30 organisms. This training set was generated from experimentally proven prokaryotic lipoproteins. HMM Lipo identified 30 ORFs that are putative lipoproteins (Table VII).

Table VII. Open Reading Frames (ORFs) encoding putative lipoproteins

68	601	747	1659	1789	1983
309	678	1157	1664	1818	2417
347	685	1202	1723	1878	2452
540	704	1284	1755	1882	2459
554	729	1495	1788	1918	2601

In addition, 15 ORFs were predicted to have a LPXTG motif and were classified as proteins that might be targeted by sortase (Table VIII).

Table VIII. Open Reading Frames (ORFs) encoding putative proteins containing the LPXTG motif

433	1218	1854	2450
608	1316	2019	2477
967	1330	2434	2497
1191	1698	2446	

SEQ ID NOS: 669-674 contain the nucleotide and amino acid sequences of the proteins Grab (ORF 608), M protein (ORF 2434), and ScpA (ORF 2446), respectively.

Furthermore, using about 70 known prokaryotic proteins containing the LPXTG cell wall sorting signal, a HMM (15) was developed to predict cell wall proteins that are anchored to the peptidoglycan layer (38, 45). The model used not only the LPXTG sequence, but also included two features of the downstream sequence, the hydrophobic transmembrane domain and the positively charged carboxy terminus. There were 5 proteins identified as potentially binding to the peptidoglycan layer in a non-covalent manner independently of the sortase (Table IX).

Table IX. Open Reading Frames (ORFs) encoding putative peptidoglycan binding proteins

898	1569	1675	2266	2311
-----	------	------	------	------

The proteins encoded by the identified ORFs were also evaluated for other characteristics. A tandem repeat finder (5) identified ORFs containing repeated DNA

WO 02/083859

PCT/US02/11610

sequences such as those found in MSCRAMMs (20) and phase variable surface proteins of *Neisseria meningitidis* (51). There were 23 ORFs found to encode proteins containing such repeat regions (Table X).

Table X. Open Reading Frames (ORFs) encoding putative proteins containing repeat regions

218	433	1149	1783	2422	2513
265	555	1562	1972	2434	2590
336	699	1583	2137	2437	2618
431	783	1683	2231	2477	

5

In addition, proteins that contain the Arg-Gly-Asp (RGD) attachment motif, together with integrins that serve as their receptor, constitute a major recognition system for cell adhesion. RGD recognition is one mechanism used by microbes to gain entry into eukaryotic tissues (29, 63). There were 65 ORFs identified that encoded RGD-containing proteins

10 (Table XI).

Table XI. Open Reading Frames (ORFs) encoding putative proteins containing the RGD motif.

18	544	885	1149	1504	1957	2379
201	626	889	1161	1626	2042	2414
209	641	899	1200	1643	2054	2446
302	654	967	1274	1657	2082	2558
344	667	968	1313	1675	2148	2570
350	668	1010	1316	1773	2205	
396	695	1027	1373	1779	2247	
397	726	1074	1401	1885	2253	
413	787	1108	1416	1891	2287	
526	829	1110	1431	1901	2335	

A graphical representation of the results of the genomic analysis and ORF identification is depicted in Fig. 1.

15

Proteomic Approach

As stated above, a proteomic approach was also taken to identify surface localized proteins of *Streptococcus pyogenes*.

In order to identify only those proteins localized to the surface of the cell, care was taken during the preparation and digestion of the *Streptococcus pyogenes* cells with trypsin.

20

Samples of the cells were taken just prior to the addition of trypsin and at the completion of the digestion, and were examined for cell integrity by viable counts and LV-SEM. Following

WO 02/083859

PCT/US02/11610

digestion, untreated cells clearly aggregated and adhered to the side of the tube, while the treated cells formed an even cell suspension. Viable counts showed no significant difference between samples and in fact were slightly higher in the treated cells due to the aggregation of the untreated sample. LV-SEM confirmed these results (Fig. 2). Digested cells were evenly and individually distributed over the cover slip, while the untreated sample displayed large clumps of bacteria. Topographical examination at high magnification of untreated bacterial cells displayed large quantities of surface material typical of *Streptococcus pyogenes*. However, individual cells in the trypsin digested sample showed the reduction of all observable surface protein as the cells appeared bald and devoid of any surface material. Fig. 3 depicts LV-SEMS of *Streptococcus pyogenes* before (left panel, Panel A) and after (right panel, Panel B) digestion with trypsin. The cells before digestion with trypsin (Panel A) are larger and display surface material. The LV-SEM of the cells after digestion (Panel B) are smaller and appear devoid of any surface protein.

In order to identify the peptide components of the complex surface digest mixture, an analytical technique was used to separate and sequence multiple peptides with high sensitivity over a large concentration range. Tandem mass spectrometry (MS/MS) has been shown to be a powerful approach to analyze proteins from both gels and in solution (17). MS/MS first uses a mass analyzer to separate a peptide ion from a mixture of ions, then uses a second step or mass analyzer to activate and dissociate the ion of interest. This process, known as collision induced dissociation (CID), causes the peptide to fragment at the peptide bonds between the amino acids, and therefore, the fragmentation pattern of a peptide is used to determine its amino acid sequence.

In addition, the SEQUEST computer algorithm was used to search the experimental fragmentation spectrum directly against protein or translated nucleotide sequence databases. For peptides above roughly 800-900 Da in size, a single spectrum can uniquely identify a protein.

To sequence multiple peptides from a complex mixture, a reversed phase chromatography system was coupled to an electrospray ion trap mass spectrometer. In this system, it is known that high sensitivity (down to sub-femtomole levels) can be attained by minimizing both flow rate and column diameter to concentrate the elution volume and direct as much of the column effluent as possible into the orifice of the mass spectrometer detector. Initial experiments separated peptides using a reversed phase gradient of 1% acetonitrile/min.

WO 02/083859

PCT/US02/11610

In order to increase chromatographic separation, longer gradients, down to 0.28% acetonitrile/min., and slower flow rates (50 nL/min.) were later employed. To maximize the coverage of proteins present in the sample, the data-dependent acquisition feature of the ion trap was employed.

5 Dynamic exclusion was used to prevent reacquisition of tandem mass spectra of ions once a spectrum had been acquired for a particular m/z value. The isotopic exclusion function excluded the ion associated with the ^{13}C isotope of peptides from the list of ions slated for MS/MS. A 3-u mass width window was selected for this purpose. Using these data-dependent features dramatically increased the number of peptide ions that were selected
10 for CID analysis.

The LC-MS/MS data acquisition conditions described above typically resulted in fragmentation data for more than 2000 peptide ions for each run. Using the SEQUEST algorithm, this data was searched against a composite protein sequence database containing the translated ORFs from *Streptococcus pyogenes* combined with the non-redundant protein
15 sequence database OWL. SEQUEST search conditions used modified trypsin selectivity and allowed a differential search of +16 Da on methionine to account for methionine oxidation. Candidate matches identified by SEQUEST were confirmed using the following manual procedure. Those matches with Xcorr values greater than 2.5 (a measure of the similarity of the experimental ms/ms data to that generated from the sequence database) and delCn values
20 greater than 0.1 (delCn measures the normalized difference between the Xcorr values of the first and second matches) were chosen for further analysis. The fragmentation spectra from good matches were checked for reasonable signal/noise, and the list of matched ions was examined for reasonable continuity. Some matches that were not acceptable alone were included if other confirmatory ms/ms data was generated by the same sample. The ORFs
25 obtained by this proteomic approach are presented in Table XII.

Table XII. Open Reading Frames (ORFs) identified by tryptic digestion

66	678	1224	1638	1878	2214	2459
102	704	1234	1650	1902	2330	2477
145	743	1237	1654	1943	2354	2586
232	825	1238	1659	1975	2377	2593
238	850	1253	1698	2019	2379	2601
436	934	1284	1788	2064	2387	
516	993	1316	1794	2086	2417	
554	1036	1330	1816	2106	2420	
589	1140	1358	1818	2116	2422	

WO 02/083859

PCT/US02/11610

608	1157	1487	1819	2120	2434
661	1191	1495	1850	2123	2446
668	1218	1557	1854	2202	2450

Several of the ORFs identified were cloned and expressed. Mouse antisera, generated to the purified proteins, were first analyzed for reactivity by ELISA using the same preparation used for the mouse immunization as the coating antigen. To quantitate protein expression on the surface of *Streptococcus pyogenes*, these sera were then used in whole cell ELISAs. To qualify the protein expression of the specific proteins, whole *Streptococcus pyogenes* cells were labeled by immunogold and viewed by LV-SEM.

For some of the identified ORFs, the encoded proteins were observed to be expressed in a manner that was dependent upon phase of growth (mid-log versus stationary). Examples of this class are ORF 218 (Fig. 4), ORF 554 (Fig. 5), and ORF 1191 (Fig. 6). In some cases, expression level was higher in the mid-log growth, while others were greater in the stationary cells. Proteins encoded by other ORFs were expressed at low levels regardless of growth phase (ORFs 2064, 2601, and 1316) (shown in Figs. 7-9, respectively), while others were expressed at high levels independent of growth phase (ORF 1224) (Fig. 10). As a positive control, anti-C5a peptidase sera was used as it is known to be expressed and localized to the cell wall of *Streptococcus pyogenes*. All antisera showed an increase in reactivity over the respective pre-immune control sera.

Combination of Genomic and Proteomic Approaches

The ORFs identified in Tables V-XII were then categorized into one of four groups:

ORFs encoding surface localized proteins identified by proteomics (Table I); ORFs encoding putative lipoproteins (Table II); ORFs encoding putative polypeptides containing a LPXTG motif (Table III); and ORFs encoding other putative surface localized polypeptides (Table IV). Tables I-IV are provided supra. It should be apparent that the ORFs contained in Tables I-IV are non-redundant, i.e., the ORFs listed in Tables I-IV each appear once though many possess characteristics that match another table.

The nucleotide sequences of Table I encode polypeptides that have been identified by the proteomic approach as being surface localized, *Streptococcus pyogenes* proteins. The nucleotide sequences of Tables II-IV encode putative polypeptides that have been identified by the described genomic approaches as being surface localized, *Streptococcus pyogenes*

WO 02/083859

PCT/US02/11610

proteins. Specifically, the nucleotide sequences of Table II encode putative lipoproteins, the nucleotide sequences of Table III encode putative proteins having an LPXTG cell wall sorting signal, and the nucleotide sequences of Table IV encode putative surface localized proteins that include at least one of several criteria, as described herein, including similarity to other proteins for which a function and cellular location had been previously identified, match with a protein family (e.g., Pfam), and a combined analysis of the membrane spanning domains, Psort and sigP values, and the predicted molecular weight of the protein.

Each of odd numbered SEQ ID NOS: 1-667 encodes an amino acid sequence that is numbered consecutively after the nucleotide sequence. Thus, for example, the nucleotide sequence of SEQ ID NO: 1 encodes the amino acid sequence of SEQ ID NO: 2, and the nucleotide sequence of SEQ ID NO: 3 encodes the amino acid sequence of SEQ ID NO: 4, etc.

Polypeptides

The invention provides *Streptococcus pyogenes* polypeptides that are surface localized. Specifically, the polypeptides of the invention include isolated polypeptides that comprise an amino acid sequence of any of even numbered SEQ ID NOS: 2-668, i.e., SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, 268, 270, 272, 274, 276, 278, 280, 282, 284, 286, 288, 290, 292, 294, 296, 298, 300, 302, 304, 306, 308, 310, 312, 314, 316, 318, 320, 322, 324, 326, 328, 330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 370, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 392, 394, 396, 398, 400, 402, 404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426, 428, 430, 432, 434, 436, 438, 440, 442, 444, 446, 448, 450, 452, 454, 456, 458, 460, 462, 464, 466, 468, 470, 472, 474, 476, 478, 480, 482, 484, 486, 488, 490, 492, 494, 496, 498, 500, 502, 504, 506, 508, 510, 512, 514, 516, 518, 520, 522, 524, 526, 528, 530, 532, 534, 536, 538, 540, 542, 544, 546, 548, 550, 552, 554, 556, 558, 560, 562, 564, 566, 568, 570, 572, 574, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 598, 600, 602,

WO 02/083859

PCT/US02/11610

604, 606, 608, 610, 612, 614, 616, 618, 620, 622, 624, 626, 628, 630, 632, 634, 636, 638, 640, 642, 644, 646, 648, 650, 652, 654, 656, 658, 660, 662, 664, 666, or 668.

The polypeptides of the invention also include isolated polypeptides that consist essentially of the aforementioned amino acid sequences and isolated polypeptides that consist of the aforementioned amino acid sequences. The term "isolated" means altered by the hand of man from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polypeptide or a polynucleotide naturally present in a living animal is not "isolated," but the same polypeptide or polynucleotide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. As used herein, the term "isolated" contemplates a polypeptide (or other component) that is isolated from its natural source and/or prepared using recombinant technology.

A polypeptide sequence of the invention may be identical to the reference sequence of even numbered SEQ ID NOS: 2-668, that is, 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the % identity is less than 100%. Such alterations include at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. The alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference amino acid sequence or in one or more contiguous groups within the reference amino acid sequence.

Thus, the invention also provides isolated polypeptides having sequence identity to the amino acid sequences contained in the Sequence Listing (i.e., even numbered SEQ ID NOS: 2-668). Depending on the particular sequence, the degree of sequence identity is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 97%, 99% or more). These homologous proteins include mutants and allelic variants.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by

WO 02/083859

PCT/US02/11610

known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al. 1984), BLASTP, BLASTN, and FASTA (Altschul, S. F., et al., 1990. The BLASTX program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894; Altschul, S., et al., 1990). The well known Smith Waterman algorithm may also be used to determine identity.

For example, the number of amino acid alterations for a given % identity can be determined by multiplying the total number of amino acids in one of even numbered SEQ ID NOS: 2-668 by the numerical percent of the respective percent identity (divided by 100) and then subtracting that product from said total number of amino acids in the one of even numbered SEQ ID NOS: 2-668, or:

$$n_a \leq x_a - (x_a \cdot y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in the one of SEQ ID NOS: 2-668, and y is, for instance, 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

The present invention contemplates isolated polypeptides that are substantially conserved across strains of β -hemolytic streptococci. Further, isolated polypeptides that are substantially conserved across strains of β -hemolytic streptococci and that are effective in preventing or ameliorating a β -hemolytic streptococcal colonization or infection in a susceptible subject are also contemplated by the present invention. As used herein, the term

WO 02/083859

PCT/US02/11610

"conserved" refers to, for example, the number of amino acids that do not undergo insertions, substitution and/or deletions as a percentage of the total number of amino acids in a protein. For example, if a protein is 55% conserved and has, for example, 263 amino acids, then there are 144 amino acid positions in the protein at which amino acids do not undergo substitution.

- 5 Likewise, if a protein is 90% conserved and has, for example, about 280 amino acids, then there are 28 amino acid positions at which amino acids may undergo substitution and 252 (i.e., 280 minus 28) amino acid positions at which the amino acids do not undergo substitution. According to an embodiment of the present invention, the isolated polypeptide is preferably at least about 80% conserved across the strains of β -hemolytic streptococci, 10 more preferably at least about 85% conserved across the strains, even more preferably at least about 90% conserved across the strains, and most preferably at least about 95% conserved across the strains, without limitation.

- Modifications and changes can be made in the structure of the polypeptides of even numbered SEQ ID NOS: 2-668 and still obtain a molecule having β -hemolytic streptococci 15 and/or *Streptococcus pyogenes* activity and/or antigenicity. For example, certain amino acids can be substituted for other amino acids in a sequence without appreciable loss of activity and/or antigenicity. Because it is the interactive capacity and nature of a polypeptide that defines that polypeptide's biological functional activity, certain amino acid sequence substitutions can be made in a polypeptide sequence (or, of course, its underlying DNA 20 coding sequence) and nevertheless obtain a polypeptide with like properties.

- The invention includes any isolated polypeptide which is a biological equivalent that provides the desired reactivity as described herein. The term "desired reactivity" refers to reactivity that would be recognized by a person skilled in the art as being a useful result for the purposes of the invention. Examples of desired reactivity are described herein, including 25 without limitation, desired levels of protection, desired antibody titers, desired opsonophagocytic activity and/or desired cross-reactivity, such as would be recognized by a person skilled in the art as being useful for the purposes of the present invention. The desired opsonophagocytic activity is indicated by a percent killing of bacteria as measured by decrease in colony forming units (CFU) in OPA versus a negative control. Without being 30 limited thereto, the desired opsonophagocytic activity is preferably at least about 15%, more preferably at least about 20%, even more preferably at least about 40%, even more preferably at least about 50% and most preferably at least about 60%.

WO 02/083859

PCT/US02/11610

The invention includes polypeptides that are variants of the polypeptides comprising an amino acid sequence of SEQ ID NOS: 2-668. "Variant" as the term is used herein, includes a polypeptide that differs from a reference polypeptide, but retains essential properties. Generally, differences are limited so that the sequences of the reference
5 polypeptide and the variant are closely similar overall and, in many regions, identical (i.e., biologically equivalent). A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, or deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polypeptide may be a naturally occurring such as an allelic variant, or it
10 may be a variant that is not known to occur naturally. Non-naturally occurring variants of polypeptides may be made by direct synthesis or by mutagenesis techniques.

In making such changes, the hydropathic index of amino acids can be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a polypeptide is generally understood in the art (Kyte & Doolittle, 1982). It is
15 known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still result in a polypeptide with similar biological activity. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. Those indices are listed in parentheses after each amino acid as follows: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8);
20 cysteine/cysteine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is believed that the relative hydropathic character of the amino acid residue determines the secondary and tertiary structure of the resultant polypeptide, which in turn
25 defines the interaction of the polypeptide with other molecules, such as enzymes, substrates, receptors, antibodies, antigens, and the like. It is known in the art that an amino acid can be substituted by another amino acid having a similar hydropathic index and still obtain a functionally equivalent polypeptide. In such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those which are within ± 1 are particularly
30 preferred, and those within ± 0.5 are even more particularly preferred.

Substitution of like amino acids can also be made on the basis of hydrophilicity, particularly where the biological functional equivalent polypeptide or peptide thereby created

WO 02/083859

PCT/US02/11610

is intended for use in immunological embodiments. U.S. Patent Number 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a polypeptide, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, i.e., with a biological property of the polypeptide.

- 5 As detailed in U.S. Patent Number 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); proline (-0.5 $\pm$ 1); threonine (-0.4); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and
- 10 tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent and in particular, an immunologically equivalent, polypeptide. In such changes, the substitution of amino acids whose hydrophilicity values are within $\pm$ 2 is preferred, those which are within $\pm$ 1 are particularly preferred, and those within $\pm$ 0.5 are even more particularly preferred.
- 15 As outlined above, amino acid substitutions are generally, therefore, based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions which take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and
- 20 threonine; glutamine and asparagine; and valine, leucine, and isoleucine. As shown in Table XIII below, suitable amino acid substitutions include the following:

TABLE XIII:

Original Residue	Exemplary Residue Substitution
Ala	Gly; Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Ala
His	Asn; Gln
Ile	Leu; Val

WO 02/083859

PCT/US02/11610

Leu	Ile; Val
Lys	Arg
Met	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Thus, the invention includes functional or biological equivalents of the polypeptides of SEQ ID NOS: 2-668 that contain one or more amino acid substitutions.

Biological or functional equivalents of a polypeptide can also be prepared using site-specific mutagenesis. Site-specific mutagenesis is a technique useful in the preparation of second generation polypeptides, or biologically, functionally equivalent polypeptides, derived from the sequences thereof, through specific mutagenesis of the underlying DNA. As noted above, such changes can be desirable where amino acid substitutions are desirable. The technique further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is preferred, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art. As will be appreciated, the technique typically employs a phage vector which can exist in both a single-stranded and double-stranded form. Typically, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector which includes within its sequence a DNA sequence which encodes all or a portion of the *Streptococcus pyogenes* polypeptide sequence selected. An oligonucleotide primer bearing the desired mutated sequence is prepared, for example, by well known techniques (e.g., synthetically). This primer is then annealed to the single-stranded vector, and extended by the use of enzymes, such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the

WO 02/083859

PCT/US02/11610

original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutation. Commercially available kits provide the necessary reagents.

5 The polypeptides and polypeptide antigens of the invention are understood to include any polypeptide comprising substantial sequence similarity, structural similarity, and/or functional similarity to a polypeptide comprising an amino acid sequence of any of SEQ ID NOS: 2-668. In addition, a polypeptide or polypeptide antigen of the invention is not limited to a particular source. Thus, the invention provides for the general detection and isolation of
10 the polypeptides from a variety of sources.

The polypeptides of the invention may advantageously be cleaved into fragments for use in further structural or functional analysis, or in the generation of reagents such as *Streptococcus pyogenes*-related polypeptides and *Streptococcus pyogenes*-specific antibodies. This can be accomplished by treating purified or unpurified polypeptides of the invention
15 with a peptidase such as endoproteinase glu-C (Boehringer, Indianapolis, IN). Treatment with CNBr is another method by which peptide fragments may be produced from natural *Streptococcus pyogenes* polypeptides. Recombinant techniques also can be used to produce specific fragments of a *Streptococcus pyogenes* polypeptide.

In addition, the inventors contemplate that compounds sterically similar to a particular
20 *Streptococcus pyogenes* polypeptide antigen may be formulated to mimic the key portions of the peptide structure, known in the art as peptidomimetics. Mimetics are peptide-containing molecules which mimic elements of protein secondary structure. The underlying rationale behind the use of peptidomimetics is that the peptide backbone of proteins exists chiefly to orient amino acid side chains in such a way as to facilitate molecular interactions, such as
25 those of receptor and ligand.

