



(12) 发明专利

(10) 授权公告号 CN 107106635 B

(45) 授权公告日 2021.10.08

(21) 申请号 201580071394.2

(22) 申请日 2015.11.20

(65) 同一申请的已公布的文献号
申请公布号 CN 107106635 A

(43) 申请公布日 2017.08.29

(30) 优先权数据
62/082,848 2014.11.21 US

(85) PCT国际申请进入国家阶段日
2017.06.28

(86) PCT国际申请的申请数据
PCT/US2015/061859 2015.11.20

(87) PCT国际申请的公布数据
W02016/081839 EN 2016.05.26

(73) 专利权人 俄克拉何马大学董事会
地址 美国俄克拉何马州

(72) 发明人 R·K·特温坦

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

代理人 韩威威

(51) Int. Cl.
A61K 38/00 (2006.01)

(56) 对比文件

US 2014341942 A1,2014.11.20

CN 102939105 A,2013.02.20

CN 101018863 A,2007.08.15

US 2014023674 A1,2014.01.23

US 8128939 B2,2012.03.06

J A Walker等.Molecular Cloning, Characterization, and Complete Nucleotide Sequence of the Gene for Pneumolysin, the Sulfhydryl-Activated Toxin of Streptococcus Pneumoniae.《Infection and immunity》.1987,第55卷(第5期),1184-1189.
Richard M Harvey等.The Impact of Pneumolysin on the Macrophage Response to Streptococcus Pneumoniae Is Strain-Dependent.《PloS one》.2014,第9卷(第8期),1-9.

侯宏嘉.肺炎链球菌溶血素Ply蛋白疫苗及BLP技术应用的研究.《中国优秀硕士学位论文全文数据库 医药卫生科技辑》.2014,(第9期),E059-72.

审查员 邓沁

权利要求书1页 说明书32页
序列表49页 附图10页

(54) 发明名称

肺炎球菌溶血素突变体及其使用方法

(57) 摘要

本公开在多种实施方案中提供胆固醇依赖性细胞溶素(如肺炎球菌溶血素)的免疫原性突变体,与其野生型蛋白相比,所述突变体具有减少的溶血活性和减少的成孔活性。在多种实施方案中,本公开还提供编码这类突变体的核酸和使用这类突变体的方法。

Table with 2 columns: Protein Name and Sequence. Lists various pneumolysin variants and their amino acid sequences.

CN 107106635 B

1. 一种纯化的突变肺炎球菌溶血素,其由SEQ ID NO:40所示的氨基酸序列组成。
2. 免疫原性组合物,其包含根据权利要求1所述的纯化的突变肺炎球菌溶血素和可药用赋形剂。
3. 疫苗,其包含根据权利要求2所述的免疫原性组合物。
4. 根据权利要求3所述的疫苗,其还包含佐剂。
5. 核酸序列,其编码根据权利要求1所述的纯化的突变肺炎球菌溶血素。
6. 宿主细胞,其包含根据权利要求5所述的核酸序列。
7. 根据权利要求1所述的纯化的突变肺炎球菌溶血素在制备用于预防肺炎链球菌(*Streptococcus pneumoniae*)所致的疾病的药物中的用途。
8. 免疫原性组合物在制备用于预防肺炎链球菌所致的疾病的药物中的用途,其中所述免疫原性组合物包含根据权利要求1所述的纯化的突变肺炎球菌溶血素。
9. 制备根据权利要求2所述的免疫原性组合物的方法,所述方法包括将所述纯化的突变肺炎球菌溶血素与所述可药用赋形剂组合。
10. 制备根据权利要求4所述的疫苗的方法,所述方法包括将所述纯化的突变肺炎球菌溶血素与所述可药用赋形剂组合得到免疫原性组合物,并将所述免疫原性组合物与所述佐剂混合。
11. 制备根据权利要求1所述的纯化的突变肺炎球菌溶血素的方法,所述方法包括从包含权利要求5所述的核酸序列的表达载体或权利要求6的宿主细胞表达所述肺炎球菌溶血素,并且分离所表达的肺炎球菌溶血素。

肺炎球菌溶血素突变体及其使用方法

[0001] 相关申请的交叉引用/

[0002] 通过引用声明并入

[0003] 本申请根据35USC§119(e)要求2014年11月21日提交的美国系列号62/082,848的权益。上述申请的完整内容通过引用方式明确地并入本文。

[0004] 关于联邦赞助研究或开发的声明

[0005] 本发明受美国政府支持根据美国国家健康研究所(NIH)资助的合同号AI037657做出。美国政府在本发明中享有某些权利。

背景技术

[0006] 胆固醇依赖性细胞溶素(CDC)是一个庞大的成孔毒素家族,由来自属梭菌属(*Clostridium*)、链球菌属(*Streptococcus*)、李斯特氏菌属(*Listeria*)、芽孢杆菌属(*Bacillus*)和隐秘杆菌属(*Arcanobacterium*)的超过20个种产生。这些毒素的成孔机制显示两个标志性特征:绝对依赖于膜胆固醇的存在和形成十分大的孔。每种CDC作为II型分泌系统分泌的可溶性单体蛋白质产生,一个成员例外。一旦遇到真核细胞,CDC发生可溶性单体蛋白质至膜嵌入型超分子孔复合体的转化。单体转化成低聚体膜插入型孔复合体需要单体结构的某些特别变化。

[0007] 尽管CDC作为 β -溶血蛋白熟知,已经变得日益明显的是,较之用作单纯的溶血素或通用细胞溶解物质,细菌病原体以复杂得多的方式利用这些蛋白质。CDC结构还显示可塑性,所述可塑性已经允许某些CDC的独特特征进化,而不破坏根本的成孔机制。在激活补体、利用非甾体受体、显示pH敏感性成孔机制或可以作为蛋白质转运通道发挥作用的CDC中反映这些特征中的某些。

[0008] CDC是富含 β 片层的四结构域蛋白质。结构域4中存在一个高度保守的富含色氨酸的十一肽,其参与某些CDC与富含胆固醇的膜结合。此外,已经显示在结构域4顶端靠近所述十一肽的三个其他疏水性短环(环L1、L2和L3)还插入膜表面并以垂直取向令CDC锚定至膜。在膜结合后,CDC单体横向地扩散以启动膜低聚物的形成。

[0009] 一旦前孔复合体达到巨大的尺寸,假定地达到完整环结构,则它随后转变成孔复合体。当前孔复合体内部每种单体的结构域3中的两个 α -螺旋束转化成两个延长的两亲性跨膜 β -发夹(TM β H)时,形成跨膜孔。一旦前孔转化成孔,前孔结构的高度发生约40埃的垂直塌陷。前孔结构的塌陷导致结构域3TM β H处于膜表面的惊人距离范围内,此时它们协调地插入膜中,导致巨大的跨膜 β -桶孔形成。CDC孔巨大:它由35至50个单体组成并且显示250至300埃的直径。

[0010] 在CDC单体与膜相互作用过程期间,一旦CDC单体与膜表面相互作用,则结构域4 β -夹心的顶端的十一肽和三个其他短环(L1、L2、和L3)插入膜中。这些环并不深入穿透膜中并且表面上不直接参与跨膜孔的结构。这些环的一个功能似乎是令单体以直立位置锚定至膜中。结构域4以垂直于膜的取向存在并且由含水环境包围,甚至在低聚状态下也如此。

[0011] CDC的结构域4介导膜识别,无论它借助胆固醇或另一个受体发挥作用,如在ILY

(中间链球菌溶素)的情形。

[0012] CDC还能够在体外溶解多种类型的有核细胞类型,并且这种能力转而已经由多位研究者用来以CDC透化各种真核细胞类型。尽管这些毒素具有作为通用细胞溶解物质在体外发挥作用的能力,尚未证实细胞溶解是感染期间CDC的主要功能。已经例如在单核增生李斯特菌(*Listeria monocytogenes*)、化脓性链球菌(*Streptococcus pyogenes*)、肺炎链球菌(*Streptococcus pneumoniae*)、化脓隐秘杆菌(*Arcanobacterium pyogenes*)和产气荚膜梭菌(*Clostridium perfringens*)中研究了CDC对感染的贡献。这些研究中某些研究的结果表明,较之用作通用溶细胞物质,细菌以更复杂的方式利用CDC。还显而易见,CDC结构已经发生某些独特的进化转变,所述转变促进这些细菌物种的致病机制。

[0013] 肺炎链球菌是人类中、尤其婴儿、老年人、患慢性疾病的人和免疫受损人员当中重要的病原体。它是频繁从侵袭性疾病如菌血症/败血症、肺炎和脑膜炎患者分离的细菌,全球发病率和死亡率高。甚至在适宜抗生素治疗的情况下,肺炎球菌感染仍导致许多死亡。尽管抗微生物药物的到来已经降低因肺炎球菌病所致的总体死亡率,但是耐药性肺炎球菌菌株的存在如今已经变成全球重大问题并且强调需要通过抗微生物药之外的方法治疗及防止肺炎球菌感染。有效的肺炎球菌疫苗可能对肺炎链球菌病相关的发病率和死亡率产生重大影响。这疫苗将还潜在地可用于婴儿和幼儿中防止中耳炎。迫切需要提供长期免疫力的新型免疫原性肺炎球菌疫苗,特别对于2岁以下儿童,因为这个年龄组中疾病发生率高并且针对多糖疫苗抗原的抗体应答不良。

[0014] 每年在美国,肺炎球菌病估计占据3,000例脑膜炎、50,000例菌血症、500,000例肺炎和7百万例中耳炎。

[0015] 重度肺炎球菌感染因细菌散布至血流和中枢神经系产生。在1997年,来自基于社区的研究的数据显示,在美国,肺炎球菌菌血症的年总体发病率估计为每100,000人15-30例;对于大于或等于65岁的人(每100,000人50-83例)和年龄小于或等于2岁的儿童(每100,000人160例),该比率更高。在成人中,60%-87%的肺炎球菌菌血症与肺炎相关;在幼儿中,原发性感染部位往往未确定。

[0016] 在美国,白种人当中罹患菌血症的风险比其他族裔/人种群体(即,黑人、阿拉斯加原住民和美国印第安人)当中更低。与白人相比,成年黑人具有高三倍至五倍的菌血症总体发病率(每100,000人49-58例)。侵袭性肺炎球菌病的比率在阿拉斯加原住民和美国印第安人当中很高。通过前瞻性监护研究确定阿拉斯加原住民和2岁以下阿拉斯加土著儿童当中,年龄修正的侵袭性肺炎球菌感染年度发病率分别是每100,000人74例和624例。全部年龄的阿拉斯加原住民当中脑膜炎和菌血症肺炎的比率比其他美国人群高八倍至十倍。已经在特定美国印第安群体(例如,Apache)当中报道了对任何美国人群而言最高的发病率。这类群体的年总体发病率是每100,000人156例;这些群体中1-2岁儿童的发病率是每100,000人2,396例。

[0017] 在美国,估计的肺炎球菌脑膜炎年总体发病率是每100,000人一至二例。肺炎球菌脑膜炎的发病率在年龄6-24个月的儿童和年龄大于或等于65岁的人当中最高。黑人的发病率两倍于白人和西班牙人。因为儿童中b型流感嗜血杆菌(Hib)脑膜炎的发病率在引入Hib缀合物疫苗后迅速降低,所以肺炎链球菌已经变成美国最常见的细菌性脑膜炎病因(26)。

[0018] 耐药性肺炎链球菌菌株(DRSP)已经在美国和世界其他部分变得日益常见。在一些

地区中,已经报道多达35%的肺炎球菌分离株具有中等水平(最低抑菌浓度{MIC}等于0.1-1.0 μ g/mL)或高水平(MIC大于或等于2 μ g/mL)的青霉素耐药性。许多抗青霉素型肺炎球菌还抵抗其他抗微生物药物(例如,红霉素、甲氧苄啶-磺胺甲基异恶唑和超广谱头孢菌素类抗生素)。高水平的青霉素耐药性和多药耐药经常令肺炎球菌感染的管理复杂化并且致使针对脑膜炎、肺炎和中耳炎疑似病例选择经验性抗微生物疗法日益困难。治疗遭不敏感生物感染的患者可能需要使用昂贵的替代性抗微生物药并且可能导致住院延长和医疗费用增加。并未清晰确定抗微生物药耐药性对死亡率的影响。新出现的抗微生物药耐药性进一步强调需要通过接种防止肺炎球菌感染。

[0019] 目前可用的肺炎球菌疫苗, **PNEUMOVAX® 23** (Merck&Co., Inc., Kenilworth, N.J.) 和 **PNU-IMMUNE® 23** (Lederle-Praxis Biologicals, Pearl River, NY), 包含23种纯化的肺炎链球菌(血清型1、2、3、4、5、6B、7F、8、9N、9V、10A、11A、12F、14、15B、17F、18C、19A、19F、20、22F、23F和33F)荚膜多糖抗原。这些疫苗在1983年于美国注册并且替代1977年注册的较早14价制剂。一剂(0.5mL) 23价疫苗含有25 μ g溶解于等渗盐水溶液中的每种荚膜多糖抗原,具有作为防腐剂添加的苯酚(0.25%)或硫柳汞(0.01%)并且无佐剂。截至1997年,该疫苗中的23种荚膜类型代表至少85%-90%在美国儿童和成人当中造成侵袭性肺炎球菌感染的血清型。在这个23价疫苗中代表了截至1997在美国最频繁造成侵袭性耐药性肺炎球菌感染的六个血清型(6B、9V、14、19A、19F和23F)。如下文所示,仅由荚膜多糖组成的疫苗的满意度有限。

[0020] 肺炎球菌溶血素尤其是链球菌肺炎(其每年全球杀死超过一百万人)发病机制中的关键组分。肺炎球菌溶血素作为肺炎链球菌肺部感染和中耳炎的疫苗组成部分的用途可能提供重要益处,原因是基于荚膜多糖的疫苗正在因遗传性变异丧失有效性并且因为存在多于90种不同的肺炎链球菌荚膜血清型而难以生产。针对一个荚膜类型的免疫力不针对另一个荚膜类型提供保护。上文讨论的目前可用的肺炎球菌疫苗,包含来自最频繁造成疾病的菌株的23种荚膜多糖,具有主要与某些荚膜多糖的免疫原性不良、血清型多样和血清型分布随时间、地理区域和年龄组的差异相关的明显缺点。目前,肺炎球菌溶血素的点突变变体已经用于疫苗开发。这种肺炎球菌溶血素突变体(称作“Pd-B”)在位置433处含有单突变(其中已经将天然色氨酸残基改变成苯丙氨酸)。肺炎球菌溶血素中的这种突变是在结构域4的保守性十一肽中,其长期以来被认为是介导与哺乳动物膜结合的胆固醇依赖性细胞溶素(CDC)中的结构。

[0021] 尽管肺炎球菌溶血素Pd-B突变体常规地用于疫苗开发,但是这种蛋白质仍能够发生与哺乳动物细胞的膜结合后出现的多种结构性转变。这些变化大幅度改变其结构并且可能降低其在患者中刺激有效的中和免疫应答的能力,主要原因是患者免疫系统可以“看到”的肺炎球菌溶血素的结构将是终末细胞结合型低聚复合体的结构,而非可溶性单聚肺炎球菌溶血素的初始结构。更重要地,现有以遗传方式类毒素化的肺炎球菌溶血素仍受到不可接受水平的毒性阻碍。这种毒性的基础尚不清晰,但是或许因以下事实所致:这种类毒素仍可以与哺乳动物细胞结合并在其上寡聚化。

[0022] 因此,胆固醇依赖性细胞溶素(如(但不限于)肺炎球菌溶血素)的突变体将具有巨大益处,所述突变体具有降低的毒性和降低的溶血活性,仍刺激针对相应疾病生物的免疫应答。

[0023] 附图简述

[0024] 本公开的几种实施方案由此在附图中说明。然而应当指出,附图仅说明了几个常见实施方案并且因此不意在视为限制本公开的范围。另外,在附图中,相似或相同的附图标记可以用来确定共同或相似的要素,并且并非全部这些要素可以如此编号。附图不必然是按比例的和附图的某些特征和某些视图可以在比例上夸张显示并且为清晰和简洁起见以示意图显示。

[0025] 图1A-E含有各种胆固醇依赖性细胞溶素的天然氨基酸序列的氨基酸对比比较。本文中确定的每种蛋白质的氨基酸序列对应于表1中的SEQ ID NO;例如,图1A-E中的腊状芽孢杆菌溶素(cereolysin)对应于表1中的SEQ ID NO:2,并且表1中的SEQ ID NO:18 (PAF)对应于图1A-E中的草绿色链球菌溶素。

[0026] 图2显示ILY(中间链球菌溶素)的晶体结构以及ILY和PF0(产气荚膜杆菌溶素)的D4晶体结构比较。(a)中显示了ILY²⁵晶体结构的带状图,所述带状图指出这些研究中提到的多种结构和残基的位置。(b)中基于ILY和PF0的晶体结构显示了两种蛋白质的D4结构的带状图叠加^{23,24}。显示了两种蛋白质的十一肽的相对位置以及ILY和PF0的L1-L3环残基(后者在括号中)。使用VMD生成结构图像²⁵。

[0027] 图3显示ILY十一肽插入胆固醇耗尽的膜中。ILY残基A1a-486突变成半胱氨酸(ILY^{A486C})并且用NBD(碘代乙酰氨基-N,N'-二甲基-N-(7-硝基-2-氧杂-1,3-二唑基)乙二胺)衍生化。当ILY^{A486C-NBD}单独温育(实线)、与人红细胞(hRBC-短划线)温育或与耗尽胆固醇的hRBC(点线)温育时,测定NBD的荧光发射。

[0028] 图4显示ILY的环L1、L2和L3不插入胆固醇耗尽的膜中。将已知插入膜中的每个D4环残基置换为半胱氨酸并用NBD修饰。将ILY^{A428C-NBD}(a)、ILY^{A464C-NBD}(b)或ILY^{L518C-NBD}(c)单独温育(实线)、与hRBC(短划线)温育或与耗尽胆固醇的hRBC(点线)温育。随后恢复膜胆固醇并测定环L1、L2和L3的插入。将ILY^{A428C-NBD}(d)、ILY^{A464C-NBD}(e)或ILY^{L518C-NBD}(f)单独温育(实线)或与胆固醇充足的膜温育(短划线)。

[0029] 图5显示L1-L3环介导PF0与富含胆固醇的脂质体结合,(a)天然PF0(实线)和NEM修饰的PF0(短划线)的结合作用的SPR分析,(b)天然PF0(实线)、PF0^{A401D}(长短划线)、PF0^{A437D}(短短划线)和PF0^{L491D}(点线)的结合作用的SPR分析。

[0030] 图6显示化学修饰PF0十一肽半胱氨酸硫氢基阻断十一肽色氨酸插入膜和前孔转化成孔。PF0十一肽色氨酸的内在荧光发射增加已经用来测量它们向膜中的插入^{20,21}。(a)当天然PF0中色氨酸从其可溶性形式(实线)变成其膜结合状态(短划线),显示其内在荧光发射增加。(b)用已经在Cys-459经NEM修饰的天然PF0重复(a)中所示的相同实验。

[0031] 图7显示用突变肺炎球菌溶血素多肽和野生型肺炎球菌溶血素免疫、随后用肺炎链球菌免疫接种的小鼠中的免疫原性应答。

[0032] 发明详述

[0033] 在通过示例性附图、实验、结果和实验室程序详细解释发明构思的至少一个实施方案之前,应当理解本公开不限于如以下描述中所述或附图、实验和/或结果中所示的组成、组分和方法的细节。本公开能够由其他实施方案或按多种方式实施或开展。因而,本文所用的语言意在按照最广的可能范围和含义给出,并且实施方案意在为示例性质,而非穷举性质。另外,应当理解本文中所述的措辞和术语意在描述并且不应当视为限制性。

[0034] 除非本文另外定义,否则本公开中使用的科学和技术术语应当具有本领域普通技术人员通常理解的含义。另外,除非上下文所要求,单数术语应当包括复数并且复数术语应当包括单数。通常,与本文所述的细胞和组织培养、分子生物学和蛋白质及寡核苷酸或多核苷酸化学及杂交相联系所使用的命名及其技术是本领域熟知和通常使用的那些。标准技术用于重组DNA、寡核苷酸合成和组织培养和转化(例如,电穿孔法、脂质体转染法)。酶促反应和纯化技术根据生产商的说明或如本领域中通常所完成或如本文所述那样进行。前述技术和方法通常根据本领域熟知和如本说明书通篇范围内引用及讨论的多种普通和更具参考文献中所述的常规方法进行。参见,例如,Green and Sambrook (Molecular Cloning: A Laboratory Manual (第4版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)) 和 Coligan 等人 (Current Protocols in Immunology, Current Protocols, Wiley Interscience (1994)), 所述文献通过引用的方式明确并入本文。与本文所述的分析化学、合成有机化学和医用化学和药物化学结合所用的命名以及实验室方法和技术是本领域熟知并常用的那些。标准技术用于化学合成、化学分析、药物制备、配制和递送及患者治疗。

[0035] 本说明书中提到的全部专利、公开的专利申请和非专利出版物均表示本公开所属领域的技术人员的技术水平。在本申请的任何部分中提及的全部专利、公开的专利申请和非专利出版物均明确通过引用方式完整并入本文至相同的程度,如同专门且个别地指出通过引用的方式并入每份单独的专利或出版物。特别地,以下专利和专利申请的完整内容因而通过引用的方式明确并入本文:2012年2月21日提交的美国系列号13/401,460;2008年4月14日提交的美国系列号12/102,696,2012年3月6日颁布的当前美国专利号8,128,939;2007年4月13日提交的美国系列号60/923,281;和2014年11月21日提交的美国系列号62/082,848。

[0036] 可以在不进行过多实验的情况下根据本公开制得并实施本文描述和/或否则构思的全部组合物和/或方法。虽然已经根据具体实施方案描述了本文中公开和/或否则构思的组合物和/或方法,但是本领域技术人员将显而易见,变型可以适用于本文描述和/或否则构思的组合物和/或方法中以及用于本文描述和/或否则构思的方法的各步骤或系列步骤中而不脱离本公开的构思、精神和范围。对本领域技术人员显而易见的全部这类相似的代用品和修改形式视为处于所附权利要求限定的发明构思的精神、范围和设想范围内。

[0037] 如根据本公开所用,除非另外说明,否则以下术语应当理解成具有以下含义:

[0038] 在权利要求书和/或本说明书中与术语“包含”联合使用时,词“一个(a)”或“一种(an)”的用途可以意指“一个(one)”,但是它还符合以下意思:“一个或多个”、“至少一个”和“一个或多于一个”。虽然本公开内容支持仅指替代物和“和/或”的定义,但是除非明确指示仅指替代物或替代物是互斥的,否则术语“或”在权利要求书中的用途用来意指“和/或”。在本申请通篇范围内,术语“约”用来表示某个值包括正在用来测定该值的装置、方法的内在误差变异或在研究受试者之间存在的变异。例如但不限于,当使用术语“约”时,指定的值可以变动正或负12%,或11%,或10%,或9%,或8%,或7%,6%,或5%,或4%,或3%,或2%,或1%。术语“至少一个”的运用将理解为包括一个以及多于一个的任何量,包括但不限于2、3、4、5、10、15、20、30、40、50、100个等。术语“至少一个”可以扩展直至100或1000个或更多,这取决于与之相联系的术语;此外,100/1000的量不视为限制性,因为更高的限值也可能产

生令人满意的结果。此外,术语“X、Y和Z至少之一”的运用将理解为包括单独的X、单独的Y和单独的Z以及X、Y和Z的任何组合。使用序数术语(即,“第一”、“第二”、“第三”、“第四”等)的目的仅区分两个或更多个项并且例如不意在暗示任何顺序或次序或一项超过另一项的重要性或任何增加次序。

[0039] 如本说明书和权利要求书中所用,词“包含”(以及包含的任何形式,如“包含了”及“包含着”)、“具有”(以及具有的任何形式,如“具有了”及“具备”)、“包括”(以及包括的任何形式,如“包括了”及“包括”)或“含有”(以及含有的任何形式,如“含有”及“含有”)是包含性的或开放式的并且不排除额外的、未提到的因素或方法步骤。

[0040] 如本文所用的术语“或其组合”指在该术语之前所列各项的全部排列和组合。例如,“A、B、C或其组合”意在包括以下至少一种情况:A、B、C、AB、AC、BC或ABC,并且如果次序在特定情境下重要,还包括BA、CA、CB、CBA、BCA、ACB、BAC或CAB。续上例,明确包括了含有一个或多个项或条目的重复的组合,如BB、AAA、AAB、BBC、AAABCCCC、CBBAAB、CABABB等。技术人员将理解,一般不限制任意组合中项或条目的数目,除非从语境中显而易见。

[0041] 本说明书通篇范围内和权利要求书,除非语境下需要,否则术语“基本上”和“约”将理解为不限于这些形容词/副词定性的具体条目,而是理解为某值包括正在用来测定该值的装置、方法的内在误差变异和/或在研究受试者之间存在的变异。因此,所述术语允许对其不产生显著影响的少量变异和/或偏差。例如,在某些情况下,术语“约”用来表示某个值包括正在用来测定该值的装置、方法的内在误差变异和/或在研究受试者之间存在的变异。类似地,术语“基本上”还可以指80%或更高,如85%或更高,或90%或更高,或95%或更高,或99%或更高等。

[0042] 如本文所用的术语“纯化的蛋白质”或“分离的蛋白质”意指蛋白质或片段基本上不含正常情况下随蛋白质一起出现的污染物或细胞组分,从而区分开蛋白质与污染物或细胞组分。未构思“纯化的”需要具有技术上完全纯(均匀)的制品,但如本文所用的“纯化的”意指蛋白质或多肽片段与正常情况下随蛋白质一起出现的污染物或细胞组分充分分离,旨在以该蛋白质可以用于测定法如免疫沉淀法或ELISA的状态下提供该蛋白质。例如,纯化的蛋白质可以位于电泳凝胶中。

[0043] 在本文中用来描述多肽时,术语“突变体”指与相应野生型(天然)多肽的氨基酸序列小于100%相同的多肽,并且尤其指其中已经置换野生型多肽的一个或多个氨基酸残基位置的合成多肽或重组多肽。术语“变体”可以与术语“突变体”互换地使用。

[0044] 本文所述的突变CDC可以与一种或多种可药用赋形剂(包括载体、溶媒和稀释剂)组合以形成免疫原性组合物。如本文所用的术语“可药用赋形剂”意指可以在其中布置本文公开的突变CDC(例如,突变肺炎球菌溶血素多肽)以改善溶解度、递送、分散、稳定性和/或构象完整性的溶剂或其他材料。这类可药用赋形剂的例子包括但不限于水、盐水溶液(如生理盐水溶液和中性pH的缓冲盐水溶液(如磷酸盐缓冲盐水(PBS))、乙醇、糖、右旋糖、甘油和/或多元醇(如甘露糖醇和山梨醇)。其他类型的载体包括脂质体或聚合物等。

[0045] 术语“可药用的”指并非生物上不利或另外不利的材料,即,该材料可以随所选择的化合物一起施用至个体,不造成任何不利的生物学效果或不与药物组合物含有的任何其他组分以不利方式相互作用。

[0046] 突变CDC或含有所述突变CDC的免疫原性组合物还可以与佐剂如(但不限于)弗氏

不完全佐剂、弗氏完全佐剂、明矾、单磷酸酯类脂A、磷酸铝或氢氧化铝、QS-21、盐即AlK(SO₄)₂、AlNa(SO₄)₂、AlNH₄(SO₄)₂、二氧化硅、高岭土、和/或碳多核苷酸(即聚IC和聚AU)组合。佐剂的非限制性例子包括QuilA、Alhydrogel等。术语“佐剂”指随免疫原组合物一起给予时,能够增强、加速或延长免疫应答的物质。任选地,本文构思的突变CDC可以与免疫调节剂和免疫刺激剂(如但不限于白介素、干扰素等)组合。许多疫苗和其他药物制剂是本领域技术人员已知的。

[0047] “生物学上有活性的”意指调节生物的生理系统的能力。分子可以因其自身功能性能而有生物活性,或可以基于其激活或抑制具有自身生物学活性的分子的能力而有生物活性。

[0048] 在本文中使用的情况下,术语“免疫原性的”意指物质激发免疫应答的能力。例如,“免疫原性组合物”是包含突变CDC如突变肺炎球菌溶血素多肽的组合物,所述的组合物能够在施用至主题动物时在该动物中激发免疫应答,如产生抗体。术语“疫苗”指施用至受试者以激发对特定抗原的免疫应答的免疫原性组合物。例如,包含本文公开的一种或多种突变肺炎球菌溶血素多肽的疫苗是用于治疗细菌肺炎链球菌所致疾病或病状的疫苗。

[0049] 如本文所用的术语“患者”或“受试者”包括人类和兽医受试者。出于治疗目的,“哺乳动物”指划分为哺乳动物的任何动物,包括(但不限于)人、家养动物(如,但不限于犬和猫)、家畜(如,但不限于奶牛、马、猪、山羊和绵羊)、实验室动物(如,但不限于小鼠、大鼠、兔、豚鼠和栗鼠)、非人类灵长类和任何其他具有乳腺组织的动物。

[0050] “治疗”指治疗性治疗和预防性或防止性措施。需要治疗的那些个体包括但不限于已经患有特定疾病或病症的个体以及面临患有特定疾病或病症的风险的个体(例如,需要预防性/防止性措施的那些)。术语“治疗着”指出于治疗性和/或预防性/防止性目的,向患者施用某药剂。

[0051] “治疗性组合物”或“药物组合物”指可以在体内施用以引起治疗性和/或预防性/防止性效果的药剂。

[0052] 短语“施用治疗有效量”或“施用预防有效量”意在治疗中提供治疗益处、减少疾病发生、预防或管理疾病。治疗有效的具体量可以容易地由普通医疗执业者确定,并且可以根据本领域已知的因素变动,如疾病/癌症的类型、患者病史和年龄、疾病阶段和其他药剂的共施用。

[0053] “病症”是将从多肽治疗中获益的任何病状。这包括慢性和急性病症或疾病,包括使哺乳动物易罹患所讨论病症的那些病理状况。

[0054] 术语“治疗有效量”指生物活性分子或缀合物或其衍生物的数量,所述数量足以在按照本发明构思的方式使用时显示与合理收益/风险比相称的所需治疗作用,无过度的不良副作用(如毒性、刺激作用和变态反应)。治疗作用可以包括,例如但不限于,抑制不想要的组织或恶性细胞生长。受试者的有效量将取决于受试者类型、受试者的体格大小和健康、待治疗病状的性质和严重程度、施用方法、治疗持续时间、并存疗法的性质(如果有的话)、所用的具体制剂等。因此,不可能事先规定确切的“有效量”。然而,基于本文提供的信息,给定情况的有效量可以由本领域普通技术人员仅使用例行实验确定。

[0055] 如本文所用,术语“并存疗法”与术语“联合疗法”和“辅助疗法”互换使用,将理解成意指采用与本公开的药物组合物结合的另一针对疾病的药物治疗需要治疗的患者

或给予其这种药物。这种并存疗法可以是依次疗法,其中首先用一种药物并且随后用另一种药物治疗患者,或者同时给予两种药物。

[0056] 如本文所用的术语“施用”和“施用着”将理解成包括本领域已知的全部施用途径,包括但不限于口服、局部、经皮、肠胃外、皮下、鼻内、粘膜、肌内、腹膜内、玻璃体内和静脉内途径,包括局部施加和全身施加。此外,本公开的组合物(和/或其施用方法)可以设计成使用本领域熟知的配制技术,提供延迟释放、控释或缓释。

[0057] 术语“置换”、“插入”、“添加”和“缺失”在本文中用来指氨基酸序列或核苷酸序列。“置换”指将一个或多个核苷酸或氨基酸分别替换为不同的核苷酸或氨基酸。“插入”或“添加”是核苷酸序列或氨基酸序列中与天然存在的序列相比,已经导致分别添加一个或多个核苷酸残基或氨基酸残基的变化。简而言之,“缺失”定义为核苷酸序列或氨基酸序列中分别不存在一个或多个核苷酸残基或氨基酸残基的变化。

[0058] 氨基酸置换一般是单个残基性质的;插入通常将在约1至20个氨基酸残基的级别上,不过可以耐受明显更大的插入。缺失是从约1至约20个残基,不过在一些情况下,缺失可以大得多。

[0059] 置换、缺失、插入或其任意组合可以用来实现最终的突变多肽。通常,改变一些氨基酸以最大限度减少对分子的改变。然而,在某些情况下可以耐受较大的改变。

[0060] 在某些实施方案中,氨基酸置换可以是将一个氨基酸替换为具有相似结构的和/或化学特性的另一种氨基酸结果,如将异亮氨酸替换为缬氨酸,即,保守性氨基酸置换。插入或缺失可以任选地处于1至5个氨基酸的范围内。

[0061] 在实施方案中,可以根据已知的“保守性置换”进行置换。“保守性置换”指将一个类别中的氨基酸置换成相同类别中的氨基酸,其中类别由氨基酸侧链的常见物理化学特性和自然界存在的同源蛋白中的高置换频率定义。

[0062] 相反,在某些实施方案中,置换是非保守的。“非保守性置换”指将一个类别中的氨基酸置换成源自另一个类别的氨基酸。

[0063] 如本文所用,术语“多肽”指由肽键连接的氨基酸残基的单链组成的化合物。如本文所用的术语“蛋白质”可以同义于术语“多肽”或可以另外指两条或多条多肽的复合物。

[0064] 术语“核酸分子”包括RNA、DNA和cDNA分子。应当理解,作为遗传密码简并性的结果,可以产生编码给定的突变CDC蛋白的多种核苷酸序列。本公开包括其每种可能的变体核苷酸序列,鉴于遗传密码简并性,所述变体核苷酸序列均是可能的。

[0065] “异源”核酸构建体或序列具有其相对于表达它的细胞而言并非天然的部分。相对于控制序列,术语“异源”指这样的控制序列(即,启动子或增强子),所述控制序列在自然界中不发挥作用以调节该控制序列目前正在调节其表达的同基因。通常,异源核酸序列相对于其中存在它们的细胞并非内源或不是其中存在它们的天然基因组的组成部分,反而,已经通过感染、转染、转化、微量注射、电穿孔等添加至细胞。“异源”核酸构建体可以含有与天然细胞存在的控制序列/DNA编码序列组合相同或不同的控制序列/DNA编码序列组合。

[0066] 如本文所用,术语“载体”指为了不同宿主细胞之间转移所设计的核酸构建体。“表达载体”指具有在外来细胞中并入并表达异源DNA片段的能力的载体。众多原核表达载体和真核表达载体是市售的。适宜表达载体的选择处于本领域技术人员知识范围内。

[0067] 因此,“表达盒”或“表达载体”指重组或合成产生的核酸构建体,所述核酸构建体

具有一系列允许特定核酸在靶细胞中转录的指定核酸元件。重组表达盒可以并入质粒、染色体、线粒体DNA、质体DNA、病毒或核酸片段。一般而言，表达载体的重组表达盒部分包括待转录的核酸序列和启动子，连同其他序列。

[0068] 如本文所用，术语“质粒”指作为克隆载体使用并且在许多细菌和某些真核生物中形成染色体外自我复制型遗传元件的环状双链(ds)DNA构建体。

[0069] 如本文所用，术语“编码选择标记的核苷酸序列”指这样的核苷酸序列，它能够在细胞中表达并且其中该选择标记的表达引起含有所表达基因的细胞在相应选择剂存在下或相应选择性生长条件下生长的能力。

[0070] 如本文所用，术语“启动子”指发挥作用以指导下游基因转录的核酸序列。启动子通常将适用于其中正在表达靶基因的宿主细胞。启动子，连同其他转录性和翻译性调节核酸序列(也称作“控制序列”)，是表达给定基因必需的。通常，该转录性和翻译性调节序列包括，但不限于启动子序列、核糖体结合位点、转录起始和终止序列、翻译起始和终止序列和增强子或激活蛋白序列。