The invention also includes fusion proteins comprising at least one polypeptide of the invention. "Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. For example, fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof
30 have been described. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example,

WO 02/083859

PCT/US02/11610

improved pharmacokinetic properties (See, for example, EP-A 0232 2621). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected, and purified.

5 The polypeptides of the invention may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains, for example, secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

10 Fragments of the *Streptococcus pyogenes* polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence. The fragment can comprise, for example, at least 7 or more (e.g., 8, 10, 12, 14, 16, 18, 20, or more) contiguous amino acids of an amino acid sequence of any of even numbered SEQ ID NOS: 2-668. Fragments may be "freestanding" or comprised within a larger polypeptide of which they form a part or region,
15 most preferably as a single, continuous region. In one embodiment, the fragments include at least one epitope of the mature polypeptide sequence.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, and polypeptides produced by a combination of these
20 methods. Means for preparing such polypeptides are well understood in the art.

Polynucleotides

The invention also provides isolated polynucleotides comprising a nucleotide sequence that encodes a polypeptide of the invention, and polynucleotides closely related thereto. These polynucleotides include:

- 25 (i) an isolated polynucleotide comprising a nucleotide sequence of any of odd numbered SEQ ID NOS: 1-147 (Table I);
- (ii) an isolated polynucleotide comprising a nucleotide sequence of any of odd numbered SEQ ID NOS: 149-181 (Table II);
- (iii) an isolated polynucleotide comprising a nucleotide sequence of any of odd
30 numbered SEQ ID NOS: 183- 187 (Table III); and

WO 02/083859

PCT/US02/11610

(iv) an isolated polynucleotide comprising a nucleotide sequence of any of odd numbered SEQ ID NOS: 189- 667 (Table IV).

The polynucleotides encoding the polypeptides of the invention may be identical to the nucleotide sequences contained in Tables I-IV or they may have variant sequences which, as a result of the redundancy (degeneracy) of the genetic code, also encode polypeptides of the invention.

Further, the invention provides isolated polynucleotides having sequence identity to the nucleotide sequences of SEQ ID NOS: 1-667. Depending on the particular sequence, the degree of sequence identity is preferably greater than 70% (e.g., 80%, 90%, 95%, 97% 99% or more).

As discussed above, "identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. "Identity" can be readily calculated by known methods. By way of example, a polynucleotide sequence of the present invention may be identical to a reference nucleotide sequence of odd numbered SEQ ID NOS: 1-667, that is be 100% identical, or it may include up to a certain integer number of nucleotide alterations as compared to the reference nucleotide sequence. Such alterations include at least one nucleotide deletion, substitution, including transition and transversion, or insertion. The alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference nucleotide sequence. The number of nucleotide alterations is determined by multiplying the total number of nucleotides in one of odd numbered SEQ ID NOS: 1-667 by the numerical percent of the respective percent identity (divided by 100) and subtracting that product from said total number of nucleotides of the reference nucleotide sequence of any of odd numbered SEQ ID NOS: 1-667.

For example, for a polynucleotide that has at least 70% identity to a nucleotide sequence of one of odd numbered SEQ ID NOS: 1-667, the polynucleotide may include up to n_n nucleic acid alterations over the entire length of the nucleotide sequence of one of odd numbered SEQ ID NOS: 1-667, wherein n_n is calculated by the formula:

$$n_n \leq x_n - (x_n \cdot y),$$

WO 02/083859

PCT/US02/11610

and wherein x_n is the total number of nucleotides of the nucleotide sequence of one of odd numbered SEQ ID NOS: 1-667, y has a value of 0.70, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting such product from x_n . Of course, y may also have a value of 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, etc.

The invention also includes polynucleotides that encode polypeptide variants of the polypeptides comprising an amino acid sequence of SEQ ID NOS: 2-668, in which one or more amino acid residues are substituted, deleted, or added, in any combination while retaining the biological activity of the native polypeptide. "Variant" as the term is used herein, is a polynucleotide that differs from a reference polynucleotide, but retains essential properties. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions, and truncations in the polypeptide encoded by the reference sequence. A variant of a polynucleotide may be naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides may be made by mutagenesis techniques or by direct synthesis.

The invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the Stringency Conditions Table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

TABLE XIV - STRINGENCY CONDITIONS TABLE

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) ¹	Hybridization Temperature and Buffer ²	Wash Temperature and Buffer ³
A	DNA:DNA	> 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	< 50	T _B ; 1xSSC	T _B ; 1xSSC
C	DNA:RNA	> 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	< 50	T _D ; 1xSSC	T _D ; 1xSSC
E	RNA:RNA	> 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC

WO 02/083859

PCT/US02/11610

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) ¹	Hybridization Temperature and Buffer ^H	Wash Temperature and Buffer ^H
F	RNA:RNA	< 50	T _F : 1xSSC	T _F : 1xSSC
G	DNA:DNA	> 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	< 50	T _H : 4xSSC	T _H : 4xSSC
I	DNA:RNA	> 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	< 50	T _J : 4xSSC	T _J : 4xSSC
K	RNA:RNA	> 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	< 50	T _L : 2xSSC	T _L : 2xSSC
M	DNA:DNA	> 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	< 50	T _N : 6xSSC	T _N : 6xSSC
O	DNA:RNA	> 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	< 50	T _P : 6xSSC	T _P : 6xSSC
Q	RNA:RNA	> 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	< 50	T _R : 4xSSC	T _R : 4xSSC

bp¹: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

buffer^H: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

T_B through T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(EC) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(EC) = 81.5 + 16.6(log₁₀[Na⁺]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na⁺] is the concentration of sodium ions in the hybridization buffer ([Na⁺] for 1xSSC = 0.165 M).

WO 02/083859

PCT/US02/11610

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

The invention also provides polynucleotides that are fully complementary to these polynucleotides and also provides antisense sequences. The antisense sequences of the invention, also referred to as antisense oligonucleotides, include both internally generated and externally administered sequences that block expression of polynucleotides encoding the polypeptides of the invention. The antisense sequences of the invention comprise, for example, about 15-20 base pairs. The antisense sequences can be designed, for example, to inhibit transcription by preventing promoter binding to an upstream nontranslated sequence or by preventing translation of a transcript encoding a polypeptide of the invention by preventing the ribosome from binding.

The polynucleotides of the invention are prepared in many ways (e.g., by chemical synthesis, from DNA libraries, from the organism itself) and can take various forms (e.g., single-stranded, double-stranded, vectors, probes, primers). The term "polynucleotide" includes DNA and RNA, and also their analogs, such as those containing modified backbones.

When the polynucleotides of the invention are used for the recombinant production of polypeptides, the polynucleotide may include the coding sequence of the mature polypeptide or a fragment thereof, by itself, the coding sequence of the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, pro-, or prepro- protein sequence, or other fusion protein portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be linked to the coding sequence. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites, and sequences that stabilize mRNA.

Expression Systems and Vectors

For recombinant production, host cells are genetically engineered to incorporate expression systems, portions thereof, or polynucleotides of the invention. Introduction of

WO 02/083859

PCT/US02/11610

polynucleotides into host cells are effected, for example, by methods described in many standard laboratory manuals, such as Davis et al., BASIC METHODS IN MOLECULAR BIOLOGY (1986) and Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989),
5 such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, ultrasound, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction, or infection.

Representative examples of suitable hosts include bacterial cells (e.g., streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells), yeast cells (e.g., *Pichia*,
10 *Saccharomyces*), mammalian cells (e.g., vero, Chinese hamster ovary, chick embryo fibroblasts, BHK cells, human SW13 cells), and insect cells (e.g., Sf9, Sf21).

The recombinantly produced polypeptides are recovered and purified from recombinant cell cultures by well-known methods, including high performance liquid chromatography, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation
15 exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography.

A great variety of expression systems are used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial
20 plasmids, attenuated bacteria such as *Salmonella* (U.S. Patent Number 4,837,151) from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as vaccinia and other poxviruses, sindbis, adenovirus, baculoviruses, papova viruses, such as SV40, fowl pox viruses, pseudorabies viruses and retroviruses, alphaviruses such as Venezuelan equine encephalitis virus (U.S.
25 Patent Number 5,643,576), nonsegmented negative-stranded RNA viruses such as vesicular stomatitis virus (U.S. Patent Number 6,168,943), and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems should include control regions that regulate as well as engender expression, such as promoters and other regulatory elements (such as a
30 polyadenylation signal). Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-

WO 02/083859

PCT/US02/11610

known and routine techniques, such as, for example, those set forth in Sambrook et al.,
MOLECULAR CLONING, A LABORATORY MANUAL (supra).

The invention also provides vectors (e.g., expression vectors, sequencing vectors,
cloning vectors) which comprise a polynucleotide or polynucleotides of the invention, host
5 cells which are genetically engineered with vectors of the invention, and production of
polypeptides of the invention by recombinant techniques. Cell-free translation systems can
also be employed to produce such proteins using RNAs derived from the DNA constructs of
the invention.

Preferred vectors are viral vectors, such as lentiviruses, retroviruses, herpes viruses,
10 adenoviruses, adeno-associated viruses, vaccinia virus, baculovirus, and other recombinant
viruses with desirable cellular tropism. Thus, a gene encoding a functional or mutant protein
or polypeptide, or fragment thereof can be introduced *in vivo*, *ex vivo*, or *in vitro* using a viral
vector or through direct introduction of DNA. Expression in targeted tissues can be effected
by targeting the transgenic vector to specific cells, such as with a viral vector or a receptor
15 ligand, or by using a tissue-specific promoter, or both. Targeted gene delivery is described in
PCT Publication Number WO 95/28494.

Viral vectors commonly used for *in vivo* or *ex vivo* targeting and therapy procedures
are DNA-based vectors and retroviral vectors. Methods for constructing and using viral
vectors are known in the art (e.g., Miller and Rosman, BioTechniques, 1992, 7:980-990).
20 Preferably, the viral vectors are replication-defective, that is, they are unable to replicate
autonomously in the target cell. Preferably, the replication defective virus is a minimal virus,
i.e., it retains only the sequences of its genome which are necessary for encapsulating the
genome to produce viral particles.

DNA viral vectors include an attenuated or defective DNA virus, such as, but not
25 limited to, herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV),
adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or
almost entirely lack viral genes, are preferred. A defective virus is not infective after
introduction into a cell. Use of defective viral vectors allows for administration to cells in a
specific, localized area, without concern that the vector can infect other cells. Thus, a
30 specific tissue can be specifically targeted. Examples of particular vectors include, but are
not limited to, a defective herpes virus 1 (HSV1) vector (Kaplitt et al., Molec. Cell.

WO 02/083859

PCT/US02/11610

Neurosci., 1991, 2:320-330), defective herpes virus vector lacking a glycoprotein L gene, or other defective herpes virus vectors (PCT Publication Numbers WO 94/21807 and WO 92/05263); an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet et al. (J. Clin. Invest., 1992, 90:626-630; see also La Salle et al., Science, 1993, 259:988-990); and a defective adeno-associated virus vector (Samulski et al., J. Virol., 1987, 61:3096-3101; Samulski et al., J. Virol., 1989, 63:3822-3828; Lebkowski et al., Mol. Cell Biol., 1988, 8:3988-3996).

Various companies produce viral vectors commercially, including, but not limited to, Avigen, Inc. (Alameda, California; AAV vectors), Cell Genesys (Foster City, California; retroviral, adenoviral, AAV vectors, and lentiviral vectors), Clontech (retroviral and baculoviral vectors), Genovo, Inc. (Sharon Hill, Pennsylvania; adenoviral and AAV vectors), Genvec (adenoviral vectors), IntroGene (Leiden, Netherlands; adenoviral vectors), Molecular Medicine (retroviral, adenoviral, AAV, and herpes viral vectors), Norgen (adenoviral vectors), Oxford BioMedica (Oxford, United Kingdom; lentiviral vectors), and Transgene (Strasbourg, France; adenoviral, vaccinia, retroviral, and lentiviral vectors).

Adenoviruses are eukaryotic DNA viruses that can be modified to efficiently deliver a nucleotide of the invention to a variety of cell types. Various serotypes of adenovirus exist. Of these serotypes, preference is given, within the scope of the invention, to using type 2 or type 5 human adenoviruses (Ad 2 or Ad 5) or adenoviruses of animal origin (See, PCT Publication Number WO 94/26914.). Those adenoviruses of animal origin which can be used within the scope of the invention include adenoviruses of canine, bovine, murine (e.g., Mav1, Beard et al., Virology, 1990, 75-81), ovine, porcine, avian, and simian (e.g., SAV) origin. Preferably, the adenovirus of animal origin is a canine adenovirus, more preferably a CAV2 adenovirus (e.g., Manhattan or A26/61 strain, ATCC VR-800, for example). Various replication defective adenovirus and minimum adenovirus vectors have been described (e.g., PCT Publication Numbers WO 94/26914, WO 95/02697, WO 94/28938, WO 94/28152, WO 94/12649, WO 95/02697, WO 96/22378). The replication defective recombinant adenoviruses according to the invention can be prepared by any technique known to the person skilled in the art (e.g., Levrero et al., Gene, 1991, 101:195; European Publication Number EP 185 573; Graham, EMBO J., 1984, 3:2917; Graham et al., J. Gen. Virol., 1977, 36:59). Recombinant adenoviruses are recovered and purified using standard molecular biological techniques, which are well known to one of ordinary skill in the art.

WO 02/083859

PCT/US02/11610

The adeno-associated viruses (AAV) are DNA viruses of relatively small size that can integrate, in a stable and site-specific manner, into the genome of the cells which they infect. They are able to infect a wide spectrum of cells without inducing any effects on cellular growth, morphology, or differentiation, and they do not appear to be involved in human pathologies. The AAV genome has been cloned, sequenced, and characterized. The use of vectors derived from the AAVs for transferring genes *in vitro* and *in vivo* has been described (See, PCT Publication Numbers WO 91/18088 and WO 93/09239; U.S. Patent Numbers 4,797,368 and 5,139,941; European Publication Number EP 488 528). The replication defective recombinant AAVs according to the invention can be prepared by cotransfecting a plasmid containing the nucleic acid sequence of interest flanked by two AAV inverted terminal repeat (ITR) regions, and a plasmid carrying the AAV encapsidation genes (rep and cap genes), into a cell line which is infected with a human helper virus (for example, an adenovirus). The AAV recombinants which are produced are then purified by standard techniques.

In another embodiment, the gene can be introduced in a retroviral vector, e.g., as described in U.S. Patent Number 5,399,346; Mann et al., Cell, 1983, 33:153; U.S. Patent Numbers 4,650,764 and 4,980,289; Markowitz et al., J. Virol., 1988, 62:1120; U.S. Patent Number 5,124,263; European Publication Numbers EP 453 242 and EP178 220; Bernstein et al., Genet. Eng., 1985, 7:235; McCormick, BioTechnology, 1985, 3:689; PCT Publication Number WO 95/07358; and Kuo et al., Blood, 1993, 82:845. The retroviruses are integrating viruses that infect dividing cells. The retrovirus genome includes two LTRs, an encapsidation sequence, and three coding regions (gag, pol and env). In recombinant retroviral vectors, the gag, pol and env genes are generally deleted, in whole or in part, and replaced with a heterologous nucleic acid sequence of interest. These vectors can be constructed from different types of retrovirus, such as, HIV, MoMuLV ("murine Moloney leukaemia virus"), MSV ("murine Moloney sarcoma virus"), HaSV ("Harvey sarcoma virus"), SNV ("spleen necrosis virus"), RSV ("Rous sarcoma virus"), and Friend virus. Suitable packaging cell lines have been described, in particular the cell line PA317 (U.S. Patent Number 4,861,719), the PsiCRIP cell line (PCT Publication Number WO 90/02806), and the GP+envAm-12 cell line (PCT Publication Number WO 89/07150). In addition, the recombinant retroviral vectors can contain modifications within the LTRs for suppressing transcriptional activity as well as extensive encapsidation sequences which may include a part

WO 02/083859

PCT/US02/11610

of the gag gene (Bender et al., J. Virol., 1987, 61:1639). Recombinant retroviral vectors are purified by standard techniques known to those having ordinary skill in the art.

Retroviral vectors can be constructed to function as infectious particles or to undergo a single round of transfection. In the former case, the virus is modified to retain all of its genes except for those responsible for oncogenic transformation properties, and to express the heterologous gene. Non-infectious viral vectors are manipulated to destroy the viral packaging signal, but retain the structural genes required to package the co-introduced virus engineered to contain the heterologous gene and the packaging signals. Thus, the viral particles that are produced are not capable of producing additional virus.

Retrovirus vectors can also be introduced by DNA viruses, which permits one cycle of retroviral replication and amplifies transfection efficiency (See, PCT Publication Numbers WO 95/22617, WO 95/26411, WO 96/39036 and WO 97/19182.).

In another embodiment, lentiviral vectors can be used as agents for the direct delivery and sustained expression of a transgene in several tissue types, including brain, retina, muscle, liver, and blood. The vectors can efficiently transduce dividing and nondividing cells in these tissues, and maintain long-term expression of the gene of interest. For a review, see, Naldini, Curr. Opin. Biotechnol., 1998, 9:457-63; see also, Zufferey et al., J. Virol., 1998, 72:9873-80. Lentiviral packaging cell lines are available and known generally in the art. They facilitate the production of high-titer lentivirus vectors for gene therapy. An example is a tetracycline-inducible VSV-G pseudotyped lentivirus packaging cell line that can generate virus particles at titers greater than 10⁶ IU/ml for at least 3 to 4 days (Kafri et al., J. Virol., 1999, 73: 576-584). The vector produced by the inducible cell line can be concentrated as needed for efficiently transducing non-dividing cells *in vitro* and *in vivo*.

In another embodiment, the vector can be introduced *in vivo* by lipofection, as naked DNA, or with other transfection facilitating agents (peptides, polymers, etc.). Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felgner et al., Proc. Natl. Acad. Sci. U.S.A., 1987, 84:7413-7417; Felgner and Ringold, Science, 1989, 337:387-388; Mackey et al., Proc. Natl. Acad. Sci. U.S.A., 1988, 85:8027-8031; Ulmer et al., Science, 1993, 259:1745-1748). Useful lipid compounds and compositions for transfer of nucleic acids are described in PCT Patent Publication Numbers WO 95/18863 and WO 96/17823, and in U.S. Patent Number 5,459,127. Lipids may be

WO 02/083859

PCT/US02/11610

chemically coupled to other molecules for the purpose of targeting (see Mackey, et al., supra). Targeted peptides, e.g., hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

One can also introduce the vector *in vivo* as a naked DNA plasmid. Naked DNA
5 vectors for gene therapy can be introduced into the desired host cells by methods known in the art, e.g., electroporation, microinjection, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun, or use of a DNA vector transporter (e.g., Wu et al., J. Biol. Chem., 1992, 267:963-967; Wu and Wu, J. Biol. Chem., 1988, 263:14621-14624; Canadian Patent Application Number 2,012,311; Williams et al., Proc. Natl. Acad. Sci. USA, 1991,
10 88:2726-2730). Receptor-mediated DNA delivery approaches can also be used (Curiel et al., Hum. Gene Ther., 1992, 3:147-154; Wu and Wu, J. Biol. Chem., 1987, 262:4429-4432). U.S. Patent Numbers 5,580,859 and 5,589,466 disclose delivery of exogenous DNA sequences, free of transfection facilitating agents, in a mammal. Recently, a relatively low voltage, high efficiency *in vivo* DNA transfer technique, termed electrotransfer, has been
15 described (Mir et al., C.P. Acad. Sci., 1988, 321:893; PCT Publication Numbers WO 99/01157; WO 99/01158; WO 99/01175).

Other molecules are also useful for facilitating transfection of a nucleic acid *in vivo*, such as a cationic oligopeptide (e.g., PCT Patent Publication Number WO 95/21931), peptides derived from DNA binding proteins (e.g., PCT Patent Publication Number
20 WO 96/25508), or a cationic polymer (e.g., PCT Patent Publication Number WO 95/21931), or bupivacaine (U.S. Patent Number 5,593,972).

The isolated polypeptide of the present invention can be delivered to the mammal using a live vector, in particular using live recombinant bacteria, viruses, or other live agents, containing the genetic material necessary for the expression of the polypeptide or
25 immunogenic fragment as a foreign polypeptide. Particularly, bacteria that colonize the gastrointestinal tract, such as *Salmonella*, *Shigella*, *Yersinia*, *Vibrio*, *Escherichia* and BCG have been developed as vaccine vectors, and these and other examples are discussed by Holmgren et al. (1992) and McGhee et al. (1992).

The following might be used as part of a list of RNA vectors, in which one or more of
30 the immunogenic candidate proteins may be inserted.