[0071] 如本文所用的术语“嵌合基因”或“异源核酸构建体”指可以由不同基因的部分(包括调节元件)组成的非天然基因(即，已经引入宿主的那种)。用于转化宿主细胞的嵌合基因构建体一般由与异源蛋白编码序列或在选择标记嵌合基因中与选择标记基因有效连接的转录调节区(启动子)组成，其中所述选择标记基因编码向转化的细胞赋予抗生素耐药性的蛋白质。本公开的供转化入宿主细胞的常见嵌合基因包括组成型或诱导性型转录调节区、蛋白质编码序列和终止子序列。如果需要分泌靶蛋白，则嵌合基因构建体还可以包含编码信号肽的第二DNA序列。

[0072] 当一种核酸与另一个核酸序列处于功能性关系时，该核酸是“有效连接的”。例如，编码分泌性前导序列的DNA有效连接于编码多肽的DNA，若前者作为参与所述多肽分泌的前蛋白表达；启动子或增强子有效连接于编码序列，若前者影响该序列的转录；或核糖体结合位点有效连接于编码序列，若该结合位点如此安置，从而促进翻译。通常，“有效连接的”意指连接的DNA序列是连续的，并且，在分泌性前导序列情况下，是连续并符合可读框的。然而，增强子不必是连续的。连接通过在便利的限制性位点处连接而完成。如此类位点不存在，则根据常规实践使用PCR的合成性寡核苷酸衔接子、接头或引物。

[0073] 如本文所用，术语“基因”意指这样的DNA区段，所述DNA区段涉及多肽链的产生，可以包含或不包含编码区之前和之后的区域，例如，5'非翻译(5'UTR)或“前导序列”和3'UTR或“尾随序列”，以及各个编码区段(外显子)之间的间插序列(内含子)。

[0074] 如本文所用，术语“重组”包括对细胞或载体的指称，其中所述的细胞或载体已经因异源核酸序列引入而修饰；此外，术语“重组”还可以指从如此修饰的细胞衍生的细胞。因而，例如，重组细胞表达在该细胞的天然(非重组)形式中不以相同形式存在的基因或表达否则因人类故意干预而异常表达、不足表达或根本不表达的天然基因。

[0075] 如本文所用，指称细胞时，术语“转化”、“稳定转化”或“转基因”意指该细胞具有整合入其基因组中的非天然(例如，异源)核酸序列或具有经多个世代维持的附加型质粒。

[0076] 如本文中所用，术语“表达”指基于基因的核酸序列而产生多肽的过程。该过程包括转录和翻译。

[0077] 向细胞插入核酸序列的语境下，术语“引入”涉及向细胞插入核酸序列的任何方

法,包括但不限于“转染”、“转化”和/或“转导”法。术语“引入”还包括将核酸序列并入真核或原核细胞,其中所述核酸序列可以并入该细胞的基因组(例如,染色体、质粒、质体或线粒体DNA),转化成自主复制子或瞬时表达(例如,转染的mRNA)。

[0078] 现在转向本公开,某些实施方案涉及包含一种或多种无毒突变体胆固醇依赖性细胞溶素(CDC)的组合物。组合物可以例如在针对相应疾病病原体的疫苗中使用,或可以用于诊断或筛查方法或其他分析方法如检测方法中。

[0079] 产生天然形式CDC的生物具有各种病理影响,包括但不限于下文所列的那些。

[0080] 产气荚膜梭菌(*Clostridium perfringens*)是经常以肠毒素血症或软组织感染如气性坏疽为特征的多种人类疾病和动物疾病的致病因子。实验证据提示了产气荚膜杆菌溶素0在通过影响嗜中性粒细胞功能而减弱免疫应答中的作用。

[0081] 蜡样芽孢杆菌(*Bacillus cereus*) (腊状芽孢杆菌溶素0的来源)是严重非胃肠道感染的不常见病因,尤其是药瘾者、免疫抑制的新生儿和术后患者中、尤其当插入假体植入物如心室分流器时严重非胃肠道感染的不常见病因。眼感染是最常见类型的重度感染,包括眼内炎(一种眼炎)和角膜炎,通常伴以特征性形成角膜环脓肿。

[0082] 蜂房芽孢杆菌(*Bacillus alvei*)可以造成眼内炎并且可以造成肺炎和积脓。

[0083] 已经显示停乳链球菌似马亚种(*Streptococcus dysgalactiae* subsp. *equisimilis*)涉及许多不同类型的人类疾病综合征。

[0084] 犬链球菌(*Streptococcus canis*)一般在动物中、主要在犬造成疾病。它可以在人类中造成疾病,最经常是软组织感染、菌血症、尿路感染、骨感染或肺炎。

[0085] 链球菌造成多种疾病,包括链球菌性喉炎、风湿热、软组织感染(即,食肉菌)和许多其他疾病。已经显示链球菌溶素0是这些疾病中多者的主要致病因子。

[0086] 破伤风溶素由作为破伤风病因的破伤风梭菌(*Clostridium tetanus*)产生。

[0087] 伊氏李斯特氏菌(*Listeria ivanovii*)是一种动物感染并主要在羊中造成流产。

[0088] 单核增生李斯特菌(*Listeria monocytogenes*)造成人类中的食源性疾病;因其所致的最严重食源性疾病是脑膜炎。它对孕妇尤其带来问题,其中感染可能在母亲中呈亚临床性,但对胎儿致命。李斯特菌溶素是这些疾病的关键致病因子,没有它,细菌无毒力。

[0089] 猪链球菌(*Streptococcus suis*)是猪中败血症、脑膜炎、心内膜炎、关节炎和偶尔地其他感染的病因,并且日益地成为人类中的问题,正在报道越来越多的暴发,伴以多种症状,所述症状包括高热、不适感、恶心和呕吐、随后是神经症状、皮下出血、败血性休克和昏迷。

[0090] 本公开的某些实施方案提供肺炎链球菌的天然(野生型)肺炎球菌溶血素(“PLY”; SEQ ID NO:1)的无毒突变体(由SEQ ID NO:20的突变体编码)。这些PLY突变体显示出几个胜过先前已经用于疫苗开发的肺炎球菌溶血素突变体(Pd-B)的潜在优点,特别地在于相比野生型PLY蛋白,它们基本上缺少溶血活性。例如,本公开的PLY突变体缺少与哺乳动物膜结合的能力,并且因此将不发生当野生型PLY毒素与膜结合(如上文描述的Pd-B突变体(Trp433Phe)所做那样)时通常出现的任何结构性变化。

[0091] 在某些非限制性实施方案中,本公开包括这样的肺炎球菌溶血素突变体,其中(SEQ ID NO:1)的位置293和294的至少一个氨基酸和在位置458、459和460处的至少一个氨基酸已经用不同于野生型PLY序列(SEQ ID NO:1)中所存在的氨基酸置换。更具体地,位

置293和294处的任一个或两个gly残基可以用具有侧链的氨基酸置换,所述氨基酸包括但不限于ala、leu、ile、val、pro、trp、asn、gln、phe、tyr、met、cys、thr、ser、asp、glu、arg、his和lys。另外,位置458和459处的任一个或两个thr残基可以用gly、ala、leu、ile、val、pro、trp、asn、gln、phe、tyr、met、cys、ser、asp、glu、arg、his和lys置换。另外,位置460处的leu残基可以用gly、ala、ile、val、pro、trp、asn、gln、phe、tyr、met、cys、thr、ser、asp、glu、arg、his和lys置换。例如,在一个非限制性实施方案中,位置293处的甘氨酸已经用ala、leu、ile、val、pro、trp、asn、gln、phe、tyr、met、cys、thr、ser、asp、glu、arg、his和lys之一替换,并且位置460处的亮氨酸已经用gly、ala、ile、val、pro、trp、asn、gln、phe、tyr、met、cys、thr、ser、asp、glu、arg、his和lys之一替换。在突变肺炎球菌溶血素的具体非限制性实施方案中,位置293和294处的甘氨酸残基至少之一已经突变成丝氨酸残基或苏氨酸残基,和位置458,459,和460处的苏氨酸残基、苏氨酸残基和亮氨酸残基至少之一已经分别突变成天冬氨酸残基、天冬酰胺残基或谷氨酸残基。例如,当位置293已经突变成丝氨酸并且位置460处的亮氨酸已经突变成天冬氨酸时,肺炎球菌溶血素突变体命名为PLY-G293S/L460D(或PLY-L460D/G293S);其氨基酸序列作为SEQ ID NO:40提供。在替代性非限制性实施方案中,位置460可以用D、E或N置换,并且位置293可以用S或T置换,从而双突变体可以在位置460处包含D、E或N并且在位置293处包含S或T。

[0092] 除肺炎球菌溶血素的突变体之外,本公开还提供在结构域4的环1、环2和/或环3中的同功位置具有置换的其他突变CDC,包括以下溶素的突变体:蜡状芽孢杆菌溶素(蜡状芽孢杆菌(*Bacillus cereus*))、芽孢杆菌溶素(炭疽芽孢杆菌(*Bacillus anthracis*))、苏云金芽孢杆菌溶素(苏云金芽孢杆菌(*Baccillus thuringiensis*))、产气荚膜杆菌溶素(产气荚膜梭菌(*Clostridium perfringens*))、蜂房芽孢杆菌溶素(蜂房芽孢杆菌(*Bacillus alvei*))、犬链球菌溶素(犬链球菌(*Streptococcus canis*))、类马链球菌溶素(马链球菌(*Streptococcus equisimilis*))、链球菌溶素0(化脓性链球菌(*Streptococcus pyogenes*))、破伤风溶素(破伤风梭菌(*Clostridium tetani*))、伊万诺夫李斯特菌溶素(伊氏李斯特氏菌(*Listeria ivanovii*))、李斯特菌溶素(单核增生李斯特菌(*Listeria monocytogenes*))、斯氏李斯特菌溶素(斯氏李斯特菌(*Listeria seeligeri*))、猪溶素(猪链球菌(*Streptococcus suis*))、缓症链球菌溶素(缓症链球菌(*Streptococcus mitis*))、血小板聚集因子(又称作PAF和草绿色链球菌溶素)(缓症链球菌(*Streptococcus mitis*))、中间链球菌溶素(中间链球菌(*Streptococcus intermedius*))、化脓隐秘杆菌溶素(化脓隐秘杆菌(*Arcanobacterium pyogenes*))和诺维氏梭菌溶素,又称作破伤风溶素NT(诺维氏梭菌(*Clostridium novyi*))。

[0093] 蜡状芽孢杆菌溶素(Cereolysin)、芽孢杆菌溶素(Anthrolysin)、苏云金芽孢杆菌溶素(Thuringiolysin)(又称作苏云金溶素或蜡状芽孢杆菌溶素形式BT)、产气荚膜杆菌溶素(Perfringolysin)、蜂房芽孢杆菌溶素(Alveolysin)、犬链球菌溶素(Caniolysin)、类马链球菌溶素(Caniolysin)、链球菌溶素0(Streptolysin 0)、诺维氏梭菌溶素(Novyiolysin)、破伤风溶素(Tetanolysin)、伊万诺夫李斯特菌溶素(Ivanolysin)、李斯特菌溶素(Listeriolysin)、斯氏李斯特菌溶素(Seeligeriolysin)、猪溶素(Suilysin)、缓症链球菌溶素(Mitilysin)、中间链球菌溶素(Intermedilysin)、血小板聚集因子(又称作草绿色链球菌溶素(Viridanolysin)或PAF)和化脓隐秘杆菌溶素(Pyolysin)的野生型氨基酸

序列分别在SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:18和SEQ ID NO:19中显示。

[0094] 本公开的另一个实施方案涉及一种链球菌溶素O突变体,所述突变体在SEQ ID NO:9的位置561和562至少之一中及其位置395和396至少之一中包含置换并且与野生型链球菌溶素O蛋白至少90%相同。位置561和562中的置换可以是本文所述的可以在PLY (SEQ ID NO:1)的位置458-460中作出的任何置换,并且位置395和396中的置换可以是本文所述的可以在PLY的位置293或294中作出的任何置换。

[0095] 也可以根据本公开突变的SEQ ID NO:18(血小板聚集因子)的变体是凝集素溶素,其也从缓症链球菌获得。L1、L2和L3的氨基酸序列与PAF相同。凝集素溶素在12个位置不同于PAF,所述位置包括67、158、211、303、305-307、311、319、327、447和556,其中在凝集素溶素中,这些位置处的氨基酸分别是T、D、T、H、E、N、K、N、E、K、T和I。本公开因此包括与本文构思的那些其他突变体相似的凝集素溶素突变体和编码这些突变体的核酸,以及包含这些突变体的组合物。

[0096] 本文构思的肺炎球菌溶血素突变体还消除了毒素的任何毒力活性,原因在于它们不能与哺乳动物细胞结合。尽管肺炎球菌溶血素突变体Pd-B的毒性比天然肺炎球菌溶血素低约21,000倍,但它仍显示在开发包含它的任何疫苗时棘手的足够毒性。开发对抗肺炎链球菌的现代疫苗似乎以使用肺炎球菌溶血素连同其他肺炎链球菌衍生的蛋白质为中心;因此,无论疫苗中所用的其他蛋白质是什么,似乎在对抗肺炎链球菌的全部有效疫苗中均将包含肺炎球菌溶血素,原因在于其对疾病建立和进展的重要性。

[0097] 如下文描述,在产气荚膜杆菌溶素(一种与肺炎球菌溶血素相关的毒素)中显示该蛋白质的十一肽不介导这些毒素与哺乳动物细胞的结合,有悖于传统观点。介导结合的结构是靠近十一肽的三个疏水性短环。作为本公开的部分,现在已知,如果将带负电荷的天冬氨酸残基或谷氨酸残基(例如)置入内部任意单个疏水性环(在尚未包含天冬氨酸或谷氨酸的位置),则阻断CDC与膜的结合。因此,这个单点突变消除了CDC(包括肺炎球菌溶血素)与哺乳动物膜的结合。例如,将单个天冬氨酸残基或谷氨酸残基置换肺炎球菌溶血素的亮氨酸460实际上彻底地消除其溶血活性。由于已知在其他系统(下文描述)中,这种突变阻断与细胞的膜结合,它基本上消除任何毒力活性(例如令其毒性比Pd-B突变体低至少200倍),还消除了可能由其与哺乳动物膜表面结合引起的任何可能的副作用。

[0098] 在某些实施方案中,本公开的突变肺炎球菌溶血素缺少肺炎链球菌肺炎球菌溶血素蛋白中天然存在的溶血活性和成孔能力。通常,多肽组分显示小于约30%、小于约20%、小于约10%、小于约5%、小于约1%、小于约0.1%、小于约0.001%或更小的天然存在性肺炎链球菌肺炎球菌溶血素蛋白的溶血活性。

[0099] 在某些实施方案中,本公开的突变肺炎球菌溶血素在位置293、370、406或460(包括位置290、291、292、294、295、296、367、368、369、371、372、373、403、404、405、407、408、409、457、458、459、461、462和463)的任一旁侧分布的三个残基中一个或多个残基处具有置换。

[0100] 例如,这些残基可以用带负电荷的氨基酸、谷氨酸或天冬氨酸(例外是在位置403

中,所述位置已经包含天冬氨酸)或带正电荷的氨基酸赖氨酸、精氨酸或组氨酸(例外是在位置367和407中,所述位置已经包含组氨酸残基)置换。备选地,这些残基可以用消除突变体的结合活性、成孔活性和/或溶血活性的任何其他天然氨基酸(包括gly、ala、leu、ile、val、pro、trp、asn、gln、phe、tyr、met、cys、thr或ser)置换。

[0101] 如上文所示,野生型肺炎球菌溶血素的氨基酸序列是SEQ ID NO:1,并且编码SEQ ID NO:1的肺炎球菌溶血素的cDNA的反向互补物作为SEQ ID NO:20显示。本公开还包括本文所述的突变肺炎球菌溶血素(及其反向互补物)和其他突变CDC的cDNA,其中必要时,置换所述cDNA以编码本文所述或否则能够实现的置换型蛋白质(突变体),并且所述cDNA可以转而包含任何保守碱基(核苷酸)置换以产生编码这类突变体的cDNA。

[0102] 将可以理解,编码本文构思的多肽的多核苷酸序列可以用简并密码子改变,然而仍编码本公开的突变多肽。因此,本公开还提供与本文所述的多核苷酸序列(或其互补序列)杂交的多核苷酸,所述多核苷酸序列在各序列之间具有至少90%同一性、或至少95%同一性、或至少99%同一性。

[0103] 图1A至图1E显示本文鉴定的天然形式CDC的氨基酸序列的比对结果。序列沿着三个疏水性环比对,所述的疏水性环对应于肺炎球菌溶血素的位置367-373(第二环,L2)、403-409(第三环,L3)和457-463(第一环,L1),在图1A-E中作为位置586-592(第二环,L2)、622-628(第三环,L3)和676-682(第一环,L1)表示。如上文所示,这些CDC的突变体的某些具体(但是非限制性)实施方案可以在这些位置的一个或多个位置处包含带负电荷的氨基酸-谷氨酸或天冬氨酸置换(在该位置已经具有天冬氨酸时例外)或带正电荷的氨基酸组氨酸、赖氨酸或精氨酸置换(例外是其中该位置已经具有组氨酸时的组氨酸置换、其中该位置已经具有赖氨酸时的赖氨酸置换、其中该位置已经具有精氨酸时的精氨酸置换)或上文所示的任何其他15种天然氨基酸置换,其中所得到的突变体根据本公开发挥作用。

[0104] 突变体还可以包含多于一个在本文所述的置换,从而突变体在单个环(L1、L2、L3)中具有1、2、3、4、5、6或7个置换的残基,或突变体可以在两个环(例如,L1和L2、L1和L3、L2和L3)中具有一个或多个(1至7个)置换的残基,或在三个环(L1、L2和L3)的每个环中具有一个或多个(1至7个)置换的残基,其中置换选自本文中列出的那些;例如,突变体可以在第一环(L1)中具有1至7个置换,和/或在第二环(L2)中具有1至7个置换,和/或在第三环(L3)中的具有1至7个置换。例如,在某些实施方案中,其中天然残基带正电荷的情况下,置换的残基可以带负电荷,并且其中天然残基带负电荷的情况下,置换的残基可以带正电荷。备选地,天冬氨酸可以用谷氨酸、组氨酸、精氨酸或赖氨酸置换,或谷氨酸可以用天冬氨酸、赖氨酸、组氨酸或天冬酰胺置换,或精氨酸可以用不同的带正电荷的氨基酸置换。

[0105] 表1中列出了本文所述的每种CDC的环1、环2和环3的氨基酸位置。

[0106] 表1:与结构域4环相对应的氨基酸位置

	SEQ ID NO.	环 1	环 2	环 3
肺炎球菌溶素	1	457-463	367-373	403-409
腊状芽孢杆菌溶素	2	498-504	408-414	444-450
芽孢杆菌溶素	3	501-507	411-417	447-453
苏云金芽孢杆菌溶素	4	501-507	411-417	447-453
产气荚膜杆菌溶素	5	488-494	398-404	434-440
蜂房芽孢杆菌溶素	6	490-496	400-406	436-442
犬链球菌溶素	7	562-568	472-478	508-514
类马链球菌溶素	8	559-565	469-475	505-511
[0107] 链球菌溶素 O	9	559-565	469-475	505-511
诺维氏梭菌溶素	10	502-508	412-418	448-454
破伤风溶素	11	514-520	424-430	460-466
伊万诺夫李斯特菌溶素	12	512-518	422-428	458-464
李斯特菌溶素 O	13	513-519	423-429	459-465
斯氏李斯特菌溶素	14	514-520	424-430	460-466
猪溶素	15	484-490	395-401	431-437
缓症链球菌溶素	16	457-463	367-373	403-409
中间链球菌溶素	17	515-521	425-431	461-467
PAF	18	651-657	561-567	597-603
化脓隐秘杆菌溶素	19	521-527	431-437	467-473

[0108] 因此,本文中提供纯化或分离形式的蛋白质突变体及其抗原性片段、这些突变体的包含可药用赋形剂、佐剂和/或免疫刺激剂的免疫原性组合物及包含本文公开或否则构思的一种或多种突变体的疫苗和血清。突变体或其抗原性片段可以用于分析方法中,所述分析方法利用本领域已知的技术(例如ELISA)检测生物样品中可变形式蛋白质的存在。本公开还提供了包含编码本文提供的任何突变体的cDNA的核酸、宿主细胞和载体以及使用前者产生本文构思的突变体的方法。本公开还提供了施用免疫原性组合物以治疗由本文所述产生CDC的生物所致病状、疾病和感染的方法。

[0109] 如上文所示,本公开也涉及编码本文构思的突变CDC的核酸序列。本公开提供核酸,所述核酸编码本文公开的蛋白质突变体的等位变体,其中蛋白质突变体的等位变体与蛋白质突变体差异小于15%其氨基酸同一性,例如,等位变体的至少85%氨基酸与蛋白质突变体相同,并且第一环、第二环和第三环(L1、L2和L3)中的100%氨基酸与蛋白质突变体中的那些氨基酸相同。例如,等位变体可以与本文所述的蛋白质突变体差异小于12%其氨

氨基酸同一性、小于10%其氨基酸同一性、小于8%其氨基酸同一性、小于6%其氨基酸同一性、小于4%其氨基酸同一性、小于2%其氨基酸同一性、或小于1%其氨基酸同一性。另外，本公开还涉及核酸，所述核酸在严格条件下与编码本文所述的突变CDC的核酸或与编码本文所述的突变CDC的核酸的互补物杂交。

[0110] 在一个方面，本公开的CDC突变多肽或蛋白质包含这样的氨基酸序列，所述氨基酸序列与作为SEQ ID NO:1提出的序列具有至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%或更多同一性百分数(如通过序列比对程序所测定)并且具有至少一个在本文他处所述的突变。

[0111] 可以使用例如在MacVector第6.5版中以默认参数运行的CLUSTAL-W程序，执行所选择序列的比对，旨在测定的两个或更多个序列之间的“同一性%”，所述默认参数包括开放空位罚分10.0、延伸空位罚分0.1，和BLOSUM30相似性矩阵。

[0112] 在另一个实施方案中，如本文所用的术语“序列同一性”意指如下比较序列。使用默认(BLOSUM62)矩阵(值-4至+11)，连同空位开口罚分-12(对于开口的首个无效)和空位延伸罚分-4(空位中每个额外的连续无效)，使用第9版Genetic Computing Group's GAP(全局比对程序)比对序列在比对后，通过将匹配数目表述为所要求的序列中氨基酸数目的百分数，计算同一性百分数。

[0113] 本文所述或否则构思的免疫原性组合物可以包括疫苗制剂，所述的疫苗制剂可以按有效在动物中激发(刺激)保护性免疫应答的量使用。例如，可以通过抗体的形成过程测量保护性免疫应答的产生。在某些非限制性实施方案中，例如，按免疫接种之间约1周至6周的间距，本文构思的突变CDC的可以形成保护性免疫应答的量一般处于约0.001 μ g至100mg/kg体重如但不限于约0.01 μ g至1mg/kg体重、或约0.1 μ g至约10 μ g/kg体重的单位剂量形式。

[0114] 本公开还提供了刺激针对至少一种疾病生物的免疫应答的方法。在方法中，本文公开的任何免疫原性组合物可以施用至遭疾病生物感染或易遭该疾病生物感染的患者。在一个非限制性实施方案中，免疫原性组合物包含在位置293和460中具有突变如PLY_{L460D/G293S}的肺炎球菌溶血素突变体(SEQ ID NO:40)。在该方法中，免疫原性组合物基本上无毒(或与天然PLY蛋白相比基本上无毒)，基本上与细胞膜不结合，基本上无溶血性，和/或与PLY蛋白同样稳定或基本上比PLY蛋白更稳定。

[0115] 本公开还涉及至少一种在患者中减少感染发生和/或其严重程度的方法。在该方法中，将本文公开或否则构思的任何免疫原性组合物施用至感染的患者或易遭感染的患者。在一个非限制性实施方案中，免疫原性组合物包含在位置293和460中具有突变如PLY_{L460D/G293S}的肺炎球菌溶血素突变体(SEQ ID NO:40)。在该方法中，免疫原性组合物基本上无毒(或与天然PLY蛋白相比基本上无毒)，基本上与细胞膜不结合，基本上无溶血性，和/或与天然PLY蛋白同样稳定或基本上比天然PLY蛋白更稳定。在某些实施方案中，本文公开的突变肺炎球菌溶血素多肽具有比野生型肺炎球菌溶血素多肽低约100,000倍的溶血活性。在其他实施方案中，本文公开的突变肺炎球菌溶血素多肽具有比野生型肺炎球菌溶血素多肽低约150,000倍的溶血活性。在其他实施方案中，本文公开的突变肺炎球菌溶血素多肽具有比野生型肺炎球菌溶血素多肽低约200,000倍的溶血活性。在另外的实施方案中，本文公开的突变肺炎球菌溶血素多肽具有比野生型肺炎球菌溶血素多肽低约250,000倍的溶血活性。在至少某些实施方案中，相对于在氨基酸位置293、294、458、459和460中仅一个氨

基酸位置具有置换的突变肺炎球菌溶血素多肽,本文公开的在氨基酸位置293、294、458、459和460中具有至少两个置换的纯化突变肺炎球菌溶血素多肽在纯化时还具有增加的产量。增加的重组产量可以例如是至少约10倍,至少约15倍,至少约17倍,或至少约20倍。

[0116] 本文公开的免疫原性组合物可以施用至遭本文所述的疾病生物感染或可能遭其感染的动物,包括但不限于犬、猫、兔、啮齿类、马、家畜(例如,牛、绵羊、山羊和猪)、动物园动物、有蹄动物、灵长类和人。

[0117] 如上文所示,当突变体是肺炎球菌溶血素突变体时,本公开包括可以施用至受试者以在受试者中刺激免疫原性应答的免疫原性组合物(如,但不限于疫苗)。除一种或多种肺炎球菌溶血素突变体之外,免疫原性组合物/疫苗还可以包含来自肺炎链球菌的其他蛋白质或蛋白质亚基,或可以包含与免疫原性组合物/疫苗中的肺炎球菌溶血素突变体或其他蛋白质组合或缀合的荚膜多糖物质。例如,荚膜物质可以衍生自肺炎链球菌血清型1、2、3、4、5、6A、6B、7F、8、9N、9V、10A、11A、12F、14、15B、17F、18C、19A、19F、20、22F、23F、24F、27、33F或34或本领域已知的其他血清型中任一者或多者。如所示,免疫原性组合物/疫苗可以包含佐剂和/或其他可药用赋形剂。多糖可以缀合于突变体,例如,借助单聚体键(仅多糖的一个末端与多肽连接)、成环键(单一多肽与成环多糖连接)或交联(多个多糖与多个多肽连接)缀合。

[0118] 含有本公开的突变肺炎球菌溶血素多肽或其片段的免疫原性组合物或疫苗可以用来治疗与肺炎链球菌相关的疾病和病状,如,但不限于肺炎、脑膜炎、菌血症和中耳炎。

[0119] 在某些实施方案中,本文公开的突变CDC用于刺激T细胞增殖或通过刺激B细胞产生抗体。

[0120] 如上文所示,可以通过将本文构思的突变CDC与药物(生理)可接受的赋形剂如(但不限于)生理盐水或中性pH的缓冲盐水溶液(如磷酸盐缓冲盐水)组合,形成本公开的免疫原性组合物。

[0121] 本公开还包括本文所述或否则构思的突变CDC的抗原性片段。例如,对于疫苗组合物,片段大到足以刺激保护性免疫应答。多肽组分必须具有足以诱导这种增强型免疫应答的长度。对于天然存在的CDC蛋白的片段,这些片段具有至少约8个、至少约10个、至少约25个、至少约50个、至少约75个、至少约100个、至少约125个、至少约150个、至少约175个、至少约200个、至少约250个、至少约300个、至少约350个、至少约400个、至少约425个、至少约450个、至少约460个、至少约465个或更大氨基酸长度。

[0122] 片段可以包含来自突变体不同位置的已经连接在一起的肽部分。在某些具体(但非限制性)实施方案中,片段包含本文中讨论的三种环中的一者或多者。

[0123] 本文公开或否则构思的突变CDC还可用于产生中和抗体,所述中和抗体可以作为被动免疫血清用来治疗或改善患者中的症状。如上文所述的免疫原性组合物可以施用至动物(如马或人)直至生成中和抗体应答。随后可以收获、纯化这些中和抗体并用来治疗显示症状的患者。

[0124] 将这类中和抗体以有效消除病原体影响的量施用至显示疾病症状的患者。中和抗体可以静脉内、肌内、皮内、皮下等施用。具体途径是静脉内施用,或对于局限化感染,在组织损伤部位伴随清创术一起局部施用。中和抗体也可以与抗生素治疗相结合施用。中和抗体可以按单剂或多剂施用直至实现休克或组织损伤减少。一般施用的中和抗体的量是约

1mg至约1000mg抗体/kg体重,如但不限于,约50mg至约200mg抗体/kg体重。

[0125] 本公开的免疫原性组合物可以作为药物组合物制备,所述药物组合物含有处于无毒和无菌可药用赋形剂中的免疫保护性无毒量的本发明公开的至少一种突变蛋白。

[0126] 本公开的免疫原性组合物可以按本领域已知的任何适合方式施用至适宜的受试者,所述方式包括(但不限于)口服、肌内、静脉内、舌下粘膜、动脉内、鞘内、皮内、腹腔内、鼻内、肺内、眼内、阴道内、直肠内和/或皮下。可以将它们引入胃肠道或呼吸道,例如,通过吸入含有免疫原性组合物的溶液剂或散剂。如果使用,肠胃外施用法通常以注射为特征。注射剂可以按常规形式制备,或者作为液体溶液剂或混悬剂、适于在注射前形成溶液剂或混悬剂的固体形式,或者作为乳剂。

[0127] 免疫原性组合物(例如,疫苗)以足够激发抗体生成作为免疫原性应答的组成部分的量施用。对于任何给定患者的剂量取决于多种因素,包括患者的体格大小、总体健康状况、性别、身体表面积、年龄、待施用的具体化合物、施用时间和途径以及正在同时施用的其他药物。最佳剂量的确定完全处于具有普通技术的药理学专家的能力范围内。在某些实施方案中,可以施用至受试者的突变CDC的非限制性有效量范围例如是约10ng蛋白质至100mg/kg体重,如约0.1 μ g蛋白质至约1mg/kg体重。在至少一个非限制性实施方案中,提供的剂量处于约0.25 μ g至约25 μ g蛋白质范围内,伴以或不伴以佐剂。

[0128] 当通过肌内或深皮下途径肠胃外施用免疫原性组合物时,突变蛋白可以(在某些具体但非限制性实施方案中)与任何常规佐剂混合或用其吸附以吸引或增强免疫应答。这类佐剂包括但不限于氢氧化铝、磷酸铝、胞壁酰二肽、细菌脂多糖和衍生物和来自QuilA的纯化皂苷。蛋白质也可以在微粒子如脂质体或免疫刺激复合物(ISCOM)内部呈递至免疫系统。如所示,可以设计含有本公开的突变蛋白/肽片段的制剂供口腔或鼻内摄入。

[0129] 本公开的免疫原性组合物的治疗有效和无毒剂量可以由本领域普通技术人员确定。例如,任何受试者的具体剂量可能取决于多种因素,包括但不限于)年龄、总体健康状况、患者的膳食、施用时间和途径、与正在施用的其他药物的协同效应和免疫原性组合物是否反复施用。如果需要,免疫原性组合物将反复施用,每个剂量之间间隔一个月至三个月并且稍后时间任选的加强剂量。对于本领域技术人员,制备适宜剂型的实际方法是已知或将显而易见;参见例如,Remington's Pharmaceutical Sciences最新版。

[0130] 如上文所示,本公开包括编码本文所述的本公开的突变多肽和活性片段的多核苷酸。多核苷酸可以处于RNA形式或处于DNA形式(包括但不限于,cDNA、基因组DNA和合成性DNA)。DNA可以是双链或单链的,并且如果是单链的,则可以是编码链或非编码(反义)链。

[0131] 表2中显示了直接编码(或借助反向互补物编码)本文构思的天然序列CDC并且因此也可以突变以形成本文所述或否则构思的突变体形式的DNA序列(和相应的氨基酸序列)。

[0132] 表2:天然CDC形式的氨基酸序列和核酸序列

	SEQ ID NO: (氨基酸)	SEQ ID NO: (核酸)	
[0133]	肺炎球菌溶素	1	20
	腊状芽孢杆菌溶素	2	21
	芽孢杆菌溶素	3	22
	苏云金芽孢杆菌溶素	4	23
	产气荚膜杆菌溶素	5	24
	蜂房芽孢杆菌溶素	6	25
	犬链球菌溶素	7	26
	类马链球菌溶素	8	27
	链球菌溶素 O	9	28
	诺维氏梭菌溶素	10	29
[0134]	破伤风溶素	11	30
	伊万诺夫李斯特菌溶素	12	31
	李斯特菌溶素 O	13	32
	斯氏李斯特菌溶素	14	33
	猪溶素	15	34
	缓症链球菌溶素	16	35
	中间链球菌溶素	17	36
	PAF	18	37
	化脓隐秘杆菌溶素	19	38

[0135] 宿主细胞经载体基因工程化(转导、转化和/或转染),所述载体包含编码本公开的突变多肽的多核苷酸。载体例如可以处于质粒、病毒粒子、噬菌体等形式。工程化的宿主细胞可以在常规的营养培养基中培养,其中如适宜,可以调整所述的营养培养基以激活启动子、选择转化体或扩增编码这类多肽的多核苷酸。培养条件如温度、pH等是先前随被选择用于表达的宿主细胞所用的那些培养条件,并且将对普通技术人员是显而易见的。载体包括染色体性、非染色体性及合成性DNA序列,例如SV40的衍生物;细菌质粒;噬菌体DNA;杆状病毒;酵母质粒;从质粒和噬菌体DNA的组合中衍生的载体、病毒DNA如痘苗病毒、腺病毒、禽痘病毒和伪狂犬病病毒。然而,可以使用任何其他的载体,只要它在宿主中是可复制和有活力的。

[0136] 适宜的DNA序列可以通过多种方法插入载体。通常,将DNA序列通过本领域已知的方法插入适宜限制性核酸内切酶位点。认为此类方法和其他方法处于本领域技术人员的能力范围内。

[0137] 表达载体中的DNA序列与适宜的表达控制序列(启动子)有效连接以指导mRNA合成。作为这类启动子的代表性例子,可以提到:LTR或SV40启动子、大肠杆菌lac或trp、噬菌体 λ P_L启动子以及已知控制基因在原核细胞或真核细胞或其病毒中表达的其他启动子。表达载体还含有用于翻译起始的核糖体结合位点和转录终止子。载体还可以包括用于扩增表达的适宜序列。

[0138] 此外,在某些非限制性实施方案中,表达载体含有一个或多个选择标记基因以提供表型性状供选择转化的宿主细胞,如用于真核细胞培养物的二氢叶酸还原酶或新霉素抗性,或如大肠杆菌中的四环素抗性或氨苄青霉素抗性。

[0139] 含有如上文所述的适宜DNA序列以及适宜启动子或控制序列的载体可以用来转化适宜的宿主,以引起宿主表达蛋白质。

[0140] 作为适宜宿主的代表性(但非限制性)例子,可以提到:细菌细胞,如大肠杆菌、链霉菌、鼠伤寒沙门氏菌(*Salmonella typhimurium*);真菌细胞,如酵母;昆虫细胞,如果蝇S2和夜蛾属(*Spodoptera*) Sf9;动物细胞,如CHO、COS或Bowes黑素瘤;腺病毒;植物细胞等。从本文的教导内容中认为适宜宿主细胞的选择处于本领域技术人员的能力范围内。