WO 02/083859

PCT/US02/11610

Classification of nonsegmented, negative-sense, single stranded RNA Viruses of the Order
Mononegavirales

Family Paramyxoviridae

Subfamily Paramyxovirinae

- 5 Genus *Paramyxovirus*
- Sendai virus (mouse parainfluenza virus type 1)
- Human parainfluenza virus (PIV) types 1 and 3
- Bovine parainfluenza virus (BPV) type 3
- Genus *Rubulavirus*
- 10 Simian virus 5 (SV) (Canine parainfluenza virus type 2)
- Mumps virus
- Newcastle disease virus (NDV) (avian Paramyxovirus 1)
- Human parainfluenza virus (PIV-types 2, 4a and 4b)
- Genus *Morbillivirus*
- 15 Measles virus (MV)
- Dolphin Morbillivirus
- Canine distemper virus (CDV)
- Peste-des-petits-ruminants virus
- Phocine distemper virus
- 20 Rinderpest virus
- Unclassified
- Hendra virus
- Nipah virus
- Subfamily Pneumovirinae
- 25 Genus *Pneumovirus*
- Human respiratory syncytial virus (RSV)
- Bovine respiratory syncytial virus
- Pneumonia virus of mice
- Genus *Metapneumovirus*
- 30 Human metapneumovirus
- Avian pneumovirus (formerly Turkey rhinotracheitis virus)
- Family Rhabdoviridae
- Genus *Lyssavirus*
- Rabies virus

WO 02/083859

PCT/US02/11610

Genus *Vesiculovirus*

Vesicular stomatitis virus (VSV)

Genus *Ephemerovirus*

Bovine ephemeral fever virus

5 Family FiloviridaeGenus *Filovirus*

Marburg virus

The RNA virus vector is basically an isolated nucleic acid molecule that comprises a sequence which encodes at least one genome or antigenome of a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales. The isolated nucleic acid molecule may comprise a polynucleotide sequence which encodes a genome, antigenome, or a modified version thereof. In one embodiment, the polynucleotide encodes an operably linked promoter, the desired genome or antigenome, and a transcriptional terminator.

In a preferred embodiment of this invention, the polynucleotide encodes a genome or antigenome that has been modified from a wild-type RNA virus by a nucleotide insertion, rearrangement, deletion, or substitution. The genome or antigenome sequence can be derived from a human or non-human virus. The polynucleotide sequence may also encode a chimeric genome formed from recombinantly joining a genome or antigenome from two or more sources. For example, one or more genes from the A group of RSV are inserted in place of the corresponding genes of the B group of RSV; or one or more genes from bovine PIV (BPIV), PIV-1 or PIV-2 are inserted in the place of the corresponding genes of PIV-3; or RSV may replace genes of PIV and so forth. In additional embodiments, the polynucleotide encodes a genome or anti-genome for an RNA virus of the Order Mononegavirales which is a human, bovine, or murine virus. Since the recombinant viruses formed by the methods of this invention are employed for therapeutic or prophylactic purposes, the polynucleotide may also encode an attenuated or an infectious form of the RNA virus selected. In many embodiments, the polynucleotide encodes an attenuated, infectious form of the RNA virus. In particularly preferred embodiments, the polynucleotide encodes a genome or antigenome of a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene, as described by published International patent application WO 98/13501, which is hereby incorporated by reference.

WO 02/083859

PCT/US02/11610

As vectors, the polynucleotide sequences encoding the modified forms of the desired genome and antigenome as described above also encode one or more genes or nucleotide sequences for the immunogenic proteins of this invention. In addition, one or more heterologous genes may also be included in forming a desired immunogenic composition/vector, as desired. Depending on the application of the desired recombinant virus, the heterologous gene may encode a co-factor, cytokine (such as interleukin), a T-helper epitope, a restriction marker, adjuvant, or a protein of a different microbial pathogen (e.g., virus, bacterium, or fungus), especially proteins capable of eliciting a protective immune response. The heterologous gene may also be used to provide agents which are used for gene therapy. In preferred embodiments, the heterologous genes encode cytokines, such as interleukin-12, which are selected to improve the prophylactic or therapeutic characteristics of the recombinant virus.

Antibodies

The polypeptides of the invention, including the amino acid sequences of even numbered SEQ ID NOS: 2-668, their fragments, and analogs thereof, or cells expressing them, can also be used as immunogens to produce antibodies immunospecific for the polypeptides of the invention. The invention includes antibodies immunospecific for β -hemolytic streptococci and *Streptococcus pyogenes* polypeptides and the use of such antibodies to detect the presence of, or measure the quantity or concentration of, β -hemolytic streptococci and *Streptococcus pyogenes* polypeptides in a cell, a cell or tissue extract, or a biological fluid.

The antibodies of the invention include polyclonal antibodies, monoclonal antibodies, chimeric antibodies, and anti-idiotypic antibodies. Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen. Monoclonal antibodies are a substantially homogeneous population of antibodies to specific antigens. Monoclonal antibodies may be obtained by methods known to those skilled in the art, e.g., Kohler and Milstein, 1975, Nature 256:495-497 and U.S. Patent Number 4,376,110. Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, GILD and any subclass thereof.

Chimeric antibodies are molecules, different portions of which are derived from different animal species, such as those having variable region derived from a murine

WO 02/083859

PCT/US02/11610

monoclonal antibody and a human immunoglobulin constant region. Chimeric antibodies and methods for their production are known in the art (Cabilly et al., 1984, Proc. Natl. Acad. Sci. USA 81:3273-3277; Morrison et al., 1984, Proc. Natl. Acad. Sci. USA 81:6851-6855; Boulianne et al., 1984, Nature 312:643-646; Cabilly et al., European Patent Application 125023 (published November 14, 1984); Taniguchi et al., European Patent Application 171496 (published February 19, 1985); Morrison et al., European Patent Application 173494 (published March 5, 1986); Neuberger et al., PCT Application WO 86/01533 (published March 13, 1986); Kudo et al., European Patent Application 184187 (published June 11, 1986); Morrison et al., European Patent Application 173494 (published March 5, 1986); Sahagan et al., 1986, J. Immunol. 137:1066-1074; Robinson et al., PCT/US86/02269 (published May 7, 1987); Liu et al., 1987, Proc. Natl. Acad. Sci. USA 84:3439-3443; Sun et al., 1987, Proc. Natl. Acad. Sci. USA 84:214-218; Better et al., 1988, Science 240:1041-1043). These references are hereby incorporated by reference.

An anti-idiotypic (anti-Id) antibody is an antibody which recognizes unique determinants generally associated with the antigen-binding site of an antibody. An anti-Id antibody is prepared by immunizing an animal of the same species and genetic type (e.g., mouse strain) as the source of the monoclonal antibody with the monoclonal antibody to which an anti-Id is being prepared. The immunized animal will recognize and respond to the idiotypic determinants of the immunizing antibody by producing an antibody to these isotypic determinants (the anti-Id antibody).

Accordingly, monoclonal antibodies generated against the polypeptides of the present invention may be used to induce anti-Id antibodies in suitable animals. Spleen cells from such immunized mice can be used to produce anti-Id hybridomas secreting anti-Id monoclonal antibodies. Further, the anti-Id antibodies can be coupled to a carrier such as keyhole limpet hemocyanin (KLH) and used to immunize additional BALB/c mice. Sera from these mice will contain anti-anti-Id antibodies that have the binding properties of the final mAb specific for a R-PTPase epitope. The anti-Id antibodies thus have their idiotypic epitopes, or "idiotoxes" structurally similar to the epitope being evaluated, such as *Streptococcus pyogenes* polypeptides.

The term "antibody" is also meant to include both intact molecules as well as fragments such as Fab which are capable of binding antigen. Fab fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-

WO 02/083859

PCT/US02/11610

specific tissue binding than an intact antibody (Wahl et al., 1983, J. Nucl. Med. 24:316-325). It will be appreciated that Fab and other fragments of the antibodies useful in the present invention may be used for the detection and quantitation of *Streptococcus pyogenes* polypeptides according to the methods for intact antibody molecules.

5 The anti-Id antibody may also be used as an "immunogen" to induce an immune response in yet another animal, producing a so-called anti-anti-Id antibody. The anti-anti-Id may be epitopically identical to the original mAb which induced the anti-Id. Thus, by using antibodies to the idiotypic determinants of a mAb, it is possible to identify other clones expressing antibodies of identical specificity.

10 The antibodies are used in a variety of ways, e.g., for confirmation that a protein is expressed, or to confirm where a protein is expressed. Labeled antibody (e.g., fluorescent labeling for FACS) can be incubated with intact bacteria and the presence of the label on the bacterial surface confirms the location of the protein, for instance.

15 Antibodies generated against the polypeptides of the invention can be obtained by administering the polypeptides or epitope-bearing fragments, analogs, or cells to an animal using routine protocols. For preparing monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures are used.

Immunogenic Compositions

20 Also provided are immunogenic compositions. The immunogenic compositions of the present invention can be used for the treatment of streptococcal infections in mammals, such as humans (preferably) and non-human animals. For example, the animals may be bovine, canine, equine, feline, and porcine. It is noted that SEQ ID NO: 415 (ORF 1021) corresponds to a protein which also appears in *S. equi*. Accordingly, this sequence can be used in immunogenic compositions for treating equine infections, as well as in other animals
25 or humans. Particular applications include, but are not limited to, the treatment of strangles, a highly contagious disease of the nasopharynx and draining lymph nodes of Equidae, and the treatment of respiratory infections and mastitis in bovines, equines, and swine.

30 The immunogenic compositions of the invention may either be prophylactic (i.e., to prevent infection or reduce the onset of infection) or therapeutic (i.e., to treat a disease or side effects caused by an infection after the infection).

WO 02/083859

PCT/US02/11610

The immunogenic compositions may comprise a polypeptide of the invention. To do so, one or more polypeptides are adjusted to an appropriate concentration and can be formulated with any suitable adjuvant, diluent, carrier, or any combination thereof. Physiologically acceptable media may be used as carriers and/or diluents. These include, but
5 are not limited to, water, an appropriate isotonic medium, glycerol, ethanol and other conventional solvents, phosphate buffered saline, and the like.

As used herein, an "adjuvant" is a substance that serves to enhance the immunogenicity of an antigen, whether it is a polypeptide or a polynucleotide. Thus, adjuvants are often given to boost the immune response and are well known to the skilled
10 artisan. Suitable adjuvants include, but are not limited to, aluminum salts (alum), such as aluminum phosphate and aluminum hydroxide, *Mycobacterium tuberculosis*, *Bordetella pertussis*, bacterial lipopolysaccharides, aminoalkyl glucosamine phosphate compounds (AGP), or derivatives or analogs thereof, which are available from Corixa (Hamilton, MT), and which are described in United States Patent Number 6,113,918, which is hereby
15 incorporated by reference. One such AGP is 2-ethyl 2-Deoxy-4-O-phosphono-3-O-2-b-D-glucopyranoside, which is also known as 529 (formerly known as RC529). This 529 adjuvant is formulated as an aqueous form or as a stable emulsion. Other adjuvants are MPL® (3-O-deacylated monophosphoryl lipid A) (Corixa) described in U.S. Patent Number 4,912,094, synthetic polynucleotides such as oligonucleotides containing a CpG motif (U.S.
20 Patent Number 6,207,646, saponins such as Quil A or STIMULON® QS-21 (Antigenics, Framingham, Massachusetts), described in U.S. Patent Number 5,057,540, a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., International Patent Publication Nos. WO 93/13302 and WO 92/19265, cholera toxin (either in a wild-type or mutant form, for example, wherein the glutamic acid at
25 amino acid position 29 is replaced by another amino acid, preferably a histidine, in accordance with published International Patent Application number WO 00/18434).

Various cytokines and lymphokines are suitable for use as adjuvants. One such adjuvant is granulocyte-macrophage colony stimulating factor (GM-CSF), which has a nucleotide sequence as described in U.S. Patent Number 5,078,996, which is hereby
30 incorporated by reference. A plasmid containing GM-CSF cDNA has been transformed into *E. coli* and has been deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, under Accession Number 39900. The

WO 02/083859

PCT/US02/11610

cytokine Interleukin-12 (IL-12) is another adjuvant which is described in U.S. Patent Number 5,723,127, which is hereby incorporated by reference. Other cytokines or lymphokines have been shown to have immune modulating activity, including, but not limited to, the interleukins 1-alpha, 1-beta, 2, 4, 5, 6, 7, 8, 10, 13, 14, 15, 16, 17 and 18, the interferons-
5 alpha, beta and gamma, granulocyte colony stimulating factor, and the tumor necrosis factors alpha and beta, and are suitable for use as adjuvants.

The polypeptide can also include at least a portion of the polypeptide, optionally conjugated or linked to a peptide, polypeptide, or protein, or to a polysaccharide.

The immunogenic compositions of the invention can further include immunogenic
10 conjugates as disclosed in U.S. Patent Numbers 4,673,574, 4,902,506, 5,097,020, and 5,360,897 (assigned to The University of Rochester), hereby incorporated by reference. These patents teach immunogenic conjugates which are the reductive amination product of an immunogenic capsular polymer fragment having a reducing end and derived from a bacterial capsular polymer of a bacterial pathogen, and a bacterial toxin or toxoid. The present
15 invention also includes immunogenic compositions containing these conjugates which elicit effective levels of anti-capsular polymer antibodies in humans.

Combination immunogenic compositions are provided by including two or more of the polypeptides of the invention, as well as by combining one or more of the polypeptides of the invention with one or more known *Streptococcus pyogenes* polypeptides, including, but
20 not limited to, the C5a peptidase, the M proteins, adhesins, and the like.

The immunogenic compositions of the invention also comprise a polynucleotide sequence of the invention operatively associated with a regulatory sequence that controls gene expression. The polynucleotide sequence of interest is engineered into an expression vector, such as a plasmid, under the control of regulatory elements which will promote
25 expression of the DNA, that is, promoter and/or enhancer elements. In a preferred embodiment, the human cytomegalovirus immediate-early promoter/enhancer is used (U.S. Patent Number 5,168,062). The promoter may be cell-specific and permit substantial transcription of the polynucleotide only in predetermined cells.

The polynucleotide is introduced directly into the host either as "naked" DNA (U.S.
30 Patent Number 5,580,859) or formulated in compositions with agents which facilitate

WO 02/083859

PCT/US02/11610

immunization, such as bupivacaine and other local anesthetics (U.S. Patent Number 5,593,972) and cationic polyamines (U.S. Patent Number 6,127,170).

5 In this polynucleotide immunization procedure, the polypeptides of the invention are expressed on a transient basis *in vivo*; no genetic material is inserted or integrated into the chromosomes of the host. This procedure is to be distinguished from gene therapy, where the goal is to insert or integrate the genetic material of interest into the chromosome. An assay is used to confirm that the polynucleotides administered by immunization do not give rise to a transformed phenotype in the host (U.S. Patent Number 6,168,918).

10 Once formulated, the immunogenic compositions of the invention can be administered directly to the subject, delivered *ex vivo* to cells derived from the subject, or *in vitro* for expression of recombinant proteins. For delivery directly to the subject, administration may be by any conventional form, such as intranasally, parenterally, orally, intraperitoneally, intravenously, subcutaneously, or topically applied to any mucosal surface such as intranasal, oral, eye, lung, vaginal, or rectal surface, such as by an aerosol spray.

15 The subjects can be mammals or birds. The subject can also be a human. An immunologically effective amount of the immunogenic composition in an appropriate number of doses is administered to the subject to elicit an immune response. Immunologically effective amount, as used herein, means the administration of that amount to a mammalian host (preferably human), either in a single dose or as part of a series of
20 doses, sufficient to at least cause the immune system of the individual treated to generate a response that reduces the clinical impact of the bacterial infection. Protection may be conferred by a single dose of the immunogenic composition, or may require the administration of several doses, in addition to booster doses at later times to maintain protection. This may range from a minimal decrease in bacterial burden to prevention of the
25 infection. Ideally, the treated individual will not exhibit the more serious clinical manifestations of the β -hemolytic streptococcal infection. The dosage amount can vary depending upon specific conditions of the individual, such as age and weight. This amount can be determined in routine trials by means known to those skilled in the art.

30 Various tests are used to assess the *in vitro* immunogenicity of the polypeptides of the invention. For example, the polypeptides can be expressed recombinantly or chemically synthesized and used to screen subject sera by immunoblot. A positive reaction between the

WO 02/083859

PCT/US02/11610

subject and subject serum indicates that the subject has previously mounted an immune response to the polypeptide in question, i.e., the polypeptide is an immunogen. This method can also be used to identify immunodominant polypeptides.

5 An ELISA assay is also used to assess *in vitro* immunogenicity, wherein the polypeptide antigen of interest is coated onto a plate, such as a 96 well plate, and test sera from either a vaccinated or naturally exposed animal (e.g., human) is reacted with the coating antigen. If any antibody, specific for the test polypeptide antigen, is present, it can be detected by standard methods known to one skilled in the art.

10 Alternatively, the same sera can be reacted with whole *Streptococcus pyogenes* cells. Reactive antibody present in the sera can then be detected using a colloidal gold conjugated antibody and visualized by LV-SEM.

Efficacy of vaccine antigens can be tested using two animal challenge assay models. The first addresses mucosal immunity. Mice are actively immunized, parenterally or mucosally, with the vaccine candidates following established procedures. The mice are then
15 challenged with wild-type *Streptococcus pyogenes* by intranasal administration. *Streptococcus pyogenes* persistence in the nasal/pharyngeal cavity of the mice can then be measured by standard techniques. Efficacy is reflected by an enhanced clearance of the bacteria from the throats of the animals.

20 Alternatively, subsequent to active parenteral immunization, protection against systemic infection can be evaluated by subcutaneous injection of *Streptococcus pyogenes* cells. Efficacy is measured by reduction in death and/or reduced histopathology at the site of injection.

Detection in a sample

Also provided are methods for detecting and identifying β -hemolytic streptococcus
25 and *Streptococci pyogenes* in a biological sample. In one embodiment, the method comprises the steps of (a) contacting the biological sample with a polynucleotide of the invention under conditions that permit hybridization of complementary base pairs and (b) detecting the presence of hybridization complexes in the sample. In another embodiment, the method comprises the steps of (a) contacting the biological sample with an antibody of the invention
30 under conditions suitable for the formation of immune complexes and (b) detecting the

WO 02/083859

PCT/US02/11610

presence of immune complexes in the sample. In yet another embodiment, the method comprises the steps of (a) contacting the biological sample with a polypeptide of the invention under conditions suitable for the formation of immune complexes and (b) detecting the presence of immune complexes in the sample.

5 Antigens, or antigenic fragments thereof, of the invention are used in immunoassays to detect antibody levels or, conversely, anti- *Streptococcus pyogenes* antibodies are used to detect antigen levels. Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostic methods. Antibodies to the polypeptides of the invention within biological samples, including, for example, blood or serum samples, can be
10 detected. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. The polypeptides of the invention can also be a useful in receptor-ligand studies.

15 The following examples are illustrative and the present invention is not intended to be limited thereto.

EXAMPLE 1

Bacteria, media, and reagents

E. coli was cultured and maintained in SOB (0.5% Yeast Extract, 2.0% Tryp, 10mM Sodium Chloride, 2.5mM Potassium Chloride, 10mM Magnesium Chloride, 10mM
20 Magnesium Sulfate) containing the appropriate antibiotic. Ampicillin was used at a concentration of 100 µg/mL, chloramphenicol at 30 µg/mL, and kanamycin at 50 µg/mL. The *Streptococcus pyogenes* strain SF370 (ATCC accession number 700294) was cultured in 30 g/L Todd Hewitt, 5 g/L yeast extract (THY) broth.

Bioinformatics/Gene mining

25 The genomic, unannotated sequence of *Streptococcus pyogenes* M1 strain was downloaded from the website of the University of Oklahoma and was analyzed to identify open reading frames (ORFs). This genomic sequence was reported as being submitted to GenBank and assigned accession number AE004092, and strain M1 GAS was reported as being submitted to the ATCC and given accession number ATCC 700294.

WO 02/083859

PCT/US02/11610

An ORF was defined as having either one of three potential start site codons, ATG, GTG, or TTG and either one of three potential stop codons, TAA, TAG, TGA. A unique set of three ORF finder algorithms was used to enhance the efficiency for determining all ORFs: GLIMMER (59); GeneMark (34); and a third algorithm developed by inventor's assignee.

5 In order to evaluate the accuracy of the ORFs determined, a discrete mathematical cosine function, known in the art has a discrete cosine transformation (DiCTion), was employed to assign a score for each ORF. An ORF with a DiCTion score >1.5 is considered to have a high probability of encoding a protein product. The minimum length of an ORF predicted by the three ORF finding algorithms was set to 225 nucleotides (including stop
10 codon) which would encode a protein of 74 amino acids.

As a final search for remnants of ORFs, all noncoding regions >75 nucleotides were searched against the public protein databases (described below) using tBLASTn. This helped to identify regions of genes that contained frameshifts (42) or fragments of genes that might have a role in causing antigenic variation (21). Any remnant ORFs found here were added to
15 the ORF database of *Streptococcus pyogenes*. An in-house graphical analysis program was used to show all six reading frames and the location of the predicted ORFs relative to the genomic sequence. This helped to eliminate those ORFs that had large overlaps with other ORFs, although there are known cases of ORFs being totally embedded within other ORFs (25, 33).

20 The initial annotation of the *Streptococcus pyogenes* ORFs was performed using the BLAST v. 2.0 Gapped search algorithm, BLASTp, to identify homologous sequences. A cutoff "e" value of anything $< e^{-10}$ was considered significant. Other search algorithms, including FASTA and PSI-BLAST, were also used. The non-redundant protein sequence databases used for the homology searches consisted of GenBank, SWISS-PROT, PIR, and
25 TREMBL database sequences updated daily. ORFs with a BLASTp result of $> e^{-10}$ were considered to be unique to *Streptococcus pyogenes*.

A keyword search of the entire Blast results was carried out using known or suspected vaccine target genes as well as words that identified the location of a protein or function. Additionally, a keyword search was performed of all MEDLINE references associated with
30 the initial Blast results to look for additional information regarding the ORFs.

WO 02/083859

PCT/US02/11610

For DNA analysis, the %G+C content within each gene was identified. The %G+C content of an ORF was calculated as the (G+C) content of the third nucleotide position of all the codons within an ORF. The value reported was the difference of this value from the arithmetic mean of such values obtained for all ORFs found in the organism. Any absolute value ≥ 8 was considered important for further analysis, as these ORFs may have arisen from horizontal transfer as has been shown in the case of *cag* pathogenicity island from *H. pylori* (2), a pattern in keeping with many other pathogenicity islands (22).

Several parameters were used to determine partitioning of the predicted proteins.

Proteins destined for translocation across the cytoplasmic membrane encode a leader signal (also called signal sequence) composed of a central hydrophobic region flanked at the N-terminus by positively charged residues (56). The program SignalP was used to identify signal peptides and their cleavage sites (46). To predict protein localization in bacteria, the software PSORT was used (44). This program uses a neural net algorithm to predict localization of proteins to the cytoplasm, periplasm, and cytoplasmic membrane for Gram-positive bacteria as well as outer membrane for Gram-negative bacteria. Transmembrane (TM) domains of proteins were analyzed using the software program TopPred2 (10). This program predicts regions of a protein that are hydrophobic that may potentially span the lipid bilayer of the membrane. Outer membrane proteins typically do not have an α -helical TM domain.

The Hidden Markov Model (HMM) Pfam database of multiple alignments of protein domains or conserved protein regions (61) was used to identify *Streptococcus pyogenes* proteins that may belong to an existing protein family. Keyword searching of this output was used to help identify surface localized *Streptococcus pyogenes* proteins that might have been missed by the Blast search criteria. HMM models were also developed by inventor's assignee. A computer algorithm, HMM Lipo, was developed to predict lipoproteins using 132 biologically characterized non-*Streptococcus pyogenes* bacterial lipoproteins from over 30 organisms. This training set was generated from experimentally proven prokaryotic lipoproteins. The protein sequence from the start of the protein to the cysteine amino acid plus the next two additional amino acids were used to generate the HMM. Using about 70 known prokaryotic proteins containing the LPXTG cell wall sorting signal, a HMM (15) was developed to predict cell wall proteins that are anchored to the peptidoglycan layer (38, 45). The model used not only the LPXTG sequence, but also included two features of the

WO 02/083859

PCT/US02/11610

downstream sequence, the hydrophobic transmembrane domain and the positively charged carboxy terminus. There are also a number of proteins that interact, non-covalently, with the peptidoglycan layer and are distinct from the LPXTG protein class described above. These proteins seem to have a consensus sequence at their carboxy terminus (32). A HMM of this region was developed and used to identify *Streptococcus pyogenes* proteins falling into this class.

The proteins encoded by *Streptococcus pyogenes* identified ORFs were also evaluated for other characteristics. A tandem repeat finder (5) identified ORFs containing repeated DNA sequences such as those found in MSCRAMMs (20) and phase variable surface proteins of *Neisseria meningitidis* (51). Proteins that contain the Arg-Gly-Asp (RGD) attachment motif, together with integrins that serve as their receptor, constitute a major recognition system for cell adhesion. RGD recognition is one mechanism used by microbes to gain entry into eukaryotic tissues (29, 63). However, not all RGD-containing proteins mediate cell attachment. It has been shown that RGD-containing peptides with a proline at the carboxy end (RGDP) are inactive in cell attachment assays (52) and, hence, were excluded. Geanfammer software was used to cluster proteins into homologous families (50). Preliminary analysis of the family classes provided novel ORFs within a vaccine candidate cluster as well as defining potential protein function.

Tryptic digestion of *Streptococcus pyogenes*

A starter culture of *Streptococcus pyogenes* was grown overnight in THY at 37° C, in 5% CO₂, or in atmospheric O₂. Each starter culture was then diluted 1:25 in 200 mL fresh THY, and grown to an OD₄₉₀ of 1-1.3, in either CO₂ or atmospheric O₂, respectively. The cells were then harvested by centrifugation at 4,000 x g, for 15 min., and washed three times in 10 mL 20 mM Tris, pH 8.0, 150 mM NaCl buffer. Following the last wash, each pellet was resuspended in 2 mL same buffer containing 0.8 M sucrose and distributed equally between two tubes. To one tube of each growth condition, 40 µg trypsin was added; the other tube was used as a negative digestion control. The cell suspensions were rocked at 37° C for 4 hours. A sample of each suspension was taken for viable cell counts and visualization by low-voltage scanning electron microscopy (LV-SEM). The suspensions were then centrifuged and the supernatants were collected and filtered through a low protein binding, 2 µM filter.

WO 02/083859

PCT/US02/11610

Micro-capillary HPLC Interface

Peptide extracts were analyzed on an automated microelectrospray reversed phase HPLC. The microelectrospray interface consisted of a Pico frit fused silica spray needle, 50 cm length by 75 μ m ID, 8 μ m orifice diameter (New Objective, Cambridge Massachusetts) packed with 10 μ m C18 reversed-phase beads (YMC, Wilmington, North Carolina) to a length of 10 cm. The Pico frit needle was mounted in a fiber optic holder (Melles Griot, Irvine, California) held on a base positioned at the front of the mass spectrometer detector. The rear of the column was plumbed through a titanium union to supply an electrical connection for the electrospray interface. The union was connected with a length of fused silica capillary (FSC) tubing to a FAMOS autosampler (LC-Packings, San Francisco, California) that was connected to an HPLC solvent pump (ABI 140C, Perkin-Elmer, Norwalk, Connecticut). The HPLC solvent pump delivered a flow of 50 μ L/min. which was reduced to 250 nL/min. using a PEEK microtight splitting tee (Upchurch Scientific, Oak Harbor, Washington), and then delivered to the autosampler using an FSC transfer line. The LC pump and autosampler were each controlled using their internal user programs. Samples were inserted into plastic autosampler vials, sealed, and injected using a 5 μ L sample loop.

Microcapillary HPLC-Mass Spectrometry

Extracted peptides from the surface digests were concentrated 10-fold using a Savant Speed Vac Concentrator (ThermoQuest, Holdbrook, New York), and then were separated by the microelectrospray HPLC system using a 50 min. gradient of 0-50% solvent B (A: 0.1M HoAc, B: 90% MeCN/0.1M HoAc). Peptide analyses were conducted on a Finnigan LCQ-DECA ion trap mass spectrometer (ThermoQuest, San Jose, California) operating at a spray voltage of 1.5 kV, and using a heated capillary temperature of 125° C. Data were acquired in automated MS/MS mode using the data acquisition software provided with the instrument. The acquisition method included 1 MS scan (375-600 m/z) followed by MS/MS scans of the top 2 most abundant ions in the MS scan. The instrument then conducted a second MS scan (600-1000 m/z) followed by MS/MS scans of the top 2 most abundant ions in that scan. The dynamic exclusion and isotope exclusion functions were employed to increase the number of peptide ions that were analyzed (settings: 3 amu = exclusion width, 3 min. = exclusion duration, 30 sec = pre-exclusion duration, 3 amu = isotope exclusion width).

Data Analysis

WO 02/083859

PCT/US02/11610

Automated analysis of MS/MS data was performed using the SEQUEST computer algorithm incorporated (17) into the Finnigan Bioworks data analysis package (ThermoQuest, San Jose, California) using the database of proteins derived from the complete genome of *Streptococcus pyogenes*.

5 Cloning and protein expression

Primer sets were designed for PCR amplification of desired ORFs such that the forward 5' primer would anneal at the start of the predicted mature protein. For lipoproteins, the 5' forward primer was designed to anneal just after the codon encoding a cysteine residue of the mature protein to minimize disulfide bridging. Design of the opposing reverse 3' primers was dependent upon the type of predicted protein. For those proteins that contained an LPXTG, the primer was designed such that it would anneal at the beginning (5' end) of the cell wall anchor region. For all other predicted proteins, they were designed such that they would anneal at the 3' end of the ORF. Additionally, the 5'-forward primer was initially designed to allow an in-frame fusion to thioredoxin with the opposing 3'-reverse primer allowing read-through to include a downstream his-patch and V5 epitope (pBAD/thio-TOPO®, Invitrogen, Carlsbad, California). The pBAD vector uses an arabinose inducible promoter. In parallel, these same PCR products were also cloned into pCRT7 TOPO® (Invitrogen, Carlsbad, California). This allowed for an N-terminal fusion to an Xpress epitope and a his-tag for purification.

20 All PCR reactions used the *Streptococcus pyogenes* M1 strain, SF370 (ATCC accession number 700294), as the template. PCR products were transformed into the *E. coli* host, TOP10, and plated on SOB containing 100 µg/mL ampicillin. Colonies were screened by PCR amplification using a vector specific 5' primer and the specific 3' reverse primer annealing to the gene insert. Colonies were seeded into wells of a 96 well microtiter plates containing 50 µL 50% glycerol. 10-12 colonies per gene were seeded in one row of the plate. In a second 96 well PCR plate, 50 µL reactions were set up specific to the gene of interest. One µL of the cells suspended in glycerol was used as template in the PCR reaction. Reactions that produced bands of the expected size were analyzed further. The cells that were seeded in 50% glycerol had SOB media added to them and were incubated at 37° C for 5-8 hours and frozen at -70° C.

WO 02/083859

PCT/US02/11610

PCR positive colonies were inoculated into 2 mL cultures for overnight growth. Part of the culture was used to prepare plasmid DNA that was analyzed by restriction digest to confirm the inserts while another part was used to seed 10 mL expression cultures (for pBAD plasmids) for expression. Mid-log phase cultures were induced with 0.5% L-arabinose for 2 hours. T7/NT plasmids were transformed into the expression strain BLR(DE3) pLysS before screening. T7/NT cultures were induced by the addition of 1 mM IPTG and incubated for 2 hours. Whole cell lysates of induced cultures were run on SDS-PAGE in duplicate. One gel was stained with coomassie and the other was transferred to nitrocellulose and probed with antibody to the relevant epitope tag.

Positive clones were grown in 1-2 L volumes and induced for large-scale purification. Solubility and expression level of the recombinant proteins were assessed by freeze-thaw lysis of the cells followed by DNase/RNase digestion and centrifugation at 9,000 x g for 15 min. in a RCSB refrigerated centrifuge (sorbol®, Dupont, Wilmington, Delaware). The soluble fraction was removed from the insoluble material and both were separated and evaluated for protein localization and expression by SDS-PAGE. Soluble fusion proteins were purified by passing the soluble fraction of lysed cells over Ni-NTA (Qiagen Inc., Valencia, California) resin and eluting the bound proteins with imidazole. Eluted proteins were buffer exchanged on PD-10 columns (Amersham Pharmacia Biotech, Piscataway, New Jersey).

Insoluble recombinant proteins were washed and centrifuged 3 times in PBS, 0.1% TRITON-X100. The inclusion bodies were then solubilized in PBS 4 M urea and buffer exchanged through a PD-10 column (Amersham Pharmacia, Piscataway, New Jersey) into PBS, 0.01% TRITON-X100, 0.5 M NaCl. Protein was quantitated by the Lowry assay and checked for purity and concentration by SDS-PAGE.

Generation of polyclonal antisera

Swiss Webster mice (5 per group) were immunized at weeks 0, 3, and 5 with 5 µg purified protein prepared above, 100 µg AlPO₄, and 50 µg MPL®, and were then bled at week 8.

Immunogold labeling of *Streptococcus pyogenes* and LV-SEM

WO 02/083859

PCT/US02/11610

Bacterial cells were labeled as previously described (49). Briefly, late-log phase bacterial cultures were washed twice, and resuspended to a concentration of 1×10^8 cells/ml in 10 mM phosphate buffered saline (PBS) (pH 7.4) and placed on poly-L-lysine coated glass coverslips. Excess bacteria were gently washed from the coverslips and unlabeled samples
5 were placed into fixative (2.0% glutaraldehyde, in a 0.1 M sodium cacodylate buffer containing 7.5% sucrose) for 30 min. Bacteria to be labeled with colloidal gold were washed with PBS containing 0.5% bovine serum albumin, and the pre-immune or hyper-immune mouse polyclonal antibody prepared above was applied for 1 hour at room temperature. Bacteria were then gently washed, and a 1:6 dilution of goat anti-mouse conjugated to 18 nm
10 colloidal gold particles (Jackson ImmunoResearch Laboratories, Inc., West Grove, Pennsylvania) was applied for 10 min. at room temperature. Finally, all samples were washed gently with PBS, and placed into the fixative described above. The fixative was washed from samples twice for 10 min. in 0.1 M sodium cacodylate buffer, and postfixed for 30 min. in 0.1 M sodium cacodylate containing 1% osmium tetroxide. The samples were
15 then washed twice with 0.1 M sodium cacodylate, dehydrated with ethanol, critical point dried by the CO₂ method of Anderson using a Samdri-780A (Tousimis, Rockville, Maryland), and coated with a 1-2 nm discontinuous layer of platinum. *Streptococcus pyogenes* cells were viewed with a LEO 1550 field emission scanning electron microscope operated at low accelerating voltages (1-4.5 keV) using a secondary electron detector for
20 conventional topographical imaging and a high-resolution Robinson backscatter detector to enhance the visualization of colloidal gold by atomic number contrast.

WO 02/083859

PCT/US02/11610

EXAMPLE 2 - IMMUNIZATION AND CHALLENGE

Parenteral immunization of mice

Six-week old, female CD1 (Charles River Breeding Laboratories, Inc., Wilmington, Mass.) or Swiss Webster (Taconic Farms Inc., Germantown, New York) mice are immunized at weeks 0, 4, and 6 with 5 µg protein of interest mixed with 50 µg MPL® (Corixa, Hamilton, MT) and 100 µg AlPO₄ per dose to a final volume of 200 µL in saline and then injected subcutaneously (s.c.) into mice. Control mice are injected with 5 µg tetanus toxoid mixed with same adjuvants. All mice are bled seven days after the last boosting; sera are then isolated and stored at -20°C.

10 Mouse intranasal challenge model

Ten days after last immunization, sixteen-hour cultures of challenge *Streptococcus pyogenes* strains (1×10^8 to 9×10^8 colony forming units (CFU)), grown in Todd-Hewitt/Yeast broth containing 20% normal rabbit serum and resuspended in 10 ml of PBS, are administered intranasally to 25 g female CD1 (Charles River Breeding Laboratories, Inc., Wilmington, Mass.) or Swiss Webster (Taconic Farms Inc., Germantown, New York) mice. Viable counts are determined by plating dilutions of cultures on blood agar plates.

Each mouse is anesthetized with 1.2 mg of ketamine HCl (Fort Dodge Animal Health, Ft. Dodge, Iowa) by i.p. injection. The bacterial suspension is inoculated to the nostril of anesthetized mice (10 µL per mouse). Sixteen hours after challenge, mice are sacrificed, the noses are removed and homogenized in 3-ml sterile saline with a tissue homogenizer (Ultra-Turax T25, Janke & Kunkel Ika-Labortechnik, Staufen, Germany). The homogenate is 10-fold serially diluted in saline and plated onto blood agar plates containing 200 mg of streptomycin per ml. After overnight incubation at 37°C, β-hemolytic colonies on plates are counted. All challenge strains are marked by streptomycin resistance to distinguish them from β-hemolytic bacteria that may persist in the normal flora.

Subcutaneous mouse challenge model

Five-week-old (20- to 30-g) outbred, immunocompetent, hairless male mice (strain Crl:SKH1-hrBR) (Charles River, Wilmington, Massachusetts) are used for subcutaneous injection. Tissue samples are collected following humane euthanasia.

WO 02/083859

PCT/US02/11610

Streptococcus pyogenes cells, grown as described in Example 1, are harvested and washed once with sterile ice-cold, pyrogen-free phosphate-buffered saline (PBS). The optical density at 600 nm (OD_{600}) is adjusted to give the required inoculum. *Streptococcus pyogenes* (1×10^8 CFU) contained in 0.1 ml are injected subcutaneously in the right flank of each animal with a tuberculin syringe. Control mice are treated with the same volume of PBS. The number of CFU inoculated per mouse is verified for each experiment by colony counts on tryptose agar plates containing 5% sheep blood (Becton Dickinson, Cockeysville, Md.). The mice are observed for 21 days after challenge. Blood is collected from each dead animal by cardiac puncture and cultured on blood agar plates.

10 Tissue collection and histology

Prior to inoculation, the animals are assigned to groups with a random number generator, and blood samples are drawn to establish baseline hematologic data. Blood and tissue samples are collected at 24, 48, and 72 h after inoculation. The methods used for blood and tissue collection are identical for all time points.

15 Blood samples are obtained from the retro-orbital sinus of the animals, and complete blood count analysis is performed with a Technicon H*1 (Tarrytown, N.Y.) hematology analyzer with species-specific software. Skin samples are collected by wide marginal excision around the abscess or the injection site. These samples always include tissue from the injection site and contiguous grossly normal tissue for comparison. Care is taken to
20 preserve the anatomic orientation of the samples. Tissue samples are also obtained from the heart, liver, spleen, and lung.

All tissues are fixed in 10% neutral buffered formalin supplemented with zinc chloride (Antech, Ltd., Battle Creek, Michigan). Whole lungs are first infused with formalin and then, along with the other organs, fixed by submersion. The samples are placed in
25 formalin for 18 to 24 h and then transferred to 70% ethyl alcohol prior to processing. Standard histologic methods of dehydration in ascending grades of ethyl alcohol, clearing in xylene, and paraffin infiltration are employed. The paraffin blocks are processed with a rotary microtome to obtain 4- μ m sections. The histologic sections are stained with hematoxylin and eosin and mounted. Selected tissues are sectioned and stained with a tissue
30 Gram stain.

WO 02/083859

PCT/US02/11610

Mouse measurements

Mice are weighed immediately before GAS inoculation. The animal weight and abscess sizes are measured 12 h after inoculation and daily thereafter for the first week. Animals are then observed at weekly intervals for a total of 21 days. The dimensions of the abscesses are measured with a caliper; length (L) and width (W) values were used to calculate abscess volume [$V = 4/3\pi(L/2)^2 \times (W/2)$] and area [$A = \pi(L/2) \times (W/2)$], employing equations for a spherical ellipsoid.

EXAMPLE 3

Seventy-seven ORFs were initially selected for characterization by "wet chemistry".

- Aspects of these studies included: 1) the ability of specific mouse polyclonal sera generated against each purified protein to react to the surface of the bacterium as measured by whole-cell ELISA, 2) the ability of these same sera to react to the bacterial cell surface during log phase or stationary phase growth as determined by LV-SEM, 3) the genetic conservation of the genes across strains (M serotypes) of *S. pyogenes* as well as other species of streptococci that include the groups C and G, 4) phenotypic expression of specific proteins by these strains as determined by dot blot, 5) expression of the genes of interest at the transcriptional level by quantitative PCR (qPCR), and 6) the ability of human antibody to these proteins to be opsonic in an *in vitro* opsonophagocytic assay.

- Seventy-four of the ORFs have been cloned and expressed in *E. coli*, and 62 of the expressed proteins have been purified. These purified proteins were injected into mice for the generation of the specific antibody for which the analysis by whole-cell ELISA and LV-SEM has been completed. Additionally, 24 ORFs have been evaluated for genetic conservation across *S. pyogenes* strains and streptococcal species; a few have been evaluated for expression at the transcriptional level by qPCR *in vitro* and *in vivo*. Lastly, human antibody specific for *S. pyogenes* proteins has been purified and evaluated in opsonophagocytic assays.

Whole-cell enzyme-linked immunosorbent assay (ELISA)

- S. pyogenes* strain SF-370 was used to inoculate Todd-Hewitt broth containing 0.5% yeast extract (THY), and was cultured overnight at 37°C. Cells were harvested by centrifugation and washed two times with phosphate buffered saline (PBS). The bacteria

WO 02/083859

PCT/US02/11610

were resuspended in PBS to an OD₆₀₀ of 0.2 with PBS and each well of a 96 well polystyrene microtiter plate was coated with 100 µl of the bacterial suspension. The plates were then air-dried at room temperature, sealed with a mylar plate sealer and stored at 4° C inverted for up to three months. In preparation for the assay, the plates were washed three times with Tris Buffered Saline (TBS)/0.1% Brij-35, 100 µl/well of ORF-specific antisera was added to each well, and incubated at 37° C for two hours. The plates were then washed three times with TBS/0.1% Brij-35, 100 µl/well of the secondary antibody conjugate was added to each well, and incubated for one hour at room temperature. Finally, after three washes with PBS, 100 µl/well of the substrate was added to each well and allowed to develop for 60 minutes at room temperature. The reaction was then stopped by adding 50 µl/well of 3N NaOH. Absorbance values (OD₄₀₅) were determined using an ELISA plate reader.

Polymerase chain reaction (PCR) analysis of genetic conservation.

The bacterial strains tested included ten from *S. pyogenes*, SF370 (M1), 90-226 (M1), 80-003 (M1), CS210 (M2), CS194 (M4), 83-112 (M5), CS204 (OF+, M11, T11), CS24 (M12), 95-0061 (M28), CS101 (M49), and a fourth M1 serotype SpeB+, two *S. zooepidemicus* strains, CS258 and GB21, and three group G streptococcal strains, CS241, CS140, and CS242. Five ml overnight cultures were grown in THY. Two and one-half ml of each culture were centrifuged and resuspended in 480 µl of 50 mM EDTA, 120 µl of 10 mg/ml lysozyme and 2 µl of 2500 unit/ml mutanolysin. Samples were incubated at 37° C for one hour. Promega's Wizard Genomic DNA Purification Kit was followed for the remainder of the genomic purifications. Primer sets for the full-length genes and secondly, primers designed for qPCR (see below) were used in the assay. PCR cycling conditions are as follows: 94°C hold for one minute, 16 cycles of 94°C for 15 seconds and 58°C for 10 min, 12 cycles, each increasing 15 seconds from the previous, of 94°C for 15 seconds and 58°C for 10 min, a ten minute hold at 72°C, and finally a 4°C hold. PCR products were verified by mobility in agarose gels. Any amplification containing an intense band of the appropriate size was considered to be a positive result.

Quantitative PCR (qPCR)

RNA was isolated from bacterial cultures described above or from infected homogenized mouse tissue. Samples were suspended in 2 ml RNeasy Lysis Buffer (Qiagen, Crawley, UK).