[0141] 更具体地,本公开还包括重组构建体,所述重组构建体包含本文所述和能够实现的一个或多个序列。构建体包括已经按正向或反向插入多核苷酸序列的载体,如质粒或病毒载体。在一个非限制性实施方案中,构建体还包含调节序列,例如,所述调节序列包括与该序列有效连接的启动子。大量合适的载体和启动子是本领域技术人员已知的,并且是市售的。通过非限制例子提供以下载体。细菌载体:pQE70、pQE60、pQE-9(Qiagen, Inc., Hilden, 德国)、pBS、pD10、phagescript、psiX174、pbluescript SK、pBS、pNH8A、pNH16a、pNH18A、pNH46A(Stratagene, San Diego, CA); ptrc99a、pKK223-3、pKK233-3、pDR540、pRIT5(Pharmacia, 斯德哥尔摩, 瑞典)。真核载体:pWLNE0、pSV2CAT、pOG44、pXT1、pSG(Stratagene, San Diego, CA) pSVK3、pBPV、pMSG、pSVL(Pharmacia, 斯德哥尔摩, 瑞典)。然而,可以使用任何其他的质粒或载体,只要它们在宿主中是可复制和有活力的。

[0142] 使用CAT(氯霉素转移酶)载体或带选择标记的其他载体,启动子区可以选自任何目的基因。两种适宜的载体是pKK232-8和pCM7。具体命名的细菌启动子包括lacI、lacZ、T3、T7、gpt、 λ P_R、P_L和TRP。真核启动子包括CMV立即早期启动子、HSV胸苷激酶启动子、SV40早期和晚期启动子、来自逆转录病毒LTR和小鼠金属硫蛋白-I启动子。选择适宜载体和启动子完全处于本领域普通技术人员的能力范围内。

[0143] 在又一个实施方案中,本公开包括含有上述构建体的宿主细胞。宿主细胞可以是高等真核细胞,如(但不限于)哺乳动物细胞;低等真核细胞,如(但不限于)酵母细胞;或原核细胞,如(但不限于)细菌细胞。可以通过磷酸钙转染、DEAE-葡聚糖介导转染、电穿孔(Davis等人, *Basic Methods in Molecular Biology* (1986) Elsevier Science Publishing Co., Inc., New York, NY) 或任何其他合适的技术实现向宿主细胞引入构建体。

[0144] 可以按常规方式使用宿主细胞中的构建体以产生重组序列编码的基因产物。备选地,本公开的多肽可以通过常规的肽合成仪合成产生。

[0145] 成熟的蛋白质可以在哺乳动物细胞、酵母细胞、细菌细胞或其他细胞中在适宜的启动子控制下表达。使用从本公开的DNA构建体衍生的RNA,也可以使用无细胞翻译系统产生此类蛋白质。随原核宿主及真核宿主一起使用的适宜克隆载体和表达载体由Green和Sambrook (Molecular Cloning:A Laboratory Manual,第4版,Cold Spring Harbor,N.Y.,(2012)描述,所述文献的完整公开内容因而通过引用的方式并入。

[0146] 可以通过将增强子序列插入载体,增加高等真核生物转录编码本公开的突变多肽的DNA。增强子是作用于启动子以增加其转录的顺式作用DNA元件,通常约10至约300bp。例子包括在复制起点晚期侧bp 100至270上的SV40增强子、细胞巨化病毒早期启动子增强子、在复制起点晚期侧上的多瘤增强子和腺病毒增强子。

[0147] 通常,重组表达载体将包含允许转化宿主细胞的复制起点和选择标记,例如,大肠杆菌氨苄青霉素耐药基因和酿酒酵母 (*S.cerevisiae*) TRP1基因,和从高度表达的基因衍生以指导下游结构性序列转录的启动子。这类启动子可以从编码糖酵解酶如3-磷酸甘油酸激酶(PGK)、 α -因子、酸性磷酸酶或热休克蛋白的操纵子等衍生。异源结构性序列在适宜阶段中随翻译起始序列和终止序列一起装配。任选地,异源序列可以编码包含N端识别肽的融合蛋白,所述识别肽赋予所需的特征,例如,稳定作用或简化已表达重组产物的纯化。

[0148] 通过将编码所需蛋白质的结构性DNA序列连同合适的翻译起始信号和终止信号在有效读取阶段,随功能启动子一并插入,构建供细菌使用的有用表达载体。载体将包含一个或多个表型选择标记和复制起点以确保维持载体并且如果需要,以在宿主内部提供扩增。

[0149] 作为代表性但非限制性例子,供细菌使用的有用表达载体可以包含从市售质粒衍生的选择标记和细菌复制起点,所述市售质粒包含熟知克隆载体pBR322(ATCC 37017)的遗传元件。这类商业载体例如包括pKK223-3(Amersham Pharmacia Biotech,Piscataway,N.J.,美国)和pGEM1(Promega, Madison,Wis.,美国)。这些pBR322“主链”片段与适宜的启动子和待表达的结构性序列组合。

[0150] 在转化合适的宿主菌株和培育宿主菌株至适宜的细胞密度后,通过适宜手段(例如,温度转变或化学诱导)诱导所选择的启动子并且将细胞培养额外一段时间。

[0151] 一般通过离心收获细胞,通过物理或化学手段破碎,并且保留所产生的粗提物供进一步纯化。

[0152] 可以通过任何便利方法破碎在表达蛋白质中所用的微生物细胞,所述方法包括冻融循环、超声处理、弗氏压碎器、机械破裂或使用细胞裂解剂,这类方法是本领域技术人员熟知的。但是,可以想要(但是不限制)使用分泌本公开的多肽并允许从培养基回收多肽的宿主细胞。

[0153] 多种哺乳动物细胞培养系统也可以用来表达重组蛋白。哺乳动物表达系统的例子包括Gluzman (Cell (1981) 23:175)描述的猴肾成纤维细胞COS-7系,和能够表达相容性载体的其他细胞系,例如,C127、3T3、CHO、HeLa和BHK细胞系。哺乳动物表达载体将包含复制起点、合适的启动子和增强子,并且还包含任何必需的核糖体结合位点、多腺苷酸化位点、剪接供体位点和受体位点、转录终止序列和5'侧翼非转录序列。从SV40剪接(SV40splice)和多腺苷酸化位点衍生的DNA序列可以用来提供必需的非转录遗传元件。

[0154] 可以通过熟知的蛋白质回收和纯化方法,从重组细胞培养物回收和/或纯化多肽。这类方法学可以包括硫酸铵或乙醇沉淀法、酸提取法、阴离子或阳离子交换色谱法、磷酸纤

纤维素色谱法、疏水相互作用色谱法、亲和色谱法、羟基磷灰石色谱法和凝集素色谱法。根据需要,可以在完成成熟蛋白的构象时使用蛋白质再折叠步骤。在这个方面,伴侣蛋白可以用于这个再折叠方法中。最后,高效液相色谱法(HPLC)可以用于最终纯化步骤。

[0155] 在本公开中可用作免疫原的突变多肽可以是化学合成方法的产物或来自原核宿主或真核宿主(例如,培养的细菌细胞、酵母细胞、高等植物细胞、昆虫细胞和哺乳动物细胞)的重组技术产物,如先前解释。取决于重组生产方法中所用的宿主,本公开的突变多肽可以是糖基化的或可以是非糖基化的。

[0156] 可以通过本领域熟知的重组表达/分离方法分离单独表达的多肽。这类分离方法的常见例子可以利用抗体,所述抗体针对蛋白质的保守区域或针对作为蛋白质结构的组成部分表达的His标签或可切割前导序列或尾随序列。

[0157] 如所示,本文公开或否则构思的片段和变体CDC突变蛋白将视为本公开的部分。片段是具有下述氨基酸序列的变体多肽,所述氨基酸序列完全与天然或突变多肽的氨基酸序列的部分(而非全部)相同。片段可以是“无支撑”或含于更大的多肽内部,所述多肽的片段形成部分或区域,如(但不限于)单一连续区域。非限制的特定片段是这样的生物活性片段,它们是介导本公开多肽的活性的那些片段,包括活性相似或活性改善或活性降低的那些。还包括是在动物、尤其人类中有抗原性或免疫原性的那些片段。在这个方面,本公开包括:(i) 突变CDC的片段,例如(但不限于)长度至少约20-100个氨基酸,或长度约100-200个氨基酸的片段,和(ii) 包含突变片段的药物组合物。

[0158] 在一个实施方案中,编码本文所述的CDC突变体的核酸与相应的天然序列在高严格性杂交条件下杂交。高严格性条件的例子包括在约42°C于50%甲酰胺,5×SSC,5×Denhardt溶液,0.5%SDS和100μg/ml变性载体DNA中杂交,随后在室温于2×SSC和0.5%SDS中洗涤两次并且在42°C于0.1×SSC和0.5%SDS中额外洗涤两次。

[0159] 核酸构建体/表达载体

[0160] 如所示,可以将本文构思的核酸并入能够向宿主细胞引入并在其中复制的异源核酸构建体或载体中。可以使用任何其他的载体,只要该载体在引入它的细胞中可复制和有活力。大量合适的载体和启动子是本领域技术人员已知的,并且是市售的。适宜的DNA序列可以通过多种方法插入质粒或载体(在本文统称为“载体”)。通常,将DNA序列通过标准方法插入适宜限制性核酸内切酶位点。认为此类方法和相关的亚克隆方法处于本领域技术人员的知识范围内。

[0161] 本公开的异源核酸构建体可以包含本文构思的突变CDC或其片段的编码性序列,所述编码性序列(i) 处于分离状态;(ii) 与额外的编码性序列组合,如(但不限于)融合蛋白编码性序列或信号肽编码性序列,其中突变CDC编码性序列是优势的编码性序列;(iii) 与合适的宿主中有效表达编码性序列的非编码序列组合,如(但不限于)内含子和控制元件,如启动子元件和终止子元件或5'和/或3'非翻译区;和/或(iv) 处于其中突变CDC编码性序列是异源基因的载体或宿主环境中。

[0162] 适宜的载体一般配备编码选择标记的核酸序列、插入位点和合适的控制元件,如启动子序列和终止序列。载体可以包含与编码性序列有效连接的调节序列,所述调节序列例如包括非编码序列如内含子和控制元件,即,启动子元件和终止子元件或5'和/或3'非翻译区,在宿主细胞中(和/或在其中修饰的可溶性蛋白抗原编码性序列正常情况下不表达的

载体或宿主细胞环境下)有效表达编码性序列。大量合适的载体和启动子是本领域技术人员已知的,其中许多是市售的。

[0163] 示例性启动子包括组成型启动子和诱导型启动子,其例子包括CMV启动子、SV40早期启动子、RSV启动子、EF-1 α 启动子,含有tet-on或tet-off系统中tet反应元件(TRE)的启动子、 β 肌动蛋白启动子和可以通过添加某些金属盐上调的金属硫蛋白启动子。启动子序列是由用于表达目的的宿主细胞识别的DNA序列。它与编码突变多肽的DNA序列有效连接。

[0164] 除非另外说明,否则本公开组合和方法的实施利用处于本领域普通技术人员能力范围内的分子生物学、微生物学、重组DNA和免疫学常规技术。

实施例

[0165] 下文提供实施例。但是,本公开的实施方案在申请时不限于本文所述的具体实验、结果和实验室方法。反而,实施例如多种实施方案当中那样单纯提供并且意在起到示例作用,而非排他作用,并且将可以理解,本公开的教授内容的额外和不同实施方案将无疑向本领域技术人员自我提示;因此,这类其他实施方案认定为已经从本公开推断。

[0166] 实施例1

[0167] 胆固醇依赖性细胞溶素(CDC)是超过20种不同物种革兰氏阳性菌产生的成孔多肽毒素大家族¹。最初,细菌将这些毒素作为稳定的水溶性单体分泌。单体与膜结合并且发生特定顺序的结构性变化,这促进寡聚化和孔形成。如名称提示,对其成孔机制而言,CDC成孔机制绝对依赖于膜胆固醇。数十年的学说是,胆固醇是这些毒素的受体并且位于CDC的结构域4(D4)的保守十一肽(图2),对CDC与胆固醇相互作用重要²⁻⁴。但是,其他研究已经提示,十一肽不介导这些CDC与富含胆固醇的膜的初始结合^{5,6}。因此,在本研究工作之前仍不清这些CDC的介导其与胆固醇结合的结构性组分。

[0168] CDC机制对氧化敏感已经已知超过80年⁷,并且这种性状是最初给予这些毒素的名称“巯基激活的细胞溶素”的原因(参考文献8中综述)。这个巯基的氧化导致溶细胞活性显著丧失,经常>99%²。随后通过众多CDC的序列分析显示,具有敏感巯基的半胱氨酸位于保守的十一肽(ECTGLAWEWWR-SEQ ID NO:39)中,因为这是大部分测序的CDC中存在的唯一半胱氨酸。已经提出与这个巯基氧化相关的溶细胞丧失活因与富含胆固醇的膜的结合作用改变所致²,因此确立了膜结合和十一肽之间的推定性联系。十一肽的高度保守性质还提示高度保守的功能,其或许介导与膜胆固醇直接相互作用。

[0169] 胆固醇是CDC的受体的学说因发现中间链球菌溶素(ILY)(一种由中间链球菌(*Streptococcus intermedius*)分泌的CDC)而复杂化。与其他CDC相反,ILY具有人细胞特异性^{9,10},一种因其与人CD59(补体攻膜复合体的物种特异性抑制物^{11,12})而非富含胆固醇的膜¹³特异性结合的能力解释的特征。因此,目前存在至少两类CDC,与特定非固醇受体结合的ILY和与富含胆固醇的膜直接结合的PF0样CDC。然而,两个类型CDC的溶细胞机制均对膜胆固醇敏感并且在胆固醇基本上耗尽的膜上均无活性¹⁴。因此,这些研究提出一个谜:胆固醇以显著不同于PF0样CDC的方式促进ILY机制吗或存在胆固醇有助于两类CDC的统一分子基础吗?

[0170] Giddings等人¹⁴显示,hRBC膜胆固醇的耗尽阻断全部CDC的前孔至孔转化过程,还影响PF0样CDC与该膜的结合。Soltani等人¹⁵显示破坏ILY的L1-L3D4环(图2)插入膜也阻断

前孔转化成孔。因此,在ILY中两种不同的现象阻断前孔转化成孔:膜胆固醇耗尽¹⁴和破坏L1-L3环插入膜¹⁵。

[0171] 基于这些观察结果,对ILY和PF0的D4环和十一肽与膜的相互作用进行详细研究。这些研究的结果显示,在结构域4基部的L1-L3环是识别富含胆固醇的膜而不识别十一肽的主要结构。这些环与富含胆固醇的膜的相互作用介导PF0与富含胆固醇的膜的相互作用,而它们插入膜中还是PF0和ILY的前孔至孔转化过程必需的。因此,这些结果现在提供了CDC的胆固醇敏感性的结构基础并且统一解释了胆固醇对利用不同膜受体的ILY和PF0样CDC的影响。

[0172] 实施例1的材料和方法

[0173] 细菌菌株、质粒和化学品

[0174] 将ILY和PF0的基因如先前所述那样克隆入pTrcHisA (Invitrogen)^{14,16}。在天然ILY(半胱氨酸天然较少)或半胱氨酸较少的PF0(PF0^{C459A})背景下进行全部突变。天然PF0在残基459含有半胱氨酸,已经将所述半胱氨酸变成丙氨酸以产生半胱氨酸较少的PF0衍生物PF0^{C459A}。PF0和PF0^{C459A}均显示出相似的溶细胞活性¹⁶。全部化学品和酶均获自Sigma、VWR和Research Organics。全部荧光探针均获自Molecular Probes (Invitrogen)。

[0175] ILY及其衍生物的产生和纯化

[0176] 使用PCR QuikChange诱变(Stratagene),在天然ILY或PF0^{C459A}中产生多种氨基酸置换。ILY基因的突变形式的DNA序列由俄克拉荷马州医学研究基金DNA测序核心机构分析。来自大肠杆菌的重组ILY及其衍生物的表达和纯化如所述那样实施^{15,16}。将洗脱的蛋白质在4°C过夜透析入缓冲液(300mM NaCl,10mM MES,1mM EDTA,pH 6.5)。随后将蛋白质在-80°C储存于5mM DTT和10% (vol/vol) 无菌甘油中。

[0177] 用巯基特异性试剂化学修饰ILY和PF0及其衍生物

[0178] 用环境敏感性探针碘代乙酰氨基-N,N'-二甲基-N-(7-硝基-2-氧杂-1,3-二唑基)乙二胺(NBD)借助巯基修饰ILY的半胱氨酸衍生物。如先前述那样实施反应¹⁴。将修饰的蛋白质储存在10% (vol/vol) 无菌甘油中、在液氮中快速冷冻并储存在-80°C。按75%或更高的效率标记蛋白质。

[0179] 荧光测量

[0180] 全部荧光强度测量均使用SLM-8100光子计数荧光光谱仪如先前所述那样进行¹⁶。对于NBD测量,使用460-480nm激发波长和540nm发射波长,带通为4nm。对于每份样品,按分辨率1nm以积分时间1秒从500-600nm实施发射扫描。将含有10μg总毒素的样品与PBS[10mM Na₂HPO₄,2mM KH₂PO₄,137mM NaCl,3mM KCl(pH 7.5)]中的人红细胞(hRBC)血影膜(等同于303.25μg膜蛋白)在37°C温育5-10分钟,之后进行广谱测量。

[0181] 脂质体制备

[0182] 如所述那样制备按比率45:55mol%含有1-棕榈酰-2-油酰-sn-甘油-3-磷酸胆碱(POPC;Avanti Polar Lipids)和胆固醇的脂质体¹⁶。

[0183] HRBC血影膜制备

[0184] HRBC血影膜如先前所述那样制备。还如前述那样,使用Bradford法(Bio-Rad Protein Assay,Bio-Rad Laboratories,Inc.)定量膜蛋白含量^{14,16}。

[0185] 胆固醇消耗和补充

[0186] 用甲基- β -环糊精 (MBCD) 进行如先前所述那样进行胆固醇提取¹⁴。简而言之,将人 huRBC 血影膜与终浓度 20mM-40mM MBCD (每次使用,新鲜制得) 在 37°C 温育 2 小时。将膜通过反复离心 (在 4°C, 14,000 转/分钟持续 20 分钟) 洗涤三次并重悬于 PBS 中,以移除过量的 MBCD。血影膜最终悬浮于 PBS 中。使用胆固醇/胆固醇酯定量试剂盒 (Calbiochem, Billerica, MA), 测量胆固醇含量。一般,膜的胆固醇含量由这种方法减少 >90%。