WO 02/083859

PCT/US02/11610

TX, USA) and quick-frozen using dry-ice/ethanol and stored at -70°C until use. Samples were thawed to room temperature and then frozen again using the above method, for a total of three freeze-thaw cycles. Samples were either treated with $100\ \mu\text{l}$ 10mg/ml lysozyme and $10\ \mu\text{l}$ 2500 unit/ml mutanolysin, and incubated at 37°C for one hour, or samples were mixed with an equal volume of $0.1\ \text{mm}$ glass beads and placed into the bead beater for one minute at $4800\ \text{rpm}$ to lyse the cells. Supernatant was recovered from the beads and an additional $400\ \mu\text{l}$ *RNAlater* was added to the beads and mixed as above. Supernatants recovered from beads or digested solution were mixed with an equal volume of *RNAqueous* Lysis/Binding Solution (Ambion) and vortexed vigorously. Samples were spun at top speed in a microcentrifuge for two minutes to pellet any remaining tissue. The supernatants were mixed with an equal volume of 64% ethanol and passed through a filter cartridge, $700\ \mu\text{l}$ at a time. Filter cartridges were washed as described in the *RNAqueous* manual. Samples were eluted using $2 \times 25\ \mu\text{l}$ 95°C Elution Solution. Two, $1.5\ \mu\text{l}$ DNase treatments were performed for one hr each at 37°C using DNA-free (Ambion) to remove any genomic contamination. Twenty μl of purified RNA was used in $40\ \mu\text{l}$ final volume RT reaction with heat denaturation as described in *RETROscript* (Ambion) protocol to generate cDNA. Samples were denatured at 85°C , and reverse transcribed by incubating for one hour at 42°C , followed by a ten minute incubation at 92°C .

Quantitative PCR was performed using primers and probes, specific to each ORF, designed using Primer Express software (Applied Biosystems, Foster City, CA, USA). Twenty-five μl reactions were set up using $2 \times$ Taqman Universal PCR Master Mix (Applied Biosystems), $300\ \text{nM}$ forward primer, $300\ \text{nM}$ reverse primer, $200\ \text{nM}$ FAM/TAMRA probe, and cDNA template. PCR reaction was as follows: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 seconds and 60°C for one minute. Ribosomal 16S RNA is used as an internal control, with all results being normalized to the 16S Ct value. Based upon results from a standard curve, the cDNA added to these wells was diluted 100 fold to produce a Ct value similar to ORFs of interest.

Purification of human polymorphonuclear leukocytes (PMN).

PMNs were purified from a pool of human whole blood from four donors using a Percoll gradient. A three-layer gradient was prepared by diluting Percoll in Hank's Balanced Salt Solution (HBSS). The densest phase was 2.7:1, middle was 1.079:1 and upper phase

WO 02/083859

PCT/US02/11610

1.07:1, Percoll:HBSS respectively. A ten ml volume of whole blood was layered onto the gradient and centrifuged at 2600 RPM for 20 minutes at 20° C. The upper layers were removed, washed in PBS with glucose to remove Percoll, centrifuged and resuspended in sterile water to lyse red blood cells. A twenty-fold concentrated solution of normal saline was added to equilibrate, re-centrifuged to remove lysed cells, the PMNs were resuspended and counted. The cells were diluted into PBS containing calcium and magnesium and brought to 37° C before use.

Blot analysis of ORF specific antibodies from human sera.

Two µg of protein were coated onto nitrocellulose and allowed to air dry for 15 minutes. The blot was incubated in BLOTTO for 30 minutes at room temperature and then incubated with 5 ml of pooled human serum plasma at 4° C for 16 hours. The nitrocellulose was rinsed in PBS with 0.2% Tween 20 and incubated with goat anti-human IgG conjugated to alkaline phosphatase for two hr at room temperature. The blot was re-washed and developed in NBT/BCIP substrate.

Affinity purification of human antibodies.

One hundred µg of each *S. pyogenes* purified protein was allowed to adhere to a strip of nitrocellulose, blocked for 15 minutes with 5% BLOTTO and then rinsed with PBS. After the sera was adsorbed overnight at 4° C, the nitrocellulose strip was washed with PBS and rinsed with 100 mM glycine at pH 3.0 to elute bound antibodies. The eluted antibodies were neutralized with 1 M Tris pH 8.8 and dialyzed in PBS. These antibodies were tested with PMNs and human whole blood for OPA to the SF-370 strain.

Opsonophagocytic assay (OPA).

S. pyogenes strain SF-370 was used to inoculate THY broth and grown static overnight. The overnight cultures were diluted into fresh medium and further cultured to an OD₆₅₀ of 0.5-0.7. The cells were centrifuged, washed 1X with PBS and resuspended in ice cold PBS to an OD₆₅₀ of 0.5. The cells were diluted to 1:5,000 in PBS and mixed with test antibody or antiserum for 30 min at 4° C. Pre-warmed PMNs were added to the bacteria and antibody at a ratios of 100 and 200 effector cells per target cell. The reactions were incubated at 37° C for one hr on a rocker and finally stopped with ice cold PBS and plated in duplicate on BHI agar.

WO 02/083859

PCT/US02/11610

OPA using whole human blood.

Individual heparin-treated human blood was obtained and incubated at 37° C for 15-30 min until used. Bacteria were prepared as described, and incubated with 50 µl test antibody at 4° C for 15 min, then 430 µl of whole blood were added. The reactions were
 5 incubated for 1.5 hr at 37° C on rocker and plated in duplicate on BHI agar. Each experiment represents an individual person's whole blood sample, not a pool.

ResultsWhole cell ELISA.

The ability of ORF-specific antibody to react to the surface of whole cells was tested
 10 by ELISA. The antibody was produced in mice as described previously. Reactivity demonstrates differences in the amount of protein expressed on the surface of the *S. pyogenes* cells and/or the exposure of the protein in a manner that allows for antibody to bind. ELISA titers are shown in Table XV and indicate a range of reactivities reflective of the differences in either amount of protein expressed or number of epitopes exposed to allow for antibody
 15 reactivity. Values well above preimmune background titers are in bold face type.

Table XV. Whole cell ELISA titer to *S. pyogenes* ORFs.

Orf #	ELISA Titer	Orf #	ELISA Titer
68	1,635	1358	6,201
73	1,702	1487	4,007
145	2,105	1659	3,240
218	1,139	1664	5,355
232	1,277	1698	2,032
309	1,456	1723	1,273
347	2,766	1788	3,324
433	1,431	1789	1,475
554	22,873	1818	40,271
661	1,727	1820	2,498
668	1,869	1878	895
678	2,144	1983	1,179
685	3,094	2015	1,800

WO 02/083859

PCT/US02/11610

704	1,716	2019	24,669
721	680	2064	1,486
729	1,381	2258	4,962
747	11,733	2379	19,220
850	4,861	2417	4,225
967	4,823	2450	4,255
1157	1,827	2452	2,256
1191	1,248	2459	2,166
1202b	1,194	2477	5,412
1218	220,289	2497	666
1224	21,170	2593	8,602
1284	1,374	2601	2,000
1316	6,407		

Gene conservation

PCR analysis of several streptococcal strains was performed to determine the extent of conservation of the various ORFs. The results from this analysis can be seen in Fig. 11.

- 5 All PCR products were analyzed by gel electrophoresis and the band size compared to the predicted value. All ORFs indicated as positive showed a PCR product migrating at the predicted size. The data show a high degree of genomic conservation, with 21 out of 24 ORFs tested being conserved across all eleven strains of *S. pyogenes*. Additionally, 18 were conserved amongst groups C and G; the lowest amount of conservation was observed in the
- 10 strains of group B streptococci.

Quantitative PCR of selected *S. pyogenes* ORFs.

Quantitative PCR was performed to verify transcription of several ORFs contained in the *S. pyogenes* genome. Further, this method was used as a means to verify gene expression *in vivo* in a simulated infection model. Two known transcriptional regulators, *rofA* and *Mga*, and one other housekeeping gene, *gyrA*, were included as additional controls. All genes

15 tested were expressed, and depending on conditions, some showed a variation in levels of transcription. The values are expressed in Ct numbers, which indicate at which PCR cycle the amplification was detectable above background. Thus, a lower Ct value indicates that a

WO 02/083859

PCT/US02/11610

greater amount of mRNA was present in the starting material. A Ct difference of one correlates to a two-fold difference in the amount mRNA detected. Figure 12 shows the results of this analysis. All ORFs showed a significantly lower Ct value than the no template control. ORF 2019 showed a 155-fold lower expression in the thigh than that observed in either the lung or *in vitro* culture. ORF 2477, on the other hand, showed a 49-fold increase, relative to the thigh or *in vitro* culture, in mRNA levels when extracted from the lung after 8 hours of infection. These data show that all ORFs tested were transcribed *in vitro* and *in vivo* and were influenced by the conditions in which the bacteria are exposed.

Reactivity of human sera to *S. pyogenes* proteins.

Antibodies were purified from human sera to test the ability of ORF specific antibody to enhance the ability of PMNs to engulf and kill *S. pyogenes*. Figure shows the reactivity of human serum to several *S. pyogenes* proteins by dot blot indicating that this serum is suitable as a source of antibodies for opsonophagocytic studies. Table XVI summarizes the results of these blots. The results of the blot indicate that 14 of the 24 ORF proteins tested positive for reactivity with human serum. In a similar experiment, a single human serum was tested against the proteins and the results were identical to the ones shown in Table XVI. Several of the proteins were selected for use in the affinity purified antibody studies based on their reactivity and quantity of available material.

Table XVI. ORF identification for reactive proteins.

	A	B	C	D	E	F	G	H
1	ScpA	145	232	554	668	721	1224	1284
2	2452	1659	1698	1788	1818	1820	2379	2459
3	2477	2593	2601	1218	433	1358	2019	1664

Notes: Bold=positive

Opsonophagocytic activity of affinity purified human anti-ORF antibodies with purified PMNs.

PMNs were purified from a pool of four human blood samples and the growth of *S. pyogenes* SF-370 were as described above. Bacteria, PBS diluent and PMNs served as a negative control. The percent killing was calculated by dividing CFUs recovered from reaction containing test antibody with CFUs recovered from the reaction containing that of

WO 02/083859

PCT/US02/11610

the negative control. The results of these studies, summarized in Table XVII, indicate that the affinity-purified antibodies have opsonic activity to SF-370 when incubated with purified PMNs. In particular, antibodies to ScpA and ORF 1224 resulted in greater than 50% killing as measured in OPA verses negative control all three times they were tested.

- 5 Table XVII. Opsonophagocytic activity of affinity purified human antibodies to *S. pyogenes* proteins with purified PMNs as effector cells.

	Opsonophagocytic Killing of ORF Antibodies (Percent) ¹					
	ScpA	1224	1218	145	2459	1698
Exp. #1	60	64	63	ND	ND	ND
Exp. #2	65	53	59	ND	ND	ND
Exp. #3	62	85	45	71	31	61
Avg.	62.3	67.3	55.7	71	31	61

¹Opsonophagocytic activity as compared to negative control. Ratio of PMNs to bacteria was 100:1. Affinity purified antibody was 10% of the reaction mixture (1:10 dilution).

ND = No data.

- 10 Opsonophagocytic activity of affinity purified human antibodies using whole blood.

Traditional OPAs with *S. pyogenes* have utilized whole blood as the source of effector cells. Experiments were conducted to determine if the affinity-purified antibodies had opsonic activity in the presence of whole blood. The results are summarized in Table XVIII and show variable results depending on the individual whose blood was used as a source for

- 15 PMNs. However, antibodies to ORF1224 and 145 gave consistently greater OPA titers with all seven of the individual blood samples tested. In contrast, antibodies to ScpA generated consistently poor OPA titers with all seven blood samples. This was unexpected because when antibodies to ScpA were tested with PMNs there was greater than 50% killing in 3 of 3 assays. Antibodies to the five other proteins had less consistent OPA against *S. pyogenes* SF-
20 370 to the homologous strain. It should be noted that antibodies to ORF 1284 generated greater than 50% killing in 4 of 7 experiments.

Table XVIII. OPA using whole blood as source of effector cells.

Person	Opsonophagocytic Killing of ORF Antibodies (Percent) ¹							
	<i>S</i> <i>cpA</i>	145	1224	1284	1698	1818	2459	1218
1	16	77	86	60	56	45	82	56

WO 02/083859

PCT/US02/11610

2	36	50	79	86	68	72	64	28
3	16	47	56	53	39	42	66	33
4	14	48	54	41	25	63	62	33
5	19	69	56	35	63	42	19	42
6	7	57	68	54	62	54	65	36
7	5	64	59	42	33	38	19	16
Mean	14	58	64	51	32	50	47	33
Std Dev	10	12	13	17	20	13	25	12

¹Oponophagocytic activity as compared to reaction containing whole blood, bacteria and PBS.

EXAMPLE 4 - BIOLOGICAL ACTIVITIES OF STREPTOCOCCAL PYROGENIC EXOTOXIN I

A study was undertaken to characterize SPE I with regard to biological activities. The data indicate that SPE I has superantigen activity and nonspecifically induces proliferation of T cells displaying T cell receptor V β regions (TCR V β) 6.7, 9, and 21.3.

SPE I

SPE I was purified by combinations of isoelectric focusing and affinity chromatography. The purified toxin was shown to be homogeneous by sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Superantigenicity Assay

Rabbit splenocytes were seeded into the wells of a 96 well microtiter plate at a concentration of 2×10^5 cells per well. Ten fold dilutions of toxin were added to wells in quadruplicate, starting with 1.0 μ g/well down to 10^{-8} μ g/well. These dilutions were compared to cells incubated in the presence of PBS alone as a negative control and other SPEs as positive controls. The splenocytes were grown at 37°C for 3 days, and pulsed with 1 μ Ci ^{3} H-thymidine overnight. The cells were harvested the next day, and cell proliferation, as determined by ^{3} H-thymidine incorporation into DNA, was measured in a scintillation counter (Beckman Instruments, Fullerton, CA).

Flow Cytometric Analysis of T cell repertoire

WO 02/083859

PCT/US02/11610

Peripheral blood mononuclear cells (PBMC) obtained from 3 normal human donors were isolated from heparinized venous blood by density gradient sedimentation over Ficoll-Hypaque (Histopaque, Sigma). Cells were then washed three times in Hank's balanced salt solution (HBSS) (Mediatech Cellgro, Herndon, VA) and resuspended in medium for cell culture. PBMC (at 1×10^6 cells/ml) were cultured in RPMI 1640 (Mediatech Cellgro) supplemented with 10% heat inactivated fetal calf serum (FCS) (Gemini Bioproducts, Woodland, CA), 20 mM HEPES buffer (Mediatech Cellgro), 100 u/ml penicillin (Mediatech Cellgro), 100 ug/ml streptomycin (Mediatech Cellgro), and 2 mM L glutamine (Mediatech Cellgro). Cells were cultured in the presence of either anti-CD3 (20ng/ml), or SPE I (100 ng/ml) for 3 days, washed and allowed to grow for an additional day in the presence of interleukin 2 (50 U/ml) before washing and staining for immunofluorescence analysis of T cell repertoire as previous described.

For flow cytometry studies, PBMC were washed in HBSS and resuspended at 10×10^6 cells/ml in a staining solution [PBS with 5% FCS (Gemini Bioproducts), 1% immunoglobulin (Alpha Therapeutic Corp., Los Angeles, CA), 0.02% sodium azide (Sigma)]. Cells were stained in 96 well, round bottomed plates with a panel of biotinylated monoclonal antibodies against human TCRV β 2, 3, 5.1, 5.2, 7, 8, 11, 12, 13.1, 13.2, 14, 16, 17, 20, 21.3, 22 (Immunotech, Westbrook, ME), TCRV β 9, 23 (Pharmingen, San Diego, CA) and TCRV β 6.7 fluorescein isothiocyanate (FITC) (Endogen, Woburn, MA), then incubated for 30 min at 37°C in the dark. After the incubation period, cells were washed twice with washing buffer [PBS, 2% FCS (Gemini Bioproducts), 0.02% sodium azide (Sigma)] by centrifugation at 300xg for 5 min at 4°C. Cell pellets were resuspended in staining solution and incubated with anti-CD3 allophycocyanin (APC), anti-CD4 phycoerythrin (PE) (Becton Dickinson, San Jose, CA), anti-CD8 (FITC) (Becton Dickinson) and a streptavidin peridinin chlorophyll protein (PerCP) conjugate (Becton Dickinson) for 30 min at 4°C. Stained cells were again washed twice in washing buffer and once in 0.02% sodium azide (Sigma) in PBS, by centrifugation at 300xg for 5 min at 4°C. Finally, the cells were fixed in 200 ul of 1% (v/v) formaldehyde (Polysciences, Warrington, PA) in PBS. Analysis was performed using four color flow cytometry (FACS Calibur, Becton Dickinson) as described previously. Methods of cytometer set up and data acquisition have also been described previously. List mode multiparameter data files (each file with forward scatter, side scatter, and 4 fluorescent parameter) were analyzed using the Cellquest program (Becton Dickinson). Analysis of activated populations was performed with the light scatter gate set on the T cell blast

WO 02/083859

PCT/US02/11610

population. Negative control reagents were used to verify the staining specificity of experimental antibodies.

Miniosmotic Pumps

Six American Dutch belted rabbits in groups of 3 were implanted with subcutaneous miniosmotic pumps on the left flanks, containing 500ug of SPE I or 200ug of TSST-1. Lethality of the toxins was assessed over a period of 15 days.

Results

SPE I was evaluated for ability to induce rabbit splenocyte proliferation in a four day assay, as measured by incorporation of ³H thymidine into DNA (Fig. 14). SPE I was comparably mitogenic as the control SPE toxins also included in the figure. The complete fall-off of mitogenic activity for SPE I was between 10⁻⁶ and 10⁻⁷ug/well, similar to that observed for other toxins.

SPE I significantly stimulated human T cells bearing TCR Vβs 6.7, 9, and 21.3 (Fig. 15) compared to cells stimulated with anti-CD3 antibodies, consistent with SPE I being a superantigen. Some T cell populations, for example T cells with TCR Vβ 14 or 17 were significantly reduced compared to cells stimulated with anti-CD3 antibodies.

The majority of pyrogenic toxin superantigens are lethal when administered to rabbits at a toxin concentration between 200 and 500ug in subcutaneously implanted miniosmotic pumps. SPE I did not exhibit this property at the 500ug dose (3/3 survived). In contrast 200ug of TSST-1 was completely lethal (3/3 succumbed).

Discussion

Pyrogenic toxin superantigens are defined by their abilities to induce T lymphocyte proliferation nonspecifically but dependent on the composition of the variable part of the beta chain of the T cell receptor (6). Thus for example, TSST-1 will stimulate proliferation of any human T cell bearing TCR Vβ2, without regard for the antigenic specificity of the responding T cells. This high level of stimulation leads to massive release of cytokines from both T cells and macrophages. Of particular importance is the release of tumor necrosis factors α and β that cause the hypotension and shock associated with TSS.

WO 02/083859

PCT/US02/11610

The data show that SPE I stimulates T cells as a superantigen. Thus, SPE I causes human peripheral blood mononuclear cells to proliferate that contain TCR V β 6.7, 9, and 21.3. This elevation of these selected T cell populations, with the concurrent relative reduction of non-stimulated T cells, is the hallmark signal of SPE I and is referred to as V β skewing.

- 5 In addition, many pyrogenic toxin superantigens are lethal when administered to rabbits in subcutaneously implanted miniosmotic pumps, as a model for TSS (8). These pumps are designed to release a constant amount of toxin over a period of 7 days. The experiments continue for 15 days, however, since rabbits may succumb to the administered toxin for up to that period of time. SPE I was not lethal in this model of TSS. Although many pyrogenic
- 10 toxin superantigens are lethal in this assay, there are notable exceptions. For example, the newly identified staphylococcal enterotoxins L and Q are not lethal in this model, yet these two toxins share all other activities expected of the family (including superantigenicity). For these latter toxins, it has been suggested that they either are not stable in the miniosmotic pumps for the entire 7 day toxin release period or precipitate in the pumps. Accordingly, SPE
- 15 I shares defining superantigenic property of pyrogenic toxin superantigens.

Although illustrated and described above with reference to specific embodiments, the invention is nevertheless not intended to be limited to the details shown. Rather, various modifications may be made in the details within the scope and range of equivalents of the claims and without departing from the spirit of the invention.