[0187] 使用加载胆固醇的 MBCD 进行胆固醇补充。先前已经描述这种方法¹⁴。简而言之,向缓冲液 A (140mM NaCl, 5mM KCl, 5mM KH₂PO₄, 1mM MgSO₄, 10mM HEPES, 5mM 葡萄糖, pH 6.5) 添加新鲜制得的 MBCD 添加至终浓度 5mM。在 1:2 (vol/vol) 的氯仿:甲醇中制得 100mM 胆固醇母液。在玻璃容器中加热缓冲液 A+MBCD 加热至 80°C。一旦加热到 80°C, 添加悬浮的胆固醇至终浓度 4mM。通过超声处理 (4X 20 秒) 均化溶液。随后使用 0.22 μ m 滤器过滤溶液。将载有胆固醇的 MBCD 添加至沉淀的耗尽胆固醇的血影膜并在 37°C 温育 2 小时。如前通过重复离心洗涤膜并且最终重悬于 PBS 中。

[0188] 在 L1 SPR 传感芯片上固定脂质体

[0189] 使用 L1 传感芯片 (BIAcore, 乌普萨拉, 瑞典), 以 BIAcore 3000 系统测量表面等离子体共振 (SPR)。L1 传感芯片含有与疏水性残基共价结合并且已经常规用于固定脂质体的葡聚糖基质。为脂质体准备 L1 芯片时, 将 10 μ l 20mM CHAPS 以流量 10 μ l/分钟注射。脂质体 (0.5mM 脂质终浓度) 随后按相同流量注射 10 分钟。在注射后脂质体, 注射 50mM NaOH 持续 3 分钟以移除多层脂质。此后注射 0.1mg/ml BSA 以包被非特异性结合位点。全部注射均在 25°C 进行。通过反复注射 20mM CHAPS 和 50mM NaOH 再生 L1 芯片并剥离其脂质体, 直至达到原始 RU 读数。再生程序未导致传感芯片结合容量损耗。

[0190] SPR 分析

[0191] 脂质体和 PF0 衍生物之间相互作用的全部分析均在 HBS 中在 25°C 进行。将野生型 PF0 (50ng/ μ l) 和 PF0 天冬氨酸突变体 (50ng/ μ l) 在脂质体包被的芯片上以流量 30 μ l/分钟注射 4 分钟。

[0192] 实施例 1 的结果

[0193] 实验策略。ILY 不依赖于膜胆固醇与天然膜结合, 但是其机制仍对胆固醇敏感。不同于不与缺少胆固醇的膜结合的 PF0 样 CDC, ILY 的受体结合和寡聚化仍在胆固醇耗尽的膜上出现¹⁴。因此, ILY 用来首先鉴定负责其胆固醇依赖性的结构。一旦鉴定对膜胆固醇敏感的 ILY 结构, 则在 PF0 中研究破坏这些结构对其与富含胆固醇的脂质体膜结合的能力的影响。以这种方式, 可以确定是否 ILY 和 PF0 二者中的相同结构负责其胆固醇依赖性。

[0194] ILY 十一肽插入膜不需要胆固醇。采用 ILY 的先前研究已经显示, 十一肽必须插入膜中, 以便前孔形成¹⁵。因此, 确定其插入过程是否对膜胆固醇敏感。将半胱氨酸残基置换位于十一肽内部的 A1a-486 并借助其硫氢基用 NBD 标记。已经显示在天然 ILY 中这个残基插入膜¹⁵。在含有胆固醇的膜或胆固醇耗尽的膜不存在和存在情况下测量 ILY^{A486C-NBD} 中 NBD 的荧光强度。如图 3 中所示, 在 huRBC 血影膜存在下, 十一肽插入膜中, 如与针对可溶性状态的 ILY 观察到的情况相比荧光发射强度增加所示。当膜耗尽胆固醇时, 观察到相同的荧光发射增加。这些结果显示十一肽区域在 A1a-486 附近插入膜与膜胆固醇含量无关。

[0195] 环 L1、L2 和 L3 的插入需要胆固醇。在 D4 顶端的三个疏水性短环插入膜 (图 2) 协同发生并且是 CDC 单体在膜上锚定和恰当定向需要的^{15,17}。与十一肽的插入协同, 这些环的插入

是D3跨膜 β -发夹(TMHS)后续插入膜(其导致跨膜 β -桶孔形成)必需的¹⁵。胆固醇还是TMH插入和孔复合体形成需要的¹⁴。因此,膜胆固醇和L1-L3环插入膜是前孔转化成孔的前提^{14,15}。由于L1-L3环插入膜先于D3TMH插入,因此似乎合理的是,膜胆固醇的耗尽可能阻断L1-L3环的插入,这转而阻止D3TMH的插入并阻断前孔转变成孔。因此,假设L1-L3环插入膜需要胆固醇。

[0196] 为了检验这个假设,分别测量L1-L3环向天然huRBC血影膜和耗尽胆固醇的huRBC血影膜的膜插入过程。最近显示,位于环L1、L2和L3内部的ILY残基Leu-518、Ala-424和Ala-464分别插入膜中¹⁷。为了测量每个环的插入过程,将每种环中的残基突变成半胱氨酸(ILY^{A428C}、ILY^{A464C}、ILY^{L518C})¹⁵并且用NBD衍生化硫氢基。由于位于这些位点的NBD进入膜,其荧光发射强度显著增加^{15,17}。在可溶性单聚体毒素、与huRBC血影膜结合的毒素和与耗尽胆固醇的血影膜结合的毒素之间比较NBD的发射强度。

[0197] 与每个环插入天然hRBC血影的膜时所见的荧光发射强度增加形成鲜明对比,大约90%膜胆固醇的耗尽消除了全部三个环插入膜(图4,小图a-c)。向胆固醇耗尽的膜恢复胆固醇恢复了环插入膜的能力(图4,小图d-f)。因此,膜胆固醇是L1-L3环插入过程需要的,并且如先前显示,这个插入过程是前孔转化成孔必需的^{14,15}。

[0198] 天冬氨酸置换PF0的环L1-L3中的残基阻止其与富含胆固醇的膜结合。ILY的L1-L3环插入膜对天然膜中的胆固醇耗尽敏感,显示在PF0中这些相同的环可能介导其与富含胆固醇的膜直接结合。但是,不能在PF0中按照与随ILY所用相似的方式解决这个问题,因为胆固醇耗尽减少PF0与膜结合。因此,确定突变这些相同环对PF0与富含胆固醇的脂质体结合的影响。通过向PF0的环L1-L3中引入天冬氨酸完成这点,先前在ILY中显示所述引入阻止这些环插入膜¹⁵。在ILY中环L1-L3的插入是偶联的,并且向任何单一环残基Ala-428(L2)、Ala-464(L3)或Leu518(L1)引入天冬氨酸阻断它们插入膜。因此,预计如果将天冬氨酸置换PF0中的任一个类似残基Ala-401、Ala-437或Leu-491,它将破坏PF0与富含胆固醇的脂质体结合。

[0199] PF0中类似残基Ala-401(L2)、Ala-437(L3)和Leu-491(L1)的单独置换导致每种突变体的溶血活性丧失大于99.7%(数据未显示)。通过表面等离子体共振(SPR)测量PF0突变体与胆固醇-PC脂质体的结合。如图5a中所示,通过SPR检查时,这些突变显著减少与胆固醇-PC脂质体的结合。天冬氨酸置换Ala-401(L2)或Leu-491(L1)彻底地消除PF0与脂质体膜结合,并且因天冬氨酸置换Ala-437(L3)所致的结合作用比野生型低7%(图5b)。这个结果显示,D4L1-L3环对PF0样CDC与富含胆固醇的膜的相互作用至关重要。

[0200] 修饰PF0的Cys-459阻断十一肽色氨酸残基插入膜,但不阻断PF0的膜结合作用。长久以来认为PF0样CDC的保守十一肽参与它们与富含胆固醇的膜结合,主要原因是,据报道化学修饰十一肽的天然半胱氨酸(Cys-459)的硫氢基显著影响PF0与低细胞数目的羊RBC结合,但不影响与高细胞数目的羊RBC结合²。然而,其他已经证实其修饰似乎不影响其他CDC与细胞的结合^{5,6}。因此,通过SPR比较了天然PF0和借助十一肽的Cys-459的硫氢基修饰的PF0与胆固醇-PC脂质体结合的能力。

[0201] 用硫氢基特异性试剂N-乙基马来酰亚胺(NEM)对PF0十一肽Cys-459巯基的修饰降低溶血活性99%(数据未显示),这类似于其中化学修饰PF0和SLO的半胱氨酸硫氢基的其他报道^{2,18}。然而NEM修饰型毒素的结合速率和程度增加超过天然毒素,如通过SPR分析所测定

(图6A-B)。因此,化学修饰Cys-459并不破坏PF0与膜结合。

[0202] 如果修饰Cys-459不影响结合作用,则它提出以下问题:有效阻断PF0活性的这种修饰对PF0做了什么。自几乎90年前发现CDC以来,已经已知其溶细胞机制对氧化敏感。敏感残基的氧化最终与高度保守的十一肽半胱氨酸残基关联¹。进一步检查半胱氨酸修饰对PF0的结构性影响以确定其修饰是否阻止可能影响其活性的PF0结构变化。十一肽色氨酸464、466和467插入膜在构象偶联于D3TMH的插入。先前研究已经显示,D3TMH1残基处的突变增加其插入速率,还增加十一肽色氨酸残基插入膜的速率¹⁹。由于Cys-459靠近色氨酸残基,因此确定化学修饰半胱氨酸巯基是否阻断色氨酸残基插入膜。

[0203] 十一肽色氨酸残基的膜插入过程可以通过这些残基移入膜的非极性环境时其固有荧光强度增加来监测^{20,21}。在NEM修饰的PF0和天然PF0中测量这些色氨酸的插入过程(图6a和6b)。修饰Cys-459阻断十一肽色氨酸插入,但是不阻止它形成抗SDS低聚物,这类类似于天然PF0(数据未显示)。因此,这些数据显示,PF0结构的构象变化反映为十一肽色氨酸残基插入过程的丧失影响前孔低聚物后续转化成孔复合体。

[0204] 用肺炎球菌溶血素突变体Leu 460Asp免疫接种

[0205] 使用明矾(氢氧化铝)作为佐剂,在第0天和第14天用5 μ g肺炎球菌溶血素或肺炎球菌溶血素突变体皮下免疫接种CBA/CAHN-XID小鼠。在第21天,用单一稀释剂(无佐剂)中的蛋白质免疫接种小鼠。全部注射均以0.2ml体积给予。在第35天,用荚膜型19F菌株EF3030攻击小鼠。七天后,用二氧化碳气体令小鼠安乐死。将肺匀浆,并且通过镀覆在血琼脂平板上涂布匀浆的组织确定每只小鼠的肺中菌落形成单位(CFU)数。还将小鼠放血。血液中未观察到肺炎球菌,显示这是肺炎模型并且不是肺炎和败血症模型。结果显示,野生型和突变肺炎球菌溶血素均能够在小鼠局灶性肺炎模型中保护小鼠免遭肺炎。(图7)。

[0206] CDC的两个长期标志是其成孔机制依赖于膜胆固醇的存在和大部分CDC因十一肽半胱氨酸氧化而可逆性失活。本文的研究解析了两种现象的分子基础。不希望受理论约束,位于结构域4基部的L1-L3环插入膜似乎是ILY对膜胆固醇的存在敏感的主要事件。一旦胆固醇耗尽,则这些环不插入膜中,并且如先前显示,从hRBC膜提取胆固醇¹⁵阻止ILY的前孔转化成孔。这些结果表明两种效应也均因这些环不能插入胆固醇耗尽的膜中所致。这些数据进一步显示,PF0中保守半胱氨酸的氧化和假定地其他PF0样CDC阻断诱使PF0处于前孔状态的色氨酸残基插入膜,但不影响与富含胆固醇的脂质体结合。

[0207] 人细胞特异性毒素ILY的发现提出如下难题:如果胆固醇是ILY的受体,它怎样区分人细胞和动物细胞。ILY的人细胞特异性因发现人CD59(一种晚期阶段的物种特异性补体抑制物)是其受体而得到解释¹³。即使胆固醇不是ILY受体,其成孔机制仍对膜胆固醇敏感¹⁴,并且显示胆固醇是ILY中成孔机制的更晚阶段需要的;膜胆固醇的实质耗尽阻断前孔转化成孔。有趣地,还对SLO和PF0¹⁴(两种可以与富含胆固醇的膜直接结合的CDC)观察到这一点。尽管来自hRBC的膜胆固醇耗尽阻断PF0前孔转化成孔,它还减少PF0结合。因此,胆固醇是全部三种CDC的前孔转化成孔必需的,此外,它还有助于PF0样CDC结合膜。

[0208] 最近Soltani等人¹⁵显示ILY的L1-L3D4环插入膜是前孔转化成孔必需的。因此,胆固醇和L1-L3环插入膜均为ILY的前孔转化成孔必需的。不希望受理论约束,本文展示的数据表示,对这些观察结果的统一解释是这些环插入膜仅在富含胆固醇的膜中出现,并且这个插入过程是ILY和PF0样CDC的前孔转化成孔必需的。此外,这些环插入富含胆固醇的膜中

的能力还介导PF0及假定地PF0样CDC与富含胆固醇的膜表面初始结合。因此,这些数据显示在ILY和PF0样CDC中,L1-L3环必须插入膜中以便成功形成孔复合体。在ILY的情况下,结合作用首先由huCD59介导,接着L1-L3环插入富含胆固醇的膜中,而这两个事件即结合和插入在PF0中是完全相同的并且主要由L1-L3环介导。

[0209] 传统上已经认为,十一肽PF0样CDC有助于或直接介导富含胆固醇的膜的识别^{2,3,21}。本文的研究显示,L1-L3环是介导CDC和富含胆固醇的膜之间相互作用的主要结构。尽管用NEM化学修饰PF0的十一肽半胱氨酸降低其溶血活性超过99%,但是PF0与胆固醇-PC脂质体结合大体上不受损害。因此,与现有学说相反,PF0和其他PF0样CDC的相互作用主要由环L1-L3介导而不由十一肽介导。十一肽内部的突变可能影响L1-L3与富含胆固醇的膜相互作用。已经显示,突变ILY的十一肽Trp-491阻断L1-L3的插入¹⁵,并且改变的天然ILY十一肽结构明显阻止L1-L3与富含胆固醇的膜直接相互作用,因此允许它首先与huCD59结合。后一种观点得到以下事实强化:向ILY引入共有的十一肽结构时,它能够实现与非人类细胞结合²²。

[0210] 令人好奇的是,类似于PF0,ILY的L1-L3环为何不介导与富含胆固醇的膜结合。如上文提出,结构域4之间的主要差异似乎是高度保守的十一肽的一级结构。显然ILY已经丧失与富含胆固醇的膜直接结合的能力;否则,它将不显示借助huCD59介导的人细胞特异性。ILY和PF0的D4的晶体结构可以解释L1-L3环中介导这两种CDC与富含胆固醇的膜直接结合的这种差异。ILY的L1-L3残基(Leu-518、Ala-428和Ala-464)的位置和取向几乎与PF0中的类似残基(Leu-491、Ala-401和Ala-437)相同(图4b)。实际上,两种CDC的大部分D4结构几乎相同(rms偏差小于**0.6Å**,参考文献23),例外是结构域4顶部的十一肽环和 β -舌结构。较之PF0十一肽,ILY的十一肽环从D4基部向下进一步延伸**4-5Å**。因此,ILY十一肽可能在空间上阻碍ILY的L1-L3环与富含胆固醇的表面相互作用。可能仅与受体结合后,ILY十一肽结构才以如此方式改变,从而允许L1-L3环的插入。

[0211] 本公开揭示了严重影响十一肽半胱氨酸氧化对PF0和假定其他PF0样CDC的溶细胞机制所显示活性的结构性基础。最初,CDC因这个特征而称作巯基激活的细胞溶素,但是这种效应的分子基础未知。早期研究提出与RBC结合受影响,但与此同时与胆固醇结合不受影响,并且仍在细胞的表面上观察到无溶解性的寡聚体²。如本文所示,这种修饰阻止十一肽色氨酸插入并导致截获前孔的低聚体结构。尽管这种影响的精确结构性基础未知,但是先前的研究已经显示,形成跨膜 α -桶孔的结构域3TMH插入膜在构象上与结构域4十一肽色氨酸残基插入膜偶联¹⁹。因此,阻止这些色氨酸插入膜可以阻止结构域3TMH的插入,因此诱使PF0处于前孔状态。

[0212] 实施例2

[0213] 使用亲和柱纯化加His标签的PLY_{野生型}和PLY突变体(PLY_{L460D}和PLY_{L460D/G293S})。蛋白质平均产率是PLY_{野生型}:1mg/ml,PLY_{L460D}:1mg/ml,和PLY_{L460D/G293S}:4mg/ml。通过连续滴定毒素与人红细胞(RBC)温育,对纯化的蛋白质测试溶血活性。从非线性S型剂量-反应曲线计算每种蛋白质的EC₅₀(溶解50%RBC所需要的有效浓度)。测定每种突变体距野生型PLY的倍数变化(倍数变化=野生型EC₅₀/突变体EC₅₀)。据报道两种PLY突变体的活性低于野生型PLY某个倍数,因为在这些衍生物的最高蛋白质浓度,精确剂量-反应曲线所要求的100%RBC溶胞作

用不可实现。与野生型蛋白相比,突变体的溶血活性下降(活性低于PLY_{野生型}的倍数)对于PLY_{L460D}是>10,000倍并且对于PLY_{L460D/G293S}是>260,000倍。

[0214] 从计算熔化温度(50%蛋白质已经解折叠的温度)(T_m 摄氏度)中推断蛋白质的相对稳定性。使用Protein Thermal Shift Assay Dye试剂盒(Applied Biosystems),生成蛋白质熔化曲线。随着温度增加,蛋白质解折叠并且染料能够结合暴露的疏水性区域并发出荧光。PLY_{野生型}、PLY_{L460D}和PLY_{L460D/G293S}的 T_m 分别是 47.50 ± 0.2 、 47.69 ± 0.2 、 47.97 ± 0.2 。三种PLY蛋白报告的 T_m 无显著差异表明,向PLY_{野生型}引入所示的突变不影响这些蛋白质的稳定性。

[0215] 如上文所示,在一个非限制性实施方案中,本公开的肺炎球菌溶血素突变体是命名为PLY-L460D/G293S(PLY_{L460D/G293S})的双突变体,其中位置460的L用D置换并且位置293的G用S置换。G293S置换本身仅降低PLY突变体的溶血活性约50倍。L460D置换本身降低PLY突变体的活性约5000-10,000倍。但是在具有两种置换的PLY突变体中,活性较天然PLY毒素下降超过260,000倍。这是一种几何式下降,不仅仅是活性的“加合式”下降。

[0216] 不希望受理论约束,活性的急剧下降归因于PLY的二种必需功能阻断。首先,L460D置换阻断与胆固醇的结合,并且其次,G293S置换诱使PLY处于不能插入 β -桶孔的前孔状态(前孔定义为已经以环样结构低聚化,但是不能插入 β -桶孔的膜结合型单体)。因此,抑制作用为几何式:毒性低50X 5000-10,000>250,000倍折叠(溶血性较低),这符合显示其毒性较天然PLY低>260,000倍的测量。G293S置换还使L460D突变蛋白的单体结构稳定并且在纯化时增加产率。例如,从1.0升大肠杆菌培养物纯化时,PLY-L460D/G293S突变体的产率是大约52mg,而PLY-L460D突变体是3mg(多约17倍)。

[0217] 因此,在某些实施方案中,本公开涉及纯化或分离的突变肺炎球菌溶血素多肽,所述多肽包含与SEQ ID NO:1至少约90%相同的氨基酸序列并且在氨基酸位置458、459和460至少之一及氨基酸位置293和294至少之一处具有氨基酸置换。如与野生型肺炎球菌溶血素多肽相比,突变肺炎球菌溶血素多肽可以具有减少的溶血活性和减少的成孔活性。氨基酸序列可以与SEQ ID NO:1至少约91%相同、至少约92%相同、至少约93%相同、至少约94%相同、至少约95%相同、至少约96%相同、至少约97%相同、至少约98%相同、或至少约99%相同。突变肺炎球菌溶血素多肽可以在氨基酸位置293和458;293和459;293和460;294和458;294和459;或294和460包含氨基酸置换。突变肺炎球菌溶血素多肽可以在氨基酸位置293、458和459;293、459和460;或293、458和460包含氨基酸置换。突变肺炎球菌溶血素多肽可以在氨基酸位置294、458和459;294、459和460;或294、458和460包含氨基酸置换。突变肺炎球菌溶血素多肽可以在氨基酸位置293、294、458和459;293、294、459和460;或293、294、458和460包含氨基酸置换。突变肺炎球菌溶血素多肽可以在氨基酸位置293、294、458、459和460包含氨基酸置换。突变肺炎球菌溶血素多肽可以在氨基酸位置293包含丝氨酸或苏氨酸并在氨基酸位置460中包含天冬氨酸、谷氨酸或天冬酰胺。纯化的突变肺炎球菌溶血素多肽的氨基酸序列可以是SEQ ID NO:40。相对于野生型肺炎球菌溶血素或相对于在氨基酸位置293、294、458、459和460中仅一个氨基酸位置具有置换的突变肺炎球菌溶血素多肽,突变肺炎球菌溶血素多肽可以具有增加的产量。突变肺炎球菌溶血素多肽可以具有比野生型肺炎球菌溶血素多肽低约250,000倍的溶血活性。

[0218] 在另一个实施方案中,上文所述或否则构思的一种或多种突变肺炎球菌溶血素多肽可以置于可药用赋形剂中以形成免疫原性组合物。另一个实施方案包括疫苗,所述疫苗

包括免疫原性组合物并且可以任选地含有佐剂。又一个实施方案是编码本文所述或否则构思的任一突变肺炎球菌溶血素多肽的核酸序列；又一个实施方案涉及包括所述核酸序列的宿主细胞。又一个实施方案涉及一种在需要这种治疗的受试者中治疗、预防性防止或减少肺炎链球菌所致的病状、疾病或感染出现的方法；在这种的方法中，将治疗有效量的本文所述或否则构思的任何免疫原性组合物施用至受试者。

[0219] 在其他实施方案中，本公开涉及纯化或分离的突变链球菌溶素0多肽，所述多肽包含与SEQ ID NO:9至少90%相同的氨基酸序列并且在氨基酸位置561和562至少之一及氨基酸位置395和396至少之一处具有氨基酸置换。如与野生型链球菌溶素0多肽相比，突变链球菌溶素0多肽可以具有减少的溶血活性和减少的成孔活性。氨基酸序列可以与SEQ ID NO:9至少约91%相同、至少约92%相同、至少约93%相同、至少约94%相同、至少约95%相同、至少约96%相同、至少约97%相同、至少约98%相同、或至少约99%相同。突变链球菌溶素0多肽可以在氨基酸位置395和561；395和562；396和561；和/或396和562包含氨基酸置换。突变链球菌溶素0多肽可以在氨基酸位置395、561和562；396、561和562；395、396和561；395、396和562；和395、396、561和562包含氨基酸置换。相对于野生型链球菌溶素0或相对于在氨基酸位置395、396、561和562中仅一个氨基酸位置具有置换的突变链球菌溶素0多肽，突变链球菌溶素0多肽可以具有增加的产量。突变链球菌溶素0多肽可以具有比野生型链球菌溶素0多肽低约250,000倍的溶血活性。

[0220] 在另一个实施方案中，上文所述或否则构思的一种或多种突变链球菌溶素0多肽可以置于可药用赋形剂中以形成免疫原性组合物。另一个实施方案包括疫苗，所述疫苗包括免疫原性组合物并且可以任选地含有佐剂。又一个实施方案是编码本文所述或否则构思的任一突变链球菌溶素0多肽的核酸序列；又一个实施方案涉及包括所述核酸序列的宿主细胞。又一个实施方案涉及一种在需要这种治疗的受试者中治疗、预防性防止或减少化脓性链球菌所致的病状、疾病或感染出现的方法；在这种方法中，将治疗有效量的本文所述或否则构思的任何免疫原性组合物施用至受试者。

[0221] 虽然已经详细描述了本公开，但应当理解可以在本文所述的实施方案中进行各种变化、替换和改变而不脱离本公开的精神和范围。另外，本公开的范围不意在限于本说明书中描述的过程、物质组成、手段、方法和步骤的具体实施方案，尤其就本文描述或能够实现的具体氨基酸或核酸序列而言是这样。如本领域普通技术人员将从本发明的公开轻易地领会目前存在的或后续待开发的过程、物质组成、手段、方法、序列或步骤，它们基本上执行与本文所述的相应实施方案相同的功能或基本上实现与之相同的结果。因而，所述所附权利要求书意图将全部这类过程、物质组成、方法或方法步骤、氨基酸序列和核酸序列纳入其范围内。

[0222] 参考文献

[0223] 以下参考文献特别通过引用的方式并入本文，尤其就示例性方法细节或补充本文所述那些细节的其他细节而言通过引用的方式并入本文。

[0224] 1. Alouf, J.E., Billington, S.J. & Jost, B.H. Repertoire and general features of the cholesterol-dependent cytolysins in Bacterial Toxins: A Comprehensive Sourcebook (eds. Alouf, J.E. & Popoff, M.R.) 643-658 (Academic Press, London, 2005).

[0225] 2. Iwamoto, M., Ohno-Iwashita, Y. & Ando, S. Role of the essential thiol

group in the thiol-activated cytolysin from *Clostridium perfringens*. *Eur J Biochem* 167,425-430 (1987) .

[0226] 3. Sekino-Suzuki, N., Nakamura, M., Mitsui, K. I. & Ohno-Iwashita, Y. Contribution of individual tryptophan residues to the structure and activity of theta-toxin (perfringolysin O), a cholesterol-binding cytolysin. *Eur J Biochem* 241,941-947 (1996) .

[0227] 4. Jacobs, T. et al. The conserved undecapeptide shared by thiol-activated cytolysins is involved in membrane binding. *FEBS Lett.* 459,463-466 (1999) .

[0228] 5. Vazquez-Boland, J. A., Dominguez, L., Rodriguez-Ferri, E. F., Fernandez-Garayzabal, J. F. & Suarez, G. Preliminary evidence that different domains are involved in cytolytic activity and receptor (cholesterol) binding in listeriolysin O, the *Listeria monocytogenes* thiol-activated toxin. *FEMS Microbiol Lett* 53,95-9 (1989) .

[0229] 6. Saunders, F. K., Mitchell, T. J., Walker, J. A., Andrew, P. W. & Boulnois, G. J. Pneumolysin, the thiol-activated toxin of *Streptococcus pneumoniae*, does not require a thiol group for in vitro activity. *Infect. Immun.* 57,2547-2552 (1989) .

[0230] 7. Neill, J. M. & Fleming, W. L. Studies on the oxidation and reduction of immunological substances: II The hematoxin of the Welch bacillus. *J. Exp. Med.* 44, 215-226 (1926) .

[0231] 8. Alouf, J. E. & Geoffrey, C. Structure activity relationships in the sulfhydryl-activated toxins. in *Bacterial Protein Toxins* (eds. Alouf, J. E., Fehrenbach, F. J., Freer, J. H. & Jeljaszewicz, J.) 165-171 (Academic Press, London, 1984) .

[0232] 9. Nagamune, H. et al. Intermedilysin. A cytolytic toxin specific for human cells of a *Streptococcus intermedius* isolated from human liver abscess. *Adv Exp Med Biol* 418,773-775 (1997) .

[0233] 10. Nagamune, H. et al. Intermedilysin, a novel cytotoxin specific for human cells secreted by *Streptococcus intermedius* UNS46 isolated from a human liver abscess. *Infect Immun* 64,3093-3100 (1996) .

[0234] 11. Rollins, S. A., Zhao, J., Ninomiya, H. & Sims, P. J. Inhibition of homologous complement by CD59 is mediated by a species-selective recognition conferred through binding to C8 within C5b-8 or C9 within C5b-9. *J Immunol* 146,2345-51 (1991) .

[0235] 12. Rollins, S. A. & Sims, P. J. The complement-inhibitory activity of CD59 resides in its capacity to block incorporation of C9 into membrane C5b-9. *J Immunol* 144,3478-83 (1990) .

[0236] 13. Giddings, K. S., Zhao, J., Sims, P. J. & Tweten, R. K. Human CD59 is a

receptor for the cholesterol-dependent cytolysin intermedilysin. *Nat Struct Mol Biol* 12,1173-1178 (2004) .

[0237] 14. Giddings, K. S., Johnson, A. E. & Tweten, R. K. Redefining cholesterol's role in the mechanism of the cholesterol-dependent cytolysins. *Proc Natl Acad Sci U S A* 100,11315-11320 (2003) .

[0238] 15. Soltani, C. E., Hotze, E. M., Johnson, A. E. & Tweten, R. K. Specific protein-membrane contacts are required for prepore and pore assembly by a cholesterol-dependent cytolysin. *J. Biol. Chem. Paper in press* 282(21), 15709-15716, April 5, 2007.

[0239] 16. Shepard, L. A. et al. Identification of a membrane-spanning domain of the thiol-activated pore-forming toxin *Clostridium perfringens* perfringolysin O: an α -helical to β -sheet transition identified by fluorescence spectroscopy. *Biochemistry* 37,14563-14574 (1998) .

[0240] 17. Ramachandran, R., Heuck, A. P., Tweten, R. K. & Johnson, A. E. Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin. *Nat Struct Biol* 9,823-7 (2002) .

[0241] 18. Harris, J. R., Adrian, M., Bhakdi, S. & Palmer, M. Cholesterol-Streptolysin O Interaction: An EM Study of Wild-Type and Mutant Streptolysin O. *J Struct Biol* 121,343-55 (1998) .

[0242] 19. Heuck, A. P., Hotze, E., Tweten, R. K. & Johnson, A. E. Mechanism of membrane insertion of a multimeric β -barrel protein: Perfringolysin O creates a pore using ordered and coupled conformational changes. *Molec. Cell* 6,1233-1242 (2000) .

[0243] 20. Heuck, A. P., Tweten, R. K. & Johnson, A. E. Assembly and topography of the prepore complex in cholesterol-dependent cytolysins. *J Biol Chem* 278,31218-31225 (2003) .

[0244] 21. Nakamura, M., Sekino, N., Iwamoto, M. & Ohno-Iwashita, Y. Interaction of theta-toxin (perfringolysin O), a cholesterol-binding cytolysin, with liposomal membranes: change in the aromatic side chains upon binding and insertion. *Biochemistry* 34,6513-6520 (1995) .

[0245] 22. Nagamune, H. et al. The human-specific action of intermedilysin, a homolog of streptolysin o, is dictated by domain 4 of the protein. *Microbiol Immunol* 48,677-92 (2004) .

[0246] 23. Polekhina, G., Giddings, K. S., Tweten, R. K. & Parker, M. W. Insights into the action of the superfamily of cholesterol-dependent cytolysins from studies of intermedilysin. *Proc Natl Acad Sci* 102,600-605 (2005) .

[0247] 24. Rossjohn, J., Feil, S. C., McKinstry, W. J., Tweten, R. K. & Parker, M. W. Structure of a cholesterol-binding thiol-activated cytolysin and a model of its membrane form. *Cell* 89,685-692 (1997) .

[0248] 25.Humphrey,W.,Dalke,A.&Schulten,K.VMD:visual molecular dynamics.J Mol Graph 14,33-8,27-8(1996) .

[0249] 26.匿名:Prevention of Pneumococcol Disease:Recommendations of the Advisory Committee on Immunization Practices(ACIP).Morbidity and Mortality Weekly Report-Recommendations and Reports:46:1-24(April 4,1997) .

[0001] 序列表
 [0002] <110> 俄克拉何马大学董事会
 [0003] <120> 肺炎球菌溶血素突变体及其使用方法
 [0004] <130> 5835.161wo
 [0005] <150> 62/082,848
 [0006] <151> 2014-11-21
 [0007] <160> 40
 [0008] <170> PatentIn 版本 3.5
 [0009] <210> 1
 [0010] <211> 471
 [0011] <212> PRT
 [0012] <213> Streptococcus pneumoniae
 [0013] <400> 1
 [0014] Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala Met Asn Tyr Asp
 [0015] 1 5 10 15
 [0016] Lys Lys Lys Leu Leu Thr His Gln Gly Glu Ser Ile Glu Asn Arg Phe
 [0017] 20 25 30
 [0018] Ile Lys Glu Gly Asn Gln Leu Pro Asp Glu Phe Val Val Ile Glu Arg
 [0019] 35 40 45
 [0020] Lys Lys Arg Ser Leu Ser Thr Asn Thr Ser Asp Ile Ser Val Thr Ala
 [0021] 50 55 60
 [0022] Thr Asn Asp Ser Arg Leu Tyr Pro Gly Ala Leu Leu Val Val Asp Glu
 [0023] 65 70 75 80
 [0024] Thr Leu Leu Glu Asn Asn Pro Thr Leu Leu Ala Val Asp Arg Ala Pro
 [0025] 85 90 95
 [0026] Met Thr Tyr Ser Ile Asp Leu Pro Gly Leu Ala Ser Ser Asp Ser Phe
 [0027] 100 105 110
 [0028] Leu Gln Val Glu Asp Pro Ser Asn Ser Ser Val Arg Gly Ala Val Asn
 [0029] 115 120 125
 [0030] Asp Leu Leu Ala Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val
 [0031] 130 135 140
 [0032] Pro Ala Arg Met Gln Tyr Glu Lys Ile Thr Ala His Ser Met Glu Gln
 [0033] 145 150 155 160
 [0034] Leu Lys Val Lys Phe Gly Ser Asp Phe Glu Lys Thr Gly Asn Ser Leu
 [0035] 165 170 175
 [0036] Asp Ile Asp Phe Asn Ser Val His Ser Gly Glu Lys Gln Ile Gln Ile
 [0037] 180 185 190
 [0038] Val Asn Phe Lys Gln Ile Tyr Tyr Thr Val Ser Val Asp Ala Val Lys
 [0039] 195 200 205
 [0040] Asn Pro Gly Asp Val Phe Gln Asp Thr Val Thr Val Glu Asp Leu Lys
 [0041] 210 215 220

[0042]	Gln Arg Gly Ile Ser Ala Glu Arg Pro Leu Val Tyr Ile Ser Ser Val
[0043]	225 230 235 240
[0044]	Ala Tyr Gly Arg Gln Val Tyr Leu Lys Leu Glu Thr Thr Ser Lys Ser
[0045]	245 250 255
[0046]	Asp Glu Val Glu Ala Ala Phe Glu Ala Leu Ile Lys Gly Val Lys Val
[0047]	260 265 270
[0048]	Ala Pro Gln Thr Glu Trp Lys Gln Ile Leu Asp Asn Thr Glu Val Lys
[0049]	275 280 285
[0050]	Ala Val Ile Leu Gly Gly Asp Pro Ser Ser Gly Ala Arg Val Val Thr
[0051]	290 295 300
[0052]	Gly Lys Val Asp Met Val Glu Asp Leu Ile Gln Glu Gly Ser Arg Phe
[0053]	305 310 315 320
[0054]	Thr Ala Asp His Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
[0055]	325 330 335
[0056]	Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr Asp Tyr Val Glu
[0057]	340 345 350
[0058]	Thr Lys Val Thr Ala Tyr Arg Asn Gly Asp Leu Leu Leu Asp His Ser
[0059]	355 360 365
[0060]	Gly Ala Tyr Val Ala Gln Tyr Tyr Ile Thr Trp Asp Glu Leu Ser Tyr
[0061]	370 375 380
[0062]	Asp His Gln Gly Lys Glu Val Leu Thr Pro Lys Ala Trp Asp Arg Asn
[0063]	385 390 395 400
[0064]	Gly Gln Asp Leu Thr Ala His Phe Thr Thr Ser Ile Pro Leu Lys Gly
[0065]	405 410 415
[0066]	Asn Val Arg Asn Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
[0067]	420 425 430
[0068]	Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Thr Asp Leu Pro Leu Val
[0069]	435 440 445
[0070]	Arg Lys Arg Thr Ile Ser Ile Trp Gly Thr Thr Leu Tyr Pro Gln Val
[0071]	450 455 460
[0072]	Glu Asp Lys Val Glu Asn Asp
[0073]	465 470
[0074]	<210> 2
[0075]	<211> 509
[0076]	<212> PRT
[0077]	<213> <i>Bacillus cereus</i>
[0078]	<400> 2
[0079]	Met Asn Ile Lys Lys Asn Thr Lys Arg Arg Lys Phe Leu Ala Cys Leu
[0080]	1 5 10 15
[0081]	Leu Val Ser Leu Cys Thr Ile Asn Tyr Ser Ser Ile Ser Phe Ala Glu
[0082]	20 25 30
[0083]	Thr Gln Ala Ser Asn Ala Thr Asp Val Thr Lys Asn Ala Ser Gly Ile