WO 02/083859

PCT/US02/11610

Bibliography

1. 1997. Case definitions for Infectious Conditions Under Public Health Surveillance. CDC.
2. Alm, R. A., L. S. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. deJonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori* [published erratum appears in Nature 1999 Feb 25;397(6721):719]. Nature. 397:176-80.
- 10 3. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-402.
4. Anderson, T. F. 1951. Techniques for the preservation of three-dimensional structure in preparing specimens for the electron microscope. Trans N Y Acad Sci. 13:130-134.
- 15 5. Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27:573-80.
6. Chen, C. C., and P. P. Cleary. 1989. Cloning and expression of the streptococcal C5a peptidase gene in *Escherichia coli*: linkage to the type 12 M protein gene.
- 20 Infect. Immun. 57:1740-1745.
7. Chmouryguina, I., A. Suvorov, P. Ferrieri, and P. P. Cleary. 1996. Conservation of the C5a peptidase genes in group A and B streptococci. Infect. Immun. 64:2387-2390.
8. Cockerill, F. R., 3rd, R. L. Thompson, J. M. Musser, P. M. Schlievert, J. Talbot, K. E. Holley, W. S. Harmsen, D. M. Ilstrup, P. C. Kohnen, M. H. Kim, B. Frankfort, J. M. Manahan, J. M. Steckelberg, F. Roberson, and W. R. Wilson. 1998. Molecular, serological, and clinical features of 16 consecutive cases of invasive streptococcal disease. Southeastern Minnesota Streptococcal Working Group. Clin Infect Dis. 26:1448-58.
- 25 9. Courtney, H. S., Y. Li, J. B. Dale, and D. L. Hasty. 1994. Cloning, sequencing, and expression of a fibronectin/fibrinogen-binding protein from group A streptococci. Infect Immun. 62:3937-46.
- 30

WO 02/083859

PCT/US02/11610

10. Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Engineering*. 10:673-6.
11. Cunningham, M. W., and A. Quinn. 1997. Immunological crossreactivity
5 between the class I epitope of streptococcal M protein and myosin. *Adv Exp Med Biol*. 418:887-92.
12. Dale, J. B., R. W. Baird, H. S. Courtney, D. L. Hasty, and M. S. Bronze. 1994. Passive protection of mice against group A streptococcal pharyngeal infection by lipoteichoic acid. *J Infect Dis*. 169:319-23.
- 10 13. Dale, J. B., M. Simmons, E. C. Chiang, and E. Y. Chiang. 1996. Recombinant, octavalent group A streptococcal M protein vaccine. *Vaccine*. 14:944-8.
14. Dale, J. B., R. G. Washburn, M. B. Marques, and M. R. Wessels. 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. *Infect Immun*. 64:1495-501.
- 15 15. Eddy, S. R. 1996. Hidden Markov models. *Cur Opin Struct Bio*. 6:361-5.
16. Ellen, R. P., and R. J. Gibbons. 1972. M protein-associated adherence of *Streptococcus pyogenes* to epithelial surfaces: prerequisite for virulence. *Infect Immun*. 5:826-830.
17. Eng, J. K., A. L. McCormack, and J. R. Yates, 3rd. 1994. An approach to
20 correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *Am Soc Mass Spectrometry*. 5:976-89.
18. Fischetti, V. A., V. Pancholi, and O. Schneewind. 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. *Mol Microbiol*. 4:1603-5.
- 25 19. Fogg, G. C., and M. G. Caparon. 1997. Constitutive expression of fibronectin binding in *Streptococcus pyogenes* as a result of anaerobic activation of *rofA*. *J Bacteriol*. 179:6172-80.
20. Foster, T. J., and M. Hook. 1998. Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol*. 6:484-8.

WO 02/083859

PCT/US02/11610

21. Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, R. Lathigra, O. White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J. F. Tomb, R. D. Fleischmann, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. van Vugt, N. Palmer, M. D. Adams, J. Gocayne, J. C. Venter, and et al. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi* [see comments]. *Nature*. 390:580-6.
22. Hacker, J., G. Blum-Oehler, I. Muhldorfer, and H. Tschape. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol*. 23:1089-97.
- 10 23. Hanski, E., and M. Caparon. 1992. Protein F, a fibronectin-binding protein, is an adhesion of the group A streptococcus *Streptococcus pyogenes*. *Proc Natl Acad Sci., USA*. 89:6172-76.
24. Hanski, E., P. A. Horwitz, and M. G. Caparon. 1992. Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. *Infect Immun*. 60:5119-5125.
- 15 25. Hernandez-Sanchez, J., J. G. Valadez, J. V. Herrera, C. Ontiveros, and G. Guarneros. 1998. lambda bar minigene-mediated inhibition of protein synthesis involves accumulation of peptidyl-tRNA and starvation for tRNA. *EMBO Journal*. 17:3758-65.
- 20 26. Huang, T. T., H. Malke, and J. J. Ferretti. 1989. The streptokinase gene of group A streptococci: cloning, expression in *Escherichia coli*, and sequence analysis. *Mol Microbiol*. 3:197-205.
27. Hynes, W. L., A. R. Dixon, S. L. Walton, and L. J. Aridides. 2000. The extracellular hyaluronidase gene (*hylA*) of *Streptococcus pyogenes*. *FEMS Microbiol Lett*. 184:109-12.
- 25 28. Hynes, W. L., L. Hancock, and J. J. Ferretti. 1995. Analysis of a second bacteriophage hyaluronidase gene from *Streptococcus pyogenes*: evidence for a third hyaluronidase involved in extracellular enzymatic activity. *Infect Immun*. 63:3015-20.
29. Isberg, R. R., and G. Tran Van Nhieu. 1994. Binding and internalization of microorganisms by integrin receptors. *Trends Microbio*. 2:10-4.
- 30

WO 02/083859

PCT/US02/11610

30. Jones, K. F., and V. A. Fischetti. 1988. The importance of the location of antibody binding on the M6 protein for opsonization and phagocytosis of group A M6 streptococci. *J Exp Med.* 167:1114-23.
31. Kihlberg, B. M., M. Collin, A. Olsen, and L. Bjorck. 1999. Protein H, an antiphagocytic surface protein in *Streptococcus pyogenes*. *Infect Immun.* 67:1708-14.
32. Koebnik, R. 1995. Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins [letter; comment]. *Molecular Microbiology.* 16:1269-70.
33. Loessner, M. J., S. Gaeng, and S. Scherer. 1999. Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of *Staphylococcus aureus* bacteriophage 187. *J Bacteriol.* 181:4452-60.
34. Lukashin, A. V., and M. Borodovsky. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107-15.
35. Lukomski, S., C. A. Montgomery, J. Rurangirwa, R. S. Geske, J. P. Barrish, G. J. Adams, and J. M. Musser. 1999. Extracellular cysteine protease produced by *Streptococcus pyogenes* participates in the pathogenesis of invasive skin infection and dissemination in mice. *Infect Immun.* 67:1779-88.
36. Madore, D. V. 1998. Characterization of immune response as an indicator of *Haemophilus influenzae* type b vaccine efficacy. *Pediatr Infect Dis J.* 17:S207-10.
37. Matsuka, Y. V., S. Pillai, S. Gubba, J. M. Musser, and S. B. Olmsted. 1999. Fibrinogen cleavage by the *Streptococcus pyogenes* extracellular cysteine protease and generation of antibodies that inhibit enzyme proteolytic activity. *Infect Immun.* 67:4326-33.
38. Mazmanian, S. K., G. Liu, H. Ton-That, and O. Schneewind. 1999. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science.* 285:760-3.
39. McAtee, C. P., K. E. Fry, and D. E. Berg. 1998. Identification of potential diagnostic and vaccine candidates of *Helicobacter pylori* by "proteome" technologies. *Helicobacter.* 3:163-9.
40. McAtee, C. P., M. Y. Lim, K. Fung, M. Velligan, K. Fry, T. Chow, and D. E. Berg. 1998. Identification of potential diagnostic and vaccine candidates of *Helicobacter*

WO 02/083859

PCT/US02/11610

pylori by two-dimensional gel electrophoresis, sequence analysis, and serum profiling. Clin Diagn Lab Immunol. 5:537-42.

41. McAtee, C. P., M. Y. Lim, K. Fung, M. Velligan, K. Fry, T. P. Chow, and D. E. Berg. 1998. Characterization of a *Helicobacter pylori* vaccine candidate by proteome techniques. J Chromatogr B Biomed Sci Appl. 714:325-33.
42. Mejlhede, N., J. F. Atkins, and J. Neuhard. 1999. Ribosomal -1 frameshifting during decoding of *Bacillus subtilis* *cdd* occurs at the sequence CGA AAG. J. Bacteriol. 181:2930-7.
43. Molinari, G., S. R. Talay, P. Valentin-Weigand, M. Rohde, and G. S. Chhatwal. 1997. The fibronectin-binding protein of *Streptococcus pyogenes*, SfbI, is involved in the internalization of group A streptococci by epithelial cells. Infect Immun. 65:1357-63.
44. Nakai, K., and M. Kanehisa. 1991. Expert system for predicting protein localization sites in gram-negative bacteria. Proteins. 11:95-110.
45. Navarre, W. W., and O. Schneewind. 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol Mol Biol Rev. 63:174-229.
46. Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering. 10:1-6.
47. Nizet, V., B. Beall, D. J. Bast, V. Datta, L. Kilburn, D. E. Low, and J. C. De Azavedo. 2000. Genetic locus for streptolysin S production by group A streptococcus. Infect Immun. 68:4245-54.
48. Nordstrand, A., W. M. McShan, J. J. Ferretti, S. E. Holm, and M. Norgren. 2000. Allele substitution of the streptokinase gene reduces the nephritogenic capacity of group A streptococcal strain NZ131. Infect Immun. 68:1019-25.
49. Olmsted, S. B., S. L. Erlandsen, G. M. Dunny, and C. L. Wells. 1993. High-resolution visualization by field emission scanning electron microscopy of *Enterococcus faecalis* surface proteins encoded by the pheromone-inducible conjugative plasmid pCF10. J Bacteriol. 175:6229-37.

WO 02/083859

PCT/US02/11610

50. Park, J., and S. A. Teichmann. 1998. DIVCLUS: an automatic method in the
GEANFAMMER package that finds homologous domains in single- and multi-domain
proteins. *Bioinformatics*. 14:144-50.
51. Parkhill, J., M. Achtman, K. D. James, S. D. Bentley, C. Churcher, S. R. Klee,
5 G. Morelli, D. Basham, D. Brown, T. Chillingworth, R. M. Davies, P. Davis, K. Devlin, T.
Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Leather, S. Moule, K. Mungall, M. A. Quail,
M. A. Rajandream, K. M. Rutherford, M. Simmonds, J. Skelton, S. Whitehead, B. G. Spratt,
and B. G. Barrell. 2000. Complete DNA sequence of a serogroup A strain of *Neisseria*
meningitidis Z2491 [see comments]. *Nature*. 404:502-6.
- 10 52. Pierschbacher, M. D., and E. Ruoslahti. 1987. Influence of stereochemistry of
the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J Biol Chem*.
262:17294-8.
53. Pizza, M., V. Scarlato, V. Maignani, M. M. Giuliani, B. Arico, M.
Comanducci, G. T. Jennings, L. Baldi, E. Bartolini, B. Capecchi, C. L. Galeotti, E. Luzzi, R.
15 Manetti, E. Marchetti, M. Mora, S. Nuti, G. Ratti, L. Santini, S. Savino, M. Scarselli, E.
Storni, P. Zuo, M. Broeker, E. Hundt, B. Knapp, E. Blair, T. Mason, H. Tettelin, D. W. Hood,
A. C. Jeffries, N. J. Saunders, D. M. Granoff, J. C. Venter, E. R. Moxon, G. Grandi, and R.
Rappuoli. 2000. Identification of vaccine candidates against serogroup B meningococcus by
whole-genome sequencing [see comments]. *Science*. 287:1816-20.
- 20 54. Podbielski, A., A. Flösdorff, and J. Weber-Heynemann. 1995. The group A
streptococcal virR49 gene controls expression of four structural vir regulon genes. *Infect*
Immun. 63:9-20.
55. Proft, T., S. Louise Moffatt, C. J. Berkahn, and J. D. Fraser. 1999.
Identification and Characterization of Novel Superantigens from *Streptococcus pyogenes*. *J*
25 *Exp Med*. 189:89-102.
56. Pugsley, A. P. 1993. The complete general secretory pathway in gram-
negative bacteria. *Microbiol Rev*. 57:50-108.
57. Quinn, A., K. Ward, V. A. Fischetti, M. Hemric, and M. W. Cunningham.
1998. Immunological relationship between the class I epitope of streptococcal M protein and
30 myosin. *Infect Immun*. 66:4418-24.

WO 02/083859

PCT/US02/11610

58. Reda, K. B., V. Kapur, D. Goela, J. G. Lamphear, J. M. Musser, and R. R. Rich. 1996. Phylogenetic distribution of streptococcal superantigen SSA allelic variants provides evidence for horizontal transfer of ssa within *Streptococcus pyogenes*. *Infect Immun.* 64:1161-5.
59. Salzberg, S. L., A. L. Delcher, S. Kasif, and O. White. 1998. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* 26:544-8.
60. Sonnenberg, M. G., and J. T. Belisle. 1997. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. *Infect Immun.* 65:4515-24.
61. Bateman, A. T., R. Birney, S. P. Durbin, K. L. Howe and E. L. L. Sonnhammer. 2000. The Pfam protein families database. *Nuc. Acids. Res.* 28:263-6.
62. Stevens, D. L. 1995. Streptococcal toxic-shock syndrome: spectrum of disease, pathogenesis, and new concepts in treatment. *Emerg Infect Dis.* 1:69-78.
63. Stockbauer, K. E., L. Magoun, M. Liu, E. H. Burns, Jr., S. Gubba, S. Renish, X. Pan, S. C. Bodary, E. Baker, J. Coburn, J. M. Leong, and J. M. Musser. 1999. A natural variant of the cysteine protease virulence factor of group A streptococcus with an arginine-glycine-aspartic acid (RGD) motif preferentially binds human integrins α 5 β 1 and α 5 β 3 [In Process Citation]. *Proc Natl Acad Sci., USA.* 96:242-7.
64. Ton-That, H., G. Liu, S. K. Mazmanian, K. F. Faull, and O. Schneewind. 1999. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc Natl Acad Sci U S A.* 96:12424-12429.
65. Weldingh, K., I. Rosenkrands, S. Jacobsen, P. B. Rasmussen, M. J. Elhay, and P. Andersen. 1998. Two-dimensional electrophoresis for analysis of *Mycobacterium tuberculosis* culture filtrate and purification and characterization of six novel proteins. *Infect Immun.* 66:3492-500.
66. Yutsudo, T., K. Okumura, M. Iwasaki, A. Hara, S. Kamitani, W. Minamide, H. Igarashi, and Y. Hinuma. 1994. The gene encoding a new mitogenic factor in a *Streptococcus pyogenes* strain is distributed only in group A streptococci. *Infection and Immunity.* 62:4000-4004.

WO 02/083859

PCT/US02/11610

67. Published International Patent Application Number WO99/27944.
68. U.S Patent Number 4,666,829.

WO 02/083859

PCT/US02/11610

What is claimed is:

1. An isolated polypeptide comprising:

(i) an amino acid sequence that has at least 70% identity to an amino acid sequence of any of even numbered SEQ ID NOS: 2-668;

5 (ii) an amino acid sequence of any of even numbered SEQ ID NOS: 2-668;

(iii) an immunogenic fragment of any amino acid sequence of (i) or (ii);

(iv) at least 7 contiguous amino acid residues of any amino acid sequence of (i) or (ii);

or

(v) a biological equivalent of any of (i), (ii), (iii) or (iv) that is effective for preventing
10 or ameliorating β -hemolytic streptococcal colonization or infection in a susceptible subject.

2. The isolated polypeptide of claim 1, wherein the biological equivalent exhibits an opsonophagocytic activity of at least about 15% percent killing of bacteria as measured by decrease in colony forming units (CFU) in OPA versus a negative control.

3. The isolated polypeptide of claim 1, wherein the biological equivalent exhibits an
15 opsonophagocytic activity of at least about 20% percent killing of bacteria as measured by decrease in colony forming units (CFU) in OPA versus a negative control.

4. The isolated polypeptide of claim 1, wherein the biological equivalent exhibits an opsonophagocytic activity of at least about 40% percent killing of bacteria as measured by decrease in colony forming units (CFU) in OPA versus a negative control.

20 5. The isolated polypeptide of claim 1, wherein the biological equivalent exhibits an opsonophagocytic activity of at least about 50% percent killing of bacteria as measured by decrease in colony forming units (CFU) in OPA versus a negative control.

6. The isolated polypeptide of claim 1, wherein the biological equivalent exhibits an opsonophagocytic activity of at least about 60% percent killing of bacteria as measured by
25 decrease in colony forming units (CFU) in OPA versus a negative control.

7. The isolated polypeptide of claim 1, wherein the biological equivalent provides a desired level of protection against β -hemolytic streptococci.

8. The isolated polypeptide of claim 7, comprising any of (i) an amino acid sequence that has at least 70% identity to an amino acid sequence of any of even numbered SEQ ID
30 NOS: 2-668, (ii) an amino acid sequence of any of even numbered SEQ ID NOS: 2-668, (iii)

WO 02/083859

PCT/US02/11610

an immunogenic fragment of any amino acid sequence of (i) or (ii), or (iv) at least 7 contiguous amino acid residues of any amino acid sequence of (i) or (ii).

9. The isolated polypeptide of claim 1, comprising an amino acid sequence that has at least 85% identity to an amino acid sequence of any of even numbered SEQ ID NOS: 2-668

5 10. The isolated polypeptide of claim 1, wherein the biological equivalent provides cross-reactivity across at least two strains of β -hemolytic streptococci.

11. The isolated polypeptide of claim 1, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 6, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 80, SEQ ID NO: 88 and combinations thereof.

10 12. The isolated polypeptide of claim 1, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 64, SEQ ID NO: 128, SEQ ID NO: 140, SEQ ID NO: 182 and combinations thereof.

13. The isolated polypeptide of claim 1, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 32, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 104, SEQ ID NO: 138 and combinations thereof.

14. The isolated polypeptide of claim 1, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 66, SEQ ID NO: 82, SEQ ID NO: 78, SEQ ID NO: 142, SEQ ID NO: 146, SEQ ID NO: 162, SEQ ID NO: 186, SEQ ID NO: 342 and combinations thereof.

20 15. The isolated polypeptide of claim 1, wherein the amino acid sequence is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 34, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 54, SEQ ID NO: 62, SEQ ID NO: 68, SEQ ID NO: 90, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 106, SEQ ID NO: 118, SEQ ID NO: 124, SEQ ID NO: 130, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 376 and combinations thereof.

25 16. The isolated polypeptide of claim 2, wherein the amino acid sequence is SEQ ID NO: 6.

17. The isolated polypeptide of claim 2, wherein the amino acid sequence is SEQ ID NO: 46.

30 18. The isolated polypeptide of claim 2, wherein the amino acid sequence is SEQ ID NO: 48.

WO 02/083859

PCT/US02/11610

19. The isolated polypeptide of claim 2, wherein the amino acid sequence is SEQ ID NO: 80.

20. The isolated polypeptide of claim 2, wherein the amino acid sequence is SEQ ID NO: 88.

5 21. The isolated polypeptide of claim 1, where said isolated polypeptide is a mature polypeptide of an amino acid sequence of any of even numbered SEQ ID NOS: 2-668.

22. An isolated polynucleotide comprising:

(i) a nucleotide sequence that encodes the isolated polypeptide of claim 1;

10 (ii) a nucleotide sequence that encodes the isolated polypeptide of claim 1, where said isolated polypeptide is a mature polypeptide;

(iii) a nucleotide sequence of any of odd numbered SEQ ID NOS: 1-147;

(iv) a nucleotide sequence of any of odd numbered SEQ ID NOS: 149-181;

(v) a nucleotide sequence of any of odd numbered SEQ ID NOS: 183-187;

(vi) a nucleotide sequence of any of odd numbered SEQ ID NOS: 189-667;

15 (vii) a nucleotide sequence that has at least 70% identity to a nucleotide sequence that encodes the polypeptide of claim 1;

(viii) a nucleotide sequence that has at least 70% identity to a nucleotide sequence of any of odd numbered SEQ ID NOS: 1-667;

20 (ix) a nucleotide sequence that hybridizes under stringent hybridization conditions to a nucleotide sequence that encodes the polypeptide of claim 1;

(x) a nucleotide sequence that hybridizes under stringent hybridization conditions to a nucleotide sequence of any of odd numbered SEQ ID NOS: 1-667; or

(xi) a nucleotide sequence that is fully complementary to a nucleotide sequence of any of (i)-(x).

25 23. The isolated polynucleotide of claim 22, wherein the nucleotide sequence is selected from the group consisting of SEQ ID NO: 5, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 79, SEQ ID NO: 87 and combinations thereof.

WO 02/083859

PCT/US02/11610

24. The isolated polynucleotide of claim 22, wherein the nucleotide sequence is selected from the group consisting of SEQ ID NO: 15, SEQ ID NO: 63, SEQ ID NO: 127, SEQ ID NO: 139, SEQ ID NO: 181 and combinations thereof.

25. The isolated polynucleotide of claim 22, wherein the nucleotide sequence is selected from the group consisting of SEQ ID NO: 31, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 103, SEQ ID NO: 137 and combinations thereof.

26. The isolated polynucleotide of claim 22, wherein the nucleotide sequence is selected from the group consisting of SEQ ID NO: 65, SEQ ID NO: 81, SEQ ID NO: 77, SEQ ID NO: 141, SEQ ID NO: 145, SEQ ID NO: 161, SEQ ID NO: 185, SEQ ID NO: 341 and combinations thereof.

27. The isolated polynucleotide of claim 22, wherein the nucleotide sequence is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 7, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 33, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 53, SEQ ID NO: 61, SEQ ID NO: 67, SEQ ID NO: 89, SEQ ID NO: 95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 105, SEQ ID NO: 117, SEQ ID NO: 123, SEQ ID NO: 129, SEQ ID NO: 147, SEQ ID NO: 157, SEQ ID NO: 375 and combinations thereof.

28. The isolated polynucleotide of claim 22, wherein the nucleotide sequence is selected from the group consisting of 25, 131, 147, 149, 151, 153, 155, 159, 163, 165, 169, 171, 173, 175, 177, 179, 183, 187, 215, 243, 301, 327, 331, 463, 541, 579, 617, 619, 665, 669 and combinations thereof.

29. A recombinant host cell comprising a polynucleotide of claim 22.

30. A recombinant expression vector comprising a polynucleotide of claim 22.

31. A recombinant host cell comprising a vector of claim 22.

32. A method for producing a polypeptide comprising:

(a) culturing a recombinant host cell comprising (i) a polynucleotide of claim 22 or (ii) a recombinant expression vector comprising a polynucleotide of claim 22, under conditions suitable to produce the polypeptide encoded by the polynucleotide; and

(b) recovering the polypeptide from the culture.

WO 02/083859

PCT/US02/11610

33. An antibody that binds immunospecifically to a polypeptide of claim 1.

34. The antibody of claim 33, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is selected from the group consisting of SEQ ID NO: 6, SEQ ID NO: 46, SEQ ID NO: 48, SEQ ID NO: 80, SEQ ID NO: 88 and combinations thereof.

35. The antibody of claim 33, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 64, SEQ ID NO: 128, SEQ ID NO: 140, SEQ ID NO: 182 and combinations thereof.

36. The antibody of claim 33, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is selected from the group consisting of SEQ ID NO: 32, SEQ ID NO: 58, SEQ ID NO: 60, SEQ ID NO: 104, SEQ ID NO: 138 and combinations thereof.

37. The antibody of claim 33, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is selected from the group consisting of SEQ ID NO: 66, SEQ ID NO: 82, SEQ ID NO: 78, SEQ ID NO: 142, SEQ ID NO: 146, SEQ ID NO: 162, SEQ ID NO: 186, SEQ ID NO: 342 and combinations thereof.

38. The antibody of claim 33, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 8, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 34, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, SEQ ID NO: 54, SEQ ID NO: 62, SEQ ID NO: 68, SEQ ID NO: 90, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 106, SEQ ID NO: 118, SEQ ID NO: 124, SEQ ID NO: 130, SEQ ID NO: 148, SEQ ID NO: 158, SEQ ID NO: 376 and combinations thereof.

39. The antibody of claim 33, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is selected from the group consisting of 26, 132, 148, 150, 152, 154, 156, 160, 164, 166, 170, 172, 174, 176, 178, 180, 184, 188, 216, 244, 302, 328, 332, 464, 542, 580, 618, 620, 666, 670 and combinations thereof.

40. The antibody of claim 34, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is SEQ ID NO: 6.

WO 02/083859

PCT/US02/11610

41. The antibody of claim 34, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is SEQ ID NO: 46.

42. The antibody of claim 34, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is SEQ ID NO: 48.

5 43. The antibody of claim 34, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is SEQ ID NO: 80.

44. The antibody of claim 34, wherein the antibody binds immunospecifically to a polypeptide having an amino acid sequence which is SEQ ID NO: 88.

10 45. An immunogenic composition comprising an immunogenic amount of a component that comprises a polypeptide of claim 1, wherein said component is in an amount effective to prevent or ameliorate a β -hemolytic streptococcal colonization or infection in a susceptible mammal.

46. The immunogenic composition of claim 45, which comprises at least a portion of said polypeptide conjugated or linked to a peptide, polypeptide, or protein.

15 47. The immunogenic composition of claim 45, which comprises at least a portion of said polypeptide conjugated or linked to a polysaccharide.

48. The immunogenic composition of claim 45, which further comprises a physiologically-acceptable vehicle.

20 49. The immunogenic composition of claim 45, which further comprises an effective amount of an adjuvant.

50. The immunogenic composition of claim 45, wherein the mammal is a human, dog, bovine, porcine, or horse.

51. The immunogenic composition of claim 50, wherein the mammal is human.

25 52. An immunogenic composition comprising an immunogenic amount of a component that comprises a polypeptide of claim 1, wherein the polypeptide is capable of generating antibody that specifically recognizes said polypeptide, and wherein the amount of said component is effective to prevent or ameliorate β -hemolytic streptococcal colonization or infection in a susceptible mammal.

WO 02/083859

PCT/US02/11610

53. An immunogenic composition comprising an immunogenic amount of a component that comprises a polynucleotide of claim 22, wherein said component is in an amount effective to prevent or ameliorate a β -hemolytic streptococcal colonization or infection in a susceptible mammal.

5 54. The immunogenic composition of claim 53, comprising a recombinant expression vector comprising a polynucleotide of claim 22.

55. The immunogenic composition of claim 53, which further comprises a physiologically-acceptable vehicle.

56. The immunogenic composition of claim 53, which further comprises an effective
10 amount of an adjuvant.

57. The immunogenic composition of claim 53, wherein the mammal is a human, dog, bovine, porcine, or horse.

58. The immunogenic composition of claim 57, wherein the mammal is human.

59. The immunogenic composition of claim 53, wherein the β -hemolytic streptococci
15 is group A streptococci, group B streptococci, group C streptococci, or group G streptococci.

60. The immunogenic composition of claim 59, wherein the β -hemolytic streptococci is *Streptococcus pyogenes*.

61. An immunogenic composition comprising:

(i) an isolated polypeptide that is substantially conserved across strains of β -
20 hemolytic streptococci and that is effective in preventing or ameliorating a β -hemolytic streptococcal colonization or infection in a susceptible subject, said isolated polypeptide having at least 70% identity to an amino acid sequence of any of even numbered SEQ ID NOS: 2-668.

(ii) an immunogenic fragment of (i); or

25 (iii) an antibody that immunospecifically binds to (i) or (ii).

WO 02/083859

PCT/US02/11610

62. The immunogenic composition of claim 61, wherein the isolated polypeptide tests positive for reactivity when tested against human serum.

63. The immunogenic composition of claim 61, wherein the β -hemolytic streptococci is group A streptococci, group B streptococci, group C streptococci, or group G streptococci.

5 64. The immunogenic composition of claim 61, wherein the β -hemolytic streptococci is *Streptococcus pyogenes*.

65. The immunogenic composition of claim 61, wherein the isolated polypeptide is at least about 80% conserved across the strains.

10 66. The immunogenic composition of claim 61, wherein the isolated polypeptide is at least about 85% conserved across the strains.

67. The immunogenic composition of claim 61, wherein the isolated polypeptide is at least about 90% conserved across the strains.

68. The immunogenic composition of claim 61, wherein the isolated polypeptide is at least about 95% conserved across the strains.

15 69. A method of protecting a susceptible mammal against β -hemolytic streptococcal colonization or infection comprising administering to the mammal an immunogenic composition that comprises an immunogenic amount of a component that comprises a polypeptide of claim 1, which amount is effective to prevent or ameliorate β -hemolytic streptococcal colonization or infection in the susceptible mammal.

20 70. The method of claim 69, wherein the immunogenic composition comprises at least a portion of said polypeptide, optionally conjugated or linked to a peptide, polypeptide, or protein.

71. The method of claim 69, wherein the immunogenic composition comprises at least a portion of said polypeptide, optionally conjugated or linked to a polysaccharide.

25 72. A method of protecting a susceptible mammal against β -hemolytic streptococcal colonization or infection comprising administering to the mammal an effective amount of an immunogenic composition comprising a polypeptide of claim 1, wherein the polypeptide is

WO 02/083859

PCT/US02/11610

capable of generating antibody specific to said polypeptide, and wherein the amount is effective to prevent or ameliorate β -hemolytic streptococcal colonization or infection in the susceptible mammal.

73. The method of claim 72, wherein the polypeptide comprises the mature
5 polypeptide of an amino acid sequence of any of SEQ ID NOS: 2-670.

74. The method of claim 72, wherein the immunogenic composition further
comprises a physiologically-acceptable vehicle.

75. The method of claim 72, wherein the immunogenic composition is administered
by subcutaneous or intramuscular injection.

10 76. The method of claim 72, wherein the immunogenic composition is administered
by oral ingestion.

77. The method of claim 72, wherein the immunogenic composition is administered
intranasally.

78. The method of claim 72, wherein the β -hemolytic streptococci is group A
15 streptococci, group B streptococci, group C streptococci, or group G streptococci.

79. The method of claim 72, wherein the β -hemolytic streptococci is *Streptococcus*
pyogenes.

80. The method of claim 72, wherein the mammal is a human, dog, bovine, porcine,
or horse.

20 81. The method of claim 80, wherein the mammal is human.

82. A method of protecting a susceptible mammal against β -hemolytic streptococcal
colonization or infection comprising administering to the mammal an effective amount of an
immunogenic composition comprising a polynucleotide of claim 22, which amount is
effective to prevent or ameliorate β -hemolytic streptococcal colonization or infection in the
25 susceptible mammal.

83. The method of claim 82, wherein said immunogenic composition comprises a
recombinant expression vector comprising the polynucleotide of claim 22.

WO 02/083859

PCT/US02/11610

84. The method of claim 82, wherein the immunogenic composition further comprises a physiologically-acceptable vehicle.

85. The method of claim 82, wherein the immunogenic composition is administered by subcutaneous or intramuscular injection.

5 86. The method of claim 82, wherein the immunogenic composition is administered by oral ingestion.

87. The method of claim 82, wherein the immunogenic composition is administered intranasally.

88. The method of claim 82, wherein the β -hemolytic streptococci is group A streptococci, group B streptococci, group C streptococci, or group G streptococci.

89. The method of claim 82, wherein the β -hemolytic streptococci is *Streptococcus pyogenes*.

90. The method of claim 82, wherein the mammal is a human, dog, bovine, porcine, or horse.

15 91. The method of claim 90, wherein the mammal is human.

92. A composition for reducing at least one of the number and the growth of β -hemolytic streptococci in a mammal having a β -hemolytic streptococcal colonization or infection, comprising an antibody that immunospecifically binds with the polypeptide of claim 1.

20 93. A composition for reducing at least one of the number and the growth of β -hemolytic streptococci in a mammal having a β -hemolytic streptococcal colonization or infection, comprising an antisense oligonucleotide capable of blocking expression of a nucleotide sequence encoding a polypeptide of claim 1.

94. The composition of claim 93, wherein the polypeptide comprises the mature polypeptide of an amino acid sequence of any of SEQ ID NOS: 2-668.

25

WO 02/083859

PCT/US02/11610

95. A composition for reducing at least one of the number and the growth of β -hemolytic streptococci, comprising an antisense oligonucleotide capable of blocking expression of a nucleotide sequence encoding a polypeptide of claim 1.

96. The composition of claim 95, wherein the polypeptide comprises the mature
5 polypeptide of an amino acid sequence of any of SEQ ID NOS: 2-668.

97. A method for reducing at least one of the number and the growth of β -hemolytic streptococci in a mammal having a β -hemolytic streptococcal colonization or infection, comprising administering to the mammal an effective amount of a composition comprising an antibody that immunospecifically bind to the polypeptide of claim 1, which amount is
10 effective to reduce the at least one of the number and the growth of β -hemolytic streptococci in the mammal.

98. A method for reducing at least one of the number and the growth of β -hemolytic streptococci in a mammal having a β -hemolytic streptococcal colonization or infection, comprising administering to the mammal an effective amount of a composition comprising an
15 antisense oligonucleotide capable of blocking expression of a nucleotide sequence encoding a polypeptide of claim 1.

99. A method for reducing side effects caused by β -hemolytic streptococcal infection in a mammal, comprising administering to the mammal an effective amount of a composition comprising an antibody that immunospecifically binds to the polypeptide of claim 1, which
20 amount is effective to reduce at least one of the number of and the growth of β -hemolytic streptococci in the mammal.

100. A method for reducing side effects caused by β -hemolytic streptococcal infection in a mammal, comprising administering to the mammal an effective amount of a composition comprising an antisense oligonucleotide capable of blocking expression of a
25 nucleotide sequence encoding a polypeptide of claim 1, which amount is effective to reduce at least one of the number of and the growth of β -hemolytic streptococci in the mammal.

101. A method for detecting and/or identifying β -hemolytic streptococci in a biological sample, comprising:

WO 02/083859

PCT/US02/11610

(a) contacting the biological sample with a polynucleotide of claim 22 under conditions that permit hybridization of complementary base pairs; and

(b) detecting the presence of hybridization complexes in the sample, wherein the detection of hybridization complexes indicates the presence of β -hemolytic streptococci in the biological sample.

102. The method of claim 101, wherein said β -hemolytic streptococci is *Streptococcus pyogenes*.

103. A method for detecting and/or identifying β -hemolytic streptococci in a biological sample, comprising:

(a) contacting the biological sample with an antibody that immunospecifically binds to the polypeptide of claim 1 under conditions suitable for the formation of immune complexes; and

(b) detecting the presence of immune complexes in the sample, wherein the detection of immune complexes indicates the presence of β -hemolytic streptococci in the biological sample.

104. The method of claim 103, wherein said β -hemolytic streptococci is *Streptococcus pyogenes*.

105. A method for detecting and/or identifying antibodies to β -hemolytic streptococci in a biological sample, comprising:

(a) contacting the biological sample with the polypeptide of claim 1 under conditions suitable for the formation of immune complexes; and

(b) detecting the presence of immune complexes in the sample, wherein the detection of immune complexes indicates the presence of antibodies to β -hemolytic streptococci in the biological sample.

106. The method of claim 105, wherein the polypeptide comprises a mature polypeptide of an amino acid sequence of any of even SEQ ID NOS: 2-670.

107. The method of claim 105, wherein said β -hemolytic streptococci is *Streptococcus pyogenes*.

WO 02/083859

PCT/US02/11610

108. An immunogenic composition comprising a polypeptide of claim 1.

109. The immunogenic composition of claim 108, said polypeptide being a mature polypeptide of an amino acid sequence of even numbered SEQ ID NOS: 2-668.

110. An immunogenic composition comprising a polynucleotide of claim 22.

5 111. The immunogenic composition of claim 110, comprising the expression vector of claim 30.

112. An immunogenic composition comprising an antibody that immunospecifically binds to the isolated polypeptide of claim 1.

10 113. The immunogenic composition of claim 112, wherein the isolated polypeptide comprises a mature polypeptide of an amino acid sequence of any of even numbered SEQ ID NOS: 2-668.

114. An isolated polynucleotide comprising a nucleotide sequence that has at least 70% identity to a nucleotide sequence that encodes an amino acid of SEQ ID NOS: 2-668, said polynucleotide being identified by the steps comprising:

15 (a) obtaining a first and second PCR primer derived from a nucleotide that encodes a mature polypeptide of SEQ ID NOS: 2-668, wherein the first and second primers are capable of initiating nucleic acid synthesis in an outward manner under PCR conditions, and wherein the first primer is capable of being extended in an antisense direction and the second primer is capable of being extended in a sense direction; and

20 (b) combining said first and second PCR primer with a cDNA library that contains said polynucleotide under PCR conditions suitable for synthesizing said nucleotide sequence from the first and second primers.

115. A method for extending a polynucleotide of claim 22 using polymerase chain reaction (PCR), comprising the steps of:

25 (a) obtaining a first and second PCR primer derived from said polynucleotide, wherein the first and second PCR primers are capable of initiating nucleic acid synthesis in an outward manner under PCR conditions, and wherein the first PCR primer is capable of being extended in an antisense direction and the second PCR primer is capable of being extended in a sense direction; and

WO 02/083859

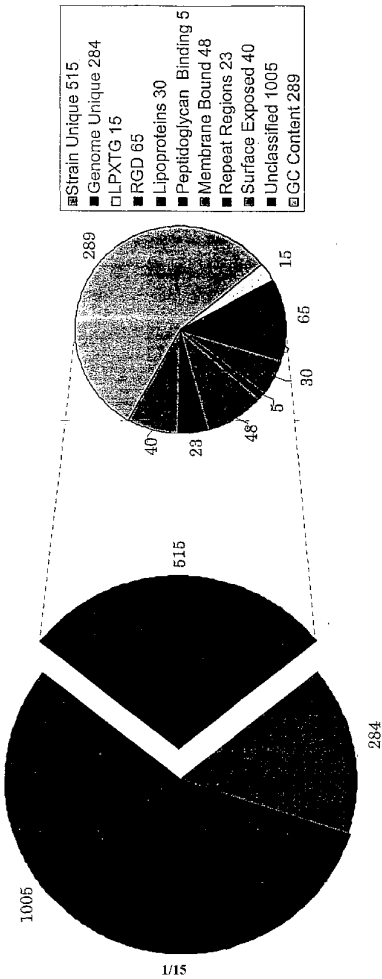
PCT/US02/11610

(b) combining said first and second PCR primers with said polynucleotide contained in a cDNA library under PCR conditions suitable for synthesizing nucleotide sequences from the first and second PCR primers, thereby extending said polynucleotide.

WO 02/083859

PCT/US02/11610

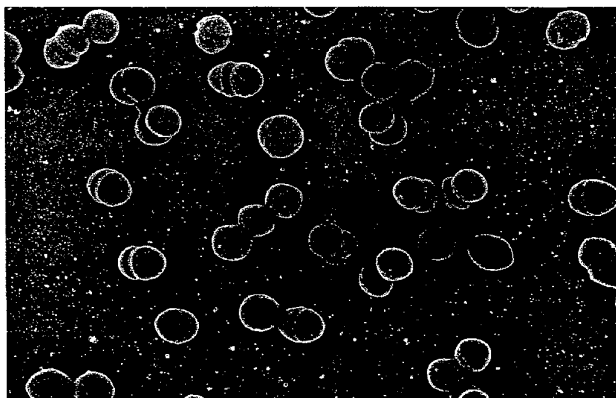
Fig. 1.



WO 02/083859

PCT/US02/11610

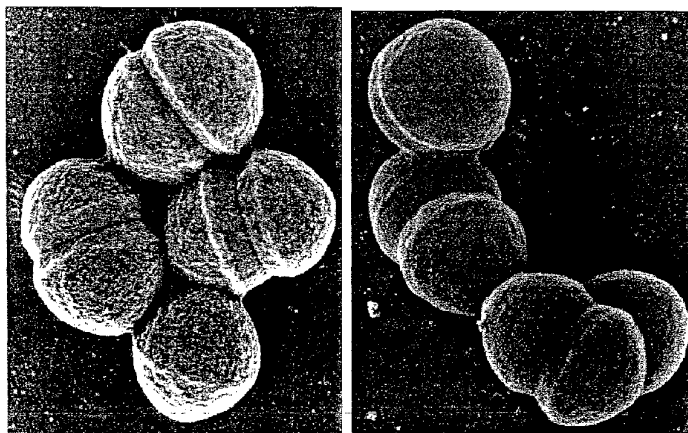
Fig. 2.



WO 02/083859

PCT/US02/11610

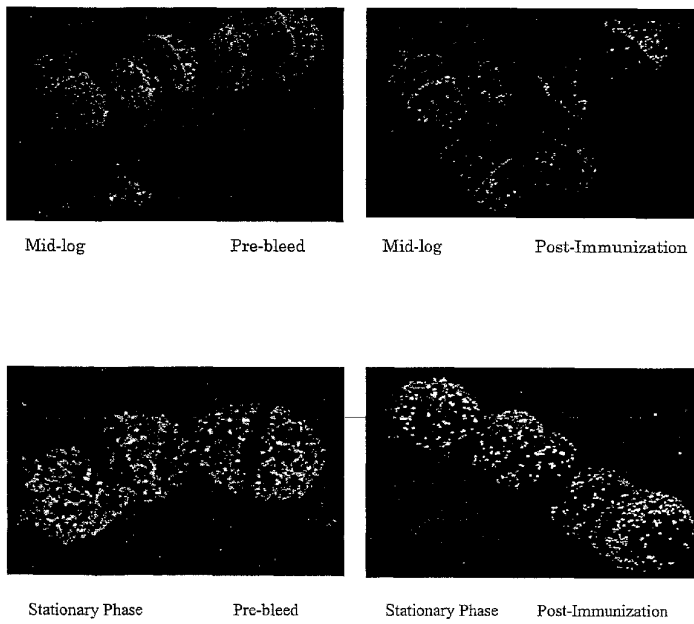
Fig. 3



WO 02/083859

PCT/US02/11610

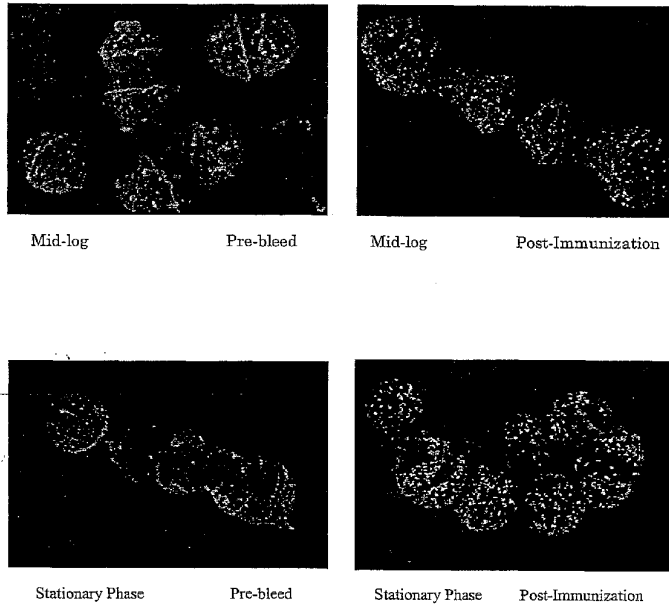
Fig. 4



WO 02/083859

PCT/US02/11610

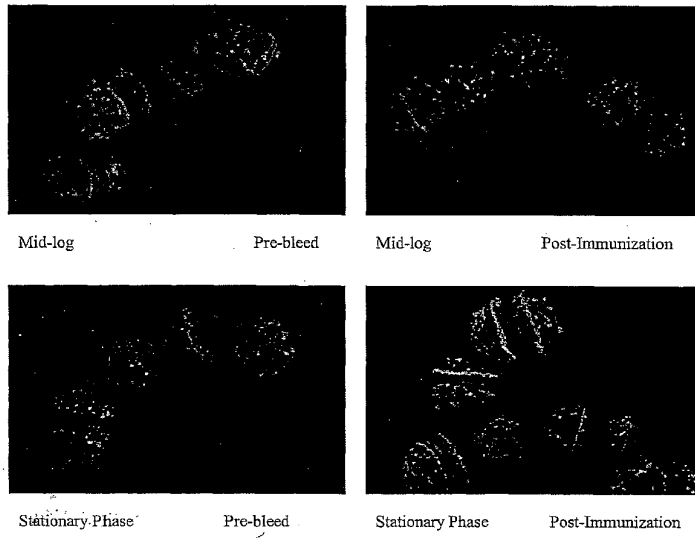
Fig. 5



WO 02/083859

PCT/US02/11610

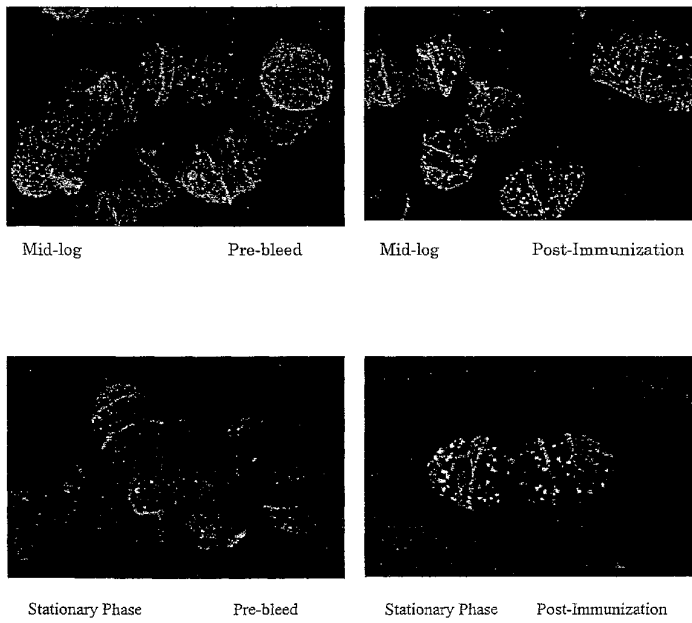
Fig. 6



WO 02/083859

PCT/US02/11610

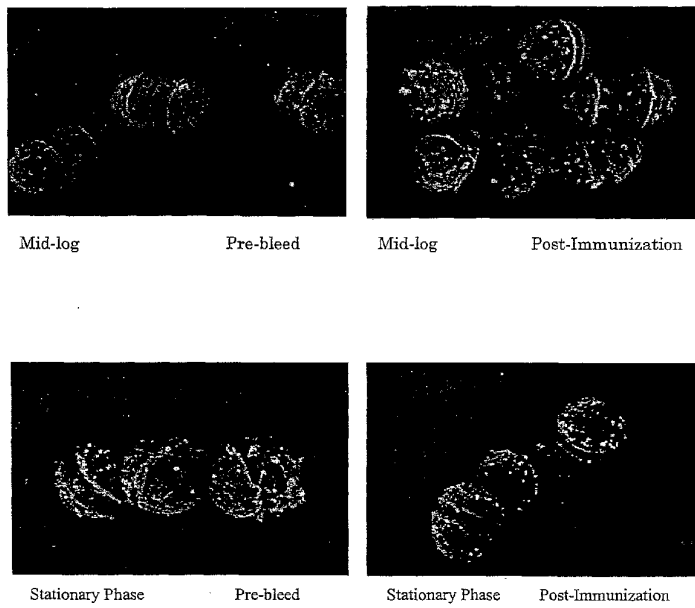
Fig. 7



WO 02/083859

PCT/US02/11610

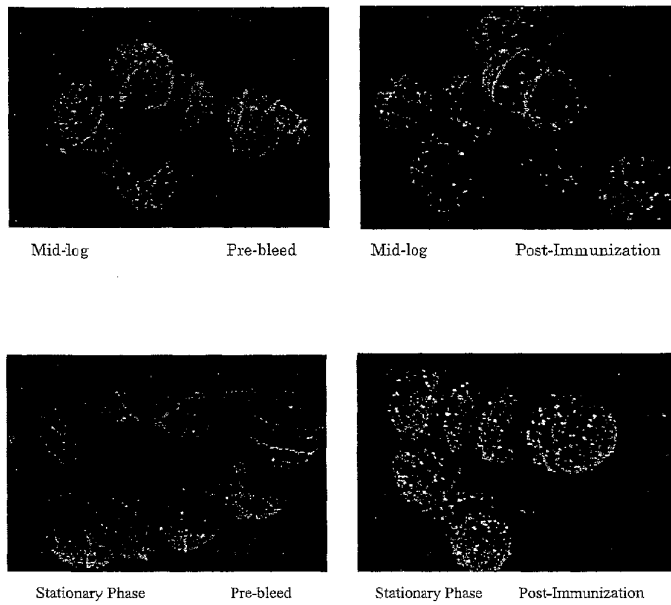
Fig. 8



WO 02/083859

PCT/US02/11610

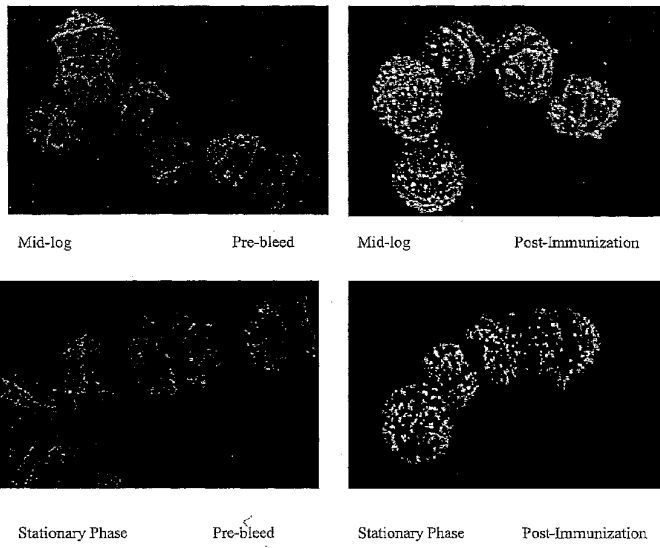
Fig. 9



WO 02/083859

PCT/US02/11610

Fig. 10



WO 02/083859

PCT/US02/11610

FIG. 11

ORF	CS204	06-0041	M1, 5yab+	CS204	60-226	BU-003	CS210	CS104	68-112	CS101	CS205	CB11	CS241	CS242	CB13	CS25	III
145	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
584	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-
665	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
721	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
850	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-
867	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1215	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-
1284	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
1310	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
1368	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
1487	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1639	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
1695	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
1728	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
1818	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2015	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2019	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2079	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2400	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2452	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2459	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2477	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2503	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

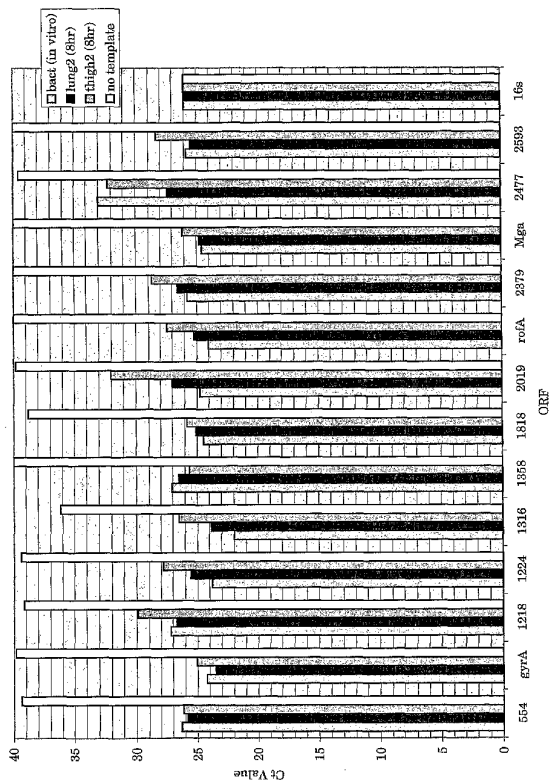
+ = positive reaction, expected size

- = negative reaction

WO 02/083859

PCT/US02/11610

FIG. 12

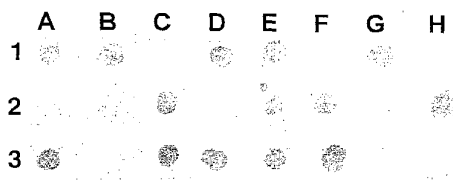


Quantitative PCR analysis of selected *S. pyogenes* ORFs. Total RNA and cDNA was prepared as described. All Ct values have been normalized to the 16s ribosomal RNA.

WO 02/083859

PCT/US02/11610

FIG. 13



WO 02/083859

PCT/US02/11610

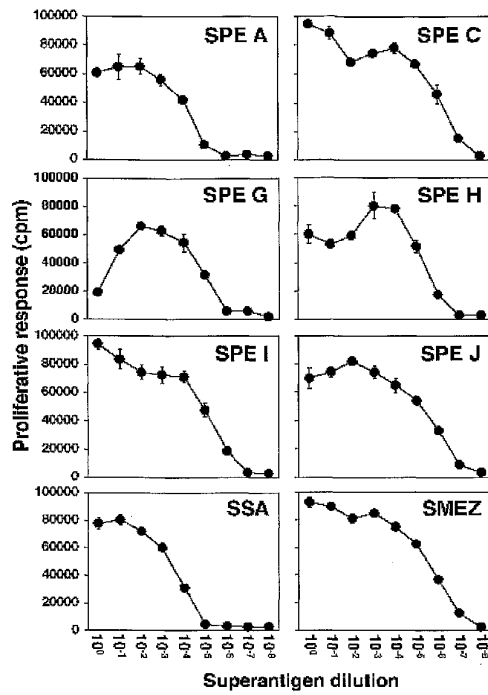


FIG. 14

WO 02/083859

PCT/US02/11610

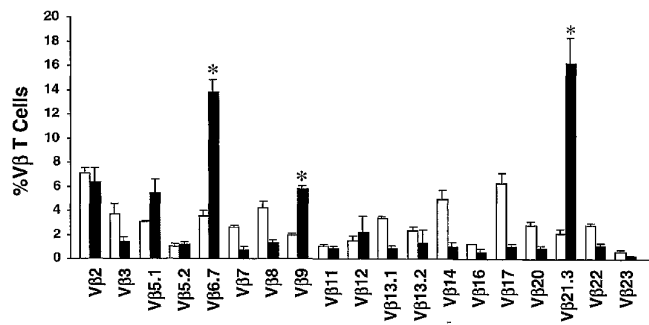
V β profile for SPE I

FIG. 15

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

REVISED VERSION

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
24 October 2002 (24.10.2002)

PCT

(10) International Publication Number
WO 2002/083859 A2

- (51) International Patent Classification⁷: C07H 21/04, A61K 39/02, C12Q 1/68
- (21) International Application Number: PCT/US2002/011610
- (22) International Filing Date: 12 April 2002 (12.04.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/283,358 13 April 2001 (13.04.2001) US
- (71) Applicant (for all designated States except US): WYETH [US/US], Five Girardeau Farms, Madison, NJ 07940 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): OLMSTEAD, Stephen, Bruce [US/US], 395 Kreag Road, Pittsford, NY 14534 (US); ZAGLIRSKY, Robert, John [US/US], 569 Fox Hunt Drive, Victor, NY 14564 (US); NICKBARG, Elliott, Bruce [US/US], 32 Audrey Road, Belmont, MA 02140 (US); WINTER, Laurie, Anne [US/US], 2915 Colby Street, Brockport, NY 14420 (US).
- (74) Agents: SCHULMAN, Robert, M. et al., Intellectual Property Department, Hunton & Williams, 1900 K Street, N.W., Suite 1200, Washington, DC 20006-1109 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with declaration under Article 17(2)(a); without abstract;
title not checked by the International Searching Authority
- (48) Date of publication of this revised version: 25 March 2004
- (15) Information about Correction:
see PCT Gazette No. 13/2004 of 25 March 2004, Section II
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SURFACE PROTEINS OF *STREPTOCOCCUS PYOGENES*

(57) Abstract:

WO 2002/083859 A2

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rule 13ter.1(c) and 39)

Applicant's or agent's file reference 38523.000017	IMPORTANT DECLARATION		Date of mailing (day/month/year)
International application No. PCT/US02/11610	International filing date (day/month/year) 12 April 2002 (12.04.2002)	(Earliest) Priority date (day/month/year) 13 April 2001 (13.04.2001)	
International Patent Classification (IPC) or both national classification and IPC IPC(7): C07H 21/04; A61K 39/02; C12Q 1/68; and US Cl.: 536/23.7, 23.1; 435/196, 195, 253.4; 530/300, 324, 832; 424/94.6, 94.61, 190.1; 514/2.8, 44, 54.			
Applicant WVEIR			

This International Searching Authority hereby declares, according to Article 17(2)(a), that no international search report will be established on the international application for the reasons indicated below.

1. ☐ The subject matter of the international application relates to:
 - a. ☐ scientific theories.
 - b. ☐ mathematical theories
 - c. ☐ plant varieties.
 - d. ☐ animal varieties.
 - e. ☐ essential biological processes for the production of plants and animals, other than microbiological processes and the products of such processes.
 - f. ☐ schemes, rules or methods of doing business.
 - g. ☐ schemes, rules or methods of performing purely mental acts.
 - h. ☐ schemes, rules or methods of playing games.
 - i. ☐ methods for treatment of the human body by surgery or therapy.
 - j. ☐ methods for treatment of the animal body by surgery or therapy.
 - k. ☐ diagnostic methods practised on the human or animal body.
 - l. ☐ mere presentations of information.
 - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art.
2. ☐ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:

☐ the description
 ☐ the claims
 ☐ the drawings
3. ☒ The failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions prevents a meaningful search from being carried out:

☒ the written form has not been furnished or does not comply with the standard.
 ☒ the computer readable form has not been furnished or does not comply with the standard.
4. Further comments:

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 505-3230	Authorized officer <i>Felicia D. Roberts</i> Patent Officer v. Bastar Telephone No. (703) 208-0196
---	---

Form PCT/ISA/203 (July 1998)

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
A 6 1 P 31/04	A 6 1 K 48/00	4 C 0 8 5
C 0 7 K 14/315	A 6 1 P 31/04	4 H 0 4 5
C 0 7 K 16/12	C 0 7 K 14/315	
C 1 2 N 1/15	C 0 7 K 16/12	
C 1 2 N 1/19	C 1 2 N 1/15	
C 1 2 N 1/21	C 1 2 N 1/19	
C 1 2 N 5/10	C 1 2 N 1/21	
C 1 2 P 21/02	C 1 2 P 21/02	C
C 1 2 Q 1/68	C 1 2 Q 1/68	A
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/53	N
G 0 1 N 33/569	G 0 1 N 33/566	
/(C 1 2 Q 1/68	G 0 1 N 33/569	C
C 1 2 R 1:46)	C 1 2 N 5/00	A
	C 1 2 Q 1/68	A
	C 1 2 R 1:46	

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

(74)代理人 100107386

弁理士 泉谷 玲子

(72)発明者 オルムステッド, スティーヴン・ブルース

アメリカ合衆国ニューヨーク州1 4 5 3 4, ピッツフォード, クリーグ・ロード 3 9 5

(72)発明者 ザガースキー, ロバート・ジョン

アメリカ合衆国ニューヨーク州1 4 5 6 4, ヴィクター, フォックス・ハント・ドライブ 5 6 9

(72)発明者 ニックバーグ, エリオット・ブルース

アメリカ合衆国マサチューセッツ州0 2 1 4 0, ベルモント, オードリー・ロード 3 2

(72)発明者 ウィンター, ローリー・アン

アメリカ合衆国ニューヨーク州1 4 4 2 0, ブロックポート, コルビー・ストリート 2 9 1 5

F ターム(参考) 4B024 AA01 AA13 BA31 CA04 CA20 GA11 HA12 HA17

4B063 QA01 QA19 QQ43 QR32 QR55 QR62 QS25 QS34

4B064 AG31 CA01 CA19 CC24 DA01

4B065 AA49Y AB01 AC14 BA02 CA24 CA45 CA46

4C084 AA02 AA07 AA13 DC50 MA17 MA52 MA59 MA66 NA14 ZB351

4C085 AA13 BB11 CC13 CC21 DD86

4H045 AA10 AA11 AA20 AA30 CA11 DA75 DA86 EA31 EA52 FA74

专利名称(译)	化脓性链球菌 (化脓性链球菌) の表面蛋白		
公开(公告)号	JP2004533236A	公开(公告)日	2004-11-04
申请号	JP2002582198	申请日	2002-04-12
[标]申请(专利权)人(译)	惠氏公司		
申请(专利权)人(译)	惠氏		
[标]发明人	オルムステッドスティーヴンブルース ザガースキーロバートジョン ニックバーグエリオットブルース ウィンターローリーアン		
发明人	オルムステッド,スティーヴン・ブルース ザガースキー,ロバート・ジョン ニックバーグ,エリオット・ブルース ウィンター,ローリー・アン		
IPC分类号	G01N33/53 A61K39/00 A61K39/39 A61K39/395 A61K48/00 A61P31/04 C07K14/135 C07K14/315 C07K16/12 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12P21/02 C12Q1/68 C12R1/46 G01N33/566 G01N33/569		
CPC分类号	A61P31/04 C07K14/315		
FI分类号	C12N15/00.ZNA.A A61K39/00.H A61K39/39 A61K39/395.D A61K39/395.N A61K48/00 A61P31/04 C07K14/315 C07K16/12 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/68.A G01N33/53.M G01N33/53.N G01N33/566 G01N33/569.C C12N5/00.A C12R1/46		
F-TERM分类号	4B024/AA01 4B024/AA13 4B024/BA31 4B024/CA04 4B024/CA20 4B024/GA11 4B024/HA12 4B024 /HA17 4B063/QA01 4B063/QA19 4B063/QQ43 4B063/QR32 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B064/AG31 4B064/CA01 4B064/CA19 4B064/CC24 4B064/DA01 4B065/AA49Y 4B065 /AB01 4B065/AC14 4B065/BA02 4B065/CA24 4B065/CA45 4B065/CA46 4C084/AA02 4C084/AA07 4C084/AA13 4C084/DC50 4C084/MA17 4C084/MA52 4C084/MA59 4C084/MA66 4C084/NA14 4C084 /ZB351 4C085/AA13 4C085/BB11 4C085/CC13 4C085/CC21 4C085/DD86 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/CA11 4H045/DA75 4H045/DA86 4H045/EA31 4H045/EA52 4H045 /FA74		
代理人(译)	小林 泰 千叶昭夫		
优先权	60/283358 2001-04-13 US		
外部链接	Espacenet		

摘要(译)

描述了β-溶血性链球菌多核苷酸，多肽，特别是化脓性链球菌多肽和多核苷酸，以及这些多肽的抗体。可以配制本发明的多核苷酸，多肽和抗体以用作免疫原性组合物。还公开了针对和减少β-溶血性链球菌感染进行免疫和检测生物样品中的β-溶血性链球菌的方法。

【 0 0 4 5 】

【 表 5 】

68	705	1202	1664	1723	2020	2385
165	729	1310	1667	1777	2046	2414
252	788	1358	1678	1909	2170	2437
510	1058	1362	1680	1972	2236	2601
601	1132	1573	1681	1975	2250	
608	1200	1638	1683	2014	2300	

さらに、ソフトウェアプログラム T o p P r e d 2 (1 0) を用いて、タンパク質の膜貫通 (T M) ドメインを解析した。このプログラムは、潜在的に膜の脂質二重層に渡る可能性のある、疎水性のタンパク質領域を予測する。T o p P r e d 2 による、潜在的に膜の脂質二重層に渡る可能性がある、タンパク質疎水性領域に関する解析によって、3 以上の膜貫通ドメインを持つ、そしてしたがって膜結合であると考えられる、4 8 の O R F (表 V I I) が同定された。

【 0 0 4 6 】

表 V I I I 3 以上の膜貫通領域を持つ、推定上のタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 4 7 】

【 表 6 】

8	307	594	752	1222	1398	2069
73	312	613	844	1266	1657	2069
80	395	650	925	1317	1708	2237
93	508	672	975	1488	1726	2283
141	551	706	1018	1496	1779	2424
265	567	708	1152	1513	1999	2562
306	593	731	1156	1596	2002	

タンパク質ドメインまたは保存されるタンパク質領域の多数並列の隠れマルコフモデル (H M M) P r a m ー ス (6 1) を用いて、保存するタンパク質の異なるタンパク質ファミリーに属する可能性がある化膿性連鎖球菌タンパク質を同定した。このアウトプットのキーワード検索を用いて、B l a s 検索標準では見過した可能性があるタンパク質を同定した。本発明者の譲受人もまた、H M M モデルを開発した。コンピュータアルゴリズム、H M M L i p o を開発して、1 5 以上の生物学的に性質決定された、3 0 を超える生物由来の非化膿性連鎖球菌細胞リポタンパク質を用いて、リポタンパク質を予測した。このトレーニングセットは、実験的に立証された膜核生物リポタンパク質から生成された。H M M L i p o は、推定上のリポタンパク質である、3 0 の O R F を同定した (表 V I I) 。

【 0 0 4 8 】

表 V I I I 推定上のリポタンパク質をコードするオープンリーディングフレーム (O R F)

【 0 0 4 9 】

【 表 7 】

68	601	747	1659	1789	1983
309	678	1137	1664	1818	2417
347	685	1202	1723	1878	2452
340	704	1264	1753	1882	2459
354	729	1495	1788	1918	2601

さらに、1 5 の O R F が L P X T G モチーフを有すると予測され、そしてソルターゼに属的と考えられる可能性があるタンパク質として分類された (表 V I I I) 。

【 0 0 5 0 】