[0084]	35	40	45
[0085]	Asp Thr Gly Ile Ala Asn Leu Lys Tyr Asn Asn Gln Glu Val Leu Ala		
[0086]	50	55	60
[0087]	Val Asn Gly Asp Lys Val Glu Ser Phe Val Pro Lys Glu Ser Ile Asn		
[0088]	65	70	75
[0089]	Ser Asn Gly Lys Phe Val Val Val Glu Arg Glu Lys Lys Ser Leu Thr		
[0090]	85	90	95
[0091]	Thr Ser Pro Val Asp Ile Ser Ile Ile Asp Ser Val Val Asn Arg Thr		
[0092]	100	105	110
[0093]	Tyr Pro Gly Ala Val Gln Leu Ala Asn Lys Ala Phe Ala Asp Asn Gln		
[0094]	115	120	125
[0095]	Pro Ser Leu Leu Val Ala Lys Arg Lys Pro Leu Asn Ile Ser Ile Asp		
[0096]	130	135	140
[0097]	Leu Pro Gly Met Arg Lys Glu Asn Thr Ile Thr Val Gln Asn Pro Thr		
[0098]	145	150	155
[0099]	Tyr Gly Asn Val Ala Gly Ala Val Asp Asp Leu Val Ser Thr Trp Asn		
[0100]	165	170	175
[0101]	Glu Lys Tyr Ser Thr Thr His Thr Leu Pro Ala Arg Met Gln Tyr Thr		
[0102]	180	185	190
[0103]	Glu Ser Met Val Tyr Ser Lys Ser Gln Ile Ala Ser Ala Leu Asn Val		
[0104]	195	200	205
[0105]	Asn Ala Lys Tyr Leu Asp Asn Ser Leu Asn Ile Asp Phe Asn Ala Val		
[0106]	210	215	220
[0107]	Ala Asn Gly Glu Lys Lys Val Met Val Ala Ala Tyr Lys Gln Ile Phe		
[0108]	225	230	235
[0109]	Tyr Thr Val Ser Ala Glu Leu Pro Asn Asn Pro Ser Asp Leu Phe Asp		
[0110]	245	250	255
[0111]	Asn Ser Val Thr Phe Asp Glu Leu Thr Arg Lys Gly Val Ser Asn Ser		
[0112]	260	265	270
[0113]	Ala Pro Pro Val Met Val Ser Asn Val Ala Tyr Gly Arg Thr Ile Tyr		
[0114]	275	280	285
[0115]	Val Lys Leu Glu Thr Thr Ser Lys Ser Lys Asp Val Gln Ala Ala Phe		
[0116]	290	295	300
[0117]	Lys Ala Leu Leu Lys Asn Asn Ser Val Glu Thr Ser Gly Gln Tyr Lys		
[0118]	305	310	315
[0119]	Asp Ile Phe Glu Glu Ser Thr Phe Thr Ala Val Val Leu Gly Gly Asp		
[0120]	325	330	335
[0121]	Ala Lys Glu His Asn Lys Val Val Thr Lys Asp Phe Asn Glu Ile Arg		
[0122]	340	345	350
[0123]	Asn Ile Ile Lys Asp Asn Ala Glu Leu Ser Leu Lys Asn Pro Ala Tyr		
[0124]	355	360	365
[0125]	Pro Ile Ser Tyr Thr Ser Thr Phe Leu Lys Asp Asn Ser Thr Ala Ala		

[0126]	370	375	380
[0127]	Val His Asn Asn Thr Asp Tyr Ile Glu Thr Thr Thr Thr Glu Tyr Ser		
[0128]	385	390	395
[0129]	Ser Ala Lys Met Thr Leu Asp His Tyr Gly Ala Tyr Val Ala Gln Phe		
[0130]		405	410
[0131]	Asp Val Ser Trp Asp Glu Phe Thr Phe Asp Gln Lys Gly Asn Glu Val		
[0132]		420	425
[0133]	Leu Thr His Lys Thr Trp Asp Gly Ser Gly Lys Asp Lys Thr Ala His		
[0134]		435	440
[0135]	Tyr Ser Thr Val Ile Pro Leu Pro Pro Asn Ser Lys Asn Ile Lys Ile		
[0136]		450	455
[0137]	Val Ala Arg Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Thr Ile		
[0138]		465	470
[0139]	Ile Asn Glu Gln Asn Val Pro Leu Thr Asn Glu Ile Lys Val Ser Ile		
[0140]		485	490
[0141]	Gly Gly Thr Thr Leu Tyr Pro Thr Ala Ser Ile Ser His		
[0142]		500	505
[0143]	<210> 3		
[0144]	<211> 512		
[0145]	<212> PRT		
[0146]	<213> Bacillus anthracis		
[0147]	<400> 3		
[0148]	Met Ile Phe Leu Asn Ile Lys Lys Asn Thr Lys Arg Arg Lys Phe Leu		
[0149]	1	5	10
[0150]	Ala Cys Leu Leu Val Ser Leu Cys Thr Ile His Tyr Ser Ser Ile Ser		
[0151]		20	25
[0152]	Phe Ala Glu Thr Gln Ala Gly Asn Ala Thr Gly Ala Ile Lys Asn Ala		
[0153]		35	40
[0154]	Ser Asp Ile Asn Thr Gly Ile Ala Asn Leu Lys Tyr Asp Ser Arg Asp		
[0155]		50	55
[0156]	Ile Leu Ala Val Asn Gly Asp Lys Val Glu Ser Phe Ile Pro Lys Glu		
[0157]		65	70
[0158]	Ser Ile Asn Ser Asn Gly Lys Phe Val Val Val Glu Arg Glu Lys Lys		
[0159]		85	90
[0160]	Ser Leu Thr Thr Ser Pro Val Asp Ile Leu Ile Ile Asp Ser Val Val		
[0161]		100	105
[0162]	Asn Arg Thr Tyr Pro Gly Ala Val Gln Leu Ala Asn Lys Ala Phe Ala		
[0163]		115	120
[0164]	Asp Asn Gln Pro Ser Leu Leu Val Ala Lys Arg Lys Pro Leu Asn Ile		
[0165]		130	135
[0166]	Ser Ile Asp Leu Pro Gly Met Arg Lys Glu Asn Thr Ile Thr Val Gln		
[0167]		145	150
			155
			160

[0168]	Asn Pro Thr Tyr Gly Asn Val Ala Gly Ala Val Asp Asp Leu Val Ser
[0169]	165 170 175
[0170]	Thr Trp Asn Glu Lys Tyr Ser Thr Thr His Thr Leu Pro Ala Arg Met
[0171]	180 185 190
[0172]	Gln Tyr Thr Glu Ser Met Val Tyr Ser Lys Ser Gln Ile Ala Ser Ala
[0173]	195 200 205
[0174]	Leu Asn Val Asn Ala Lys Tyr Leu Asp Asn Ser Leu Asn Ile Asp Phe
[0175]	210 215 220
[0176]	Asn Ala Val Ala Asn Gly Glu Lys Lys Val Met Val Ala Ala Tyr Lys
[0177]	225 230 235 240
[0178]	Gln Ile Phe Tyr Thr Val Ser Ala Glu Leu Pro Asn Asn Pro Ser Asp
[0179]	245 250 255
[0180]	Leu Phe Asp Asn Ser Val Thr Phe Asp Glu Leu Thr Arg Lys Gly Val
[0181]	260 265 270
[0182]	Ser Asn Ser Ala Pro Pro Val Met Val Ser Asn Val Ala Tyr Gly Arg
[0183]	275 280 285
[0184]	Thr Val Tyr Val Lys Leu Glu Thr Thr Ser Lys Ser Lys Asp Val Gln
[0185]	290 295 300
[0186]	Ala Ala Phe Lys Ala Leu Leu Lys Asn Asn Ser Val Glu Thr Ser Gly
[0187]	305 310 315 320
[0188]	Gln Tyr Lys Asp Ile Phe Glu Glu Ser Thr Phe Thr Ala Val Val Leu
[0189]	325 330 335
[0190]	Gly Gly Asp Ala Lys Glu His Asn Lys Val Val Thr Lys Asp Phe Asn
[0191]	340 345 350
[0192]	Glu Ile Arg Asn Ile Ile Lys Asp Asn Ala Glu Leu Ser Phe Lys Asn
[0193]	355 360 365
[0194]	Pro Ala Tyr Pro Ile Ser Tyr Thr Ser Thr Phe Leu Lys Asp Asn Ala
[0195]	370 375 380
[0196]	Thr Ala Ala Val His Asn Asn Thr Asp Tyr Ile Glu Thr Thr Thr Thr
[0197]	385 390 395 400
[0198]	Glu Tyr Ser Ser Ala Lys Met Thr Leu Asp His Tyr Gly Ala Tyr Val
[0199]	405 410 415
[0200]	Ala Gln Phe Asp Val Ser Trp Asp Glu Phe Thr Phe Asp Gln Asn Gly
[0201]	420 425 430
[0202]	Lys Glu Val Leu Thr His Lys Thr Trp Glu Gly Ser Gly Lys Asp Lys
[0203]	435 440 445
[0204]	Thr Ala His Tyr Ser Thr Val Ile Pro Leu Pro Pro Asn Ser Lys Asn
[0205]	450 455 460
[0206]	Ile Lys Ile Val Ala Arg Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp
[0207]	465 470 475 480
[0208]	Arg Thr Ile Ile Asn Glu Gln Asn Val Pro Leu Thr Asn Glu Ile Lys
[0209]	485 490 495

[0210]	Val Ser Ile Gly Gly Thr Thr Leu Tyr Pro Thr Ala Thr Ile Ser His
[0211]	500 505 510
[0212]	<210> 4
[0213]	<211> 509
[0214]	<212> PRT
[0215]	<213> Bacillus thuringiensis
[0216]	<400> 4
[0217]	Met Asn Ile Lys Lys Asn Thr Lys Arg Arg Lys Phe Leu Ala Cys Leu
[0218]	1 5 10 15
[0219]	Leu Val Ser Leu Cys Thr Ile Asn Tyr Ser Ser Ile Ser Phe Ala Glu
[0220]	20 25 30
[0221]	Thr Gln Ala Ser Asn Ala Thr Asp Val Thr Lys Asn Ala Ser Gly Ile
[0222]	35 40 45
[0223]	Asp Thr Gly Ile Ala Asn Leu Lys Tyr Asn Ile Gln Glu Val Leu Ala
[0224]	50 55 60
[0225]	Val Asn Gly Asp Lys Val Glu Ser Phe Val Pro Lys Glu Ser Ile Asn
[0226]	65 70 75 80
[0227]	Ser Asn Gly Lys Phe Val Val Val Glu Arg Glu Lys Lys Ser Leu Thr
[0228]	85 90 95
[0229]	Thr Ser Pro Val Asp Ile Ser Ile Ile Asp Ser Val Val Asn Arg Thr
[0230]	100 105 110
[0231]	Tyr Pro Gly Ala Val Gln Leu Ala Asn Lys Ala Phe Ala Asp Asn Gln
[0232]	115 120 125
[0233]	Pro Ser Leu Leu Val Ala Lys Arg Lys Pro Leu Asn Ile Ser Ile Asp
[0234]	130 135 140
[0235]	Leu Pro Gly Met Arg Lys Glu Asn Thr Ile Thr Val Gln Asn Pro Thr
[0236]	145 150 155 160
[0237]	Tyr Gly Asn Val Ala Gly Ala Val Asp Asp Leu Val Ser Thr Trp Asn
[0238]	165 170 175
[0239]	Glu Lys Tyr Ser Thr Thr His Thr Leu Pro Ala Arg Met Gln Tyr Thr
[0240]	180 185 190
[0241]	Glu Ser Met Val Tyr Ser Lys Ser Gln Ile Ala Ser Ala Leu Asn Val
[0242]	195 200 205
[0243]	Asn Ala Lys Tyr Leu Asp Asn Ser Leu Asn Ile Gly Phe Asn Ala Val
[0244]	210 215 220
[0245]	Ala Asn Gly Glu Lys Lys Val Met Val Ala Ala Tyr Lys Gln Ile Phe
[0246]	225 230 235 240
[0247]	Tyr Thr Val Ser Ala Glu Leu Pro Asn Asn Pro Ser Asp Leu Phe Asp
[0248]	245 250 255
[0249]	Asn Ser Val Thr Phe Asp Glu Leu Thr Arg Lys Gly Val Asn Asn Ser
[0250]	260 265 270
[0251]	Ala Pro Pro Val Met Val Ser Asn Val Ala Tyr Gly Arg Thr Ile Tyr

[0252]	275	280	285
[0253]	Val Lys Leu Glu Thr Thr Ser Lys Ser Lys Asp Val Gln Ala Ala Phe		
[0254]	290	295	300
[0255]	Lys Ala Leu Leu Lys Asn Asn Ser Val Glu Thr Ser Gly Gln Tyr Lys		
[0256]	305	310	315 320
[0257]	Asp Ile Phe Glu Glu Ser Thr Phe Thr Ala Val Val Leu Gly Gly Asp		
[0258]		325	330 335
[0259]	Ala Lys Glu His Asn Lys Val Val Thr Lys Asp Phe Asn Glu Ile Arg		
[0260]		340	345 350
[0261]	Asn Ile Ile Lys Asp Asn Ala Glu Leu Ser Leu Lys Asn Pro Ala Tyr		
[0262]		355	360 365
[0263]	Pro Ile Ser Tyr Thr Ser Thr Phe Leu Lys Asp Asn Ala Thr Ala Ala		
[0264]		370	375 380
[0265]	Val His Asn Asn Thr Asp Tyr Ile Glu Thr Thr Thr Thr Glu Tyr Ser		
[0266]		385	390 395 400
[0267]	Ser Ala Lys Met Thr Leu Asp His Tyr Gly Ala Tyr Val Ala Gln Phe		
[0268]		405	410 415
[0269]	Asp Val Ser Trp Asp Glu Phe Thr Phe Asp Gln Lys Gly Asn Glu Val		
[0270]		420	425 430
[0271]	Leu Thr His Lys Thr Trp Asp Gly Ser Gly Lys Asp Lys Thr Ala His		
[0272]		435	440 445
[0273]	Tyr Ser Thr Val Ile Pro Leu Pro Pro Asn Ser Lys Asn Ile Lys Ile		
[0274]		450	455 460
[0275]	Val Ala Arg Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Thr Ile		
[0276]		465	470 475 480
[0277]	Ile Asn Glu Gln Asn Val Pro Leu Thr Asn Glu Ile Lys Val Ser Ile		
[0278]		485	490 495
[0279]	Gly Gly Thr Thr Leu Tyr Pro Thr Ala Ser Ile Ser His		
[0280]		500	505
[0281]	<210> 5		
[0282]	<211> 500		
[0283]	<212> PRT		
[0284]	<213> Clostridium perfringens		
[0285]	<400> 5		
[0286]	Met Ile Arg Phe Lys Lys Thr Lys Leu Ile Ala Ser Ile Ala Met Ala		
[0287]	1	5	10 15
[0288]	Leu Cys Leu Phe Ser Gln Pro Val Ile Ser Phe Ser Lys Asp Ile Thr		
[0289]		20	25 30
[0290]	Asp Lys Asn Gln Ser Ile Asp Ser Gly Ile Ser Ser Leu Ser Tyr Asn		
[0291]		35	40 45
[0292]	Arg Asn Glu Val Leu Ala Ser Asn Gly Asp Lys Ile Glu Ser Phe Val		
[0293]		50	55 60

[0294]	Pro Lys Glu Gly Lys Lys Ala Gly Asn Lys Phe Ile Val Val Glu Arg
[0295]	65 70 75 80
[0296]	Gln Lys Arg Ser Leu Thr Thr Ser Pro Val Asp Ile Ser Ile Ile Asp
[0297]	85 90 95
[0298]	Ser Val Asn Asp Arg Thr Tyr Pro Gly Ala Leu Gln Leu Ala Asp Lys
[0299]	100 105 110
[0300]	Ala Phe Val Glu Asn Arg Pro Thr Ile Leu Met Val Lys Arg Lys Pro
[0301]	115 120 125
[0302]	Ile Asn Ile Asn Ile Asp Leu Pro Gly Leu Lys Gly Glu Asn Ser Ile
[0303]	130 135 140
[0304]	Lys Val Asp Asp Pro Thr Tyr Gly Lys Val Ser Gly Ala Ile Asp Glu
[0305]	145 150 155 160
[0306]	Leu Val Ser Lys Trp Asn Glu Lys Tyr Ser Ser Thr His Thr Leu Pro
[0307]	165 170 175
[0308]	Ala Arg Thr Gln Tyr Ser Glu Ser Met Val Tyr Ser Lys Ser Gln Ile
[0309]	180 185 190
[0310]	Ser Ser Ala Leu Asn Val Asn Ala Lys Val Leu Glu Asn Ser Leu Gly
[0311]	195 200 205
[0312]	Val Asp Phe Asn Ala Val Ala Asn Asn Glu Lys Lys Val Met Ile Leu
[0313]	210 215 220
[0314]	Ala Tyr Lys Gln Ile Phe Tyr Thr Val Ser Ala Asp Leu Pro Lys Asn
[0315]	225 230 235 240
[0316]	Pro Ser Asp Leu Phe Asp Asp Ser Val Thr Phe Asn Asp Leu Lys Gln
[0317]	245 250 255
[0318]	Lys Gly Val Ser Asn Glu Ala Pro Pro Leu Met Val Ser Asn Val Ala
[0319]	260 265 270
[0320]	Tyr Gly Arg Thr Ile Tyr Val Lys Leu Glu Thr Thr Ser Ser Ser Lys
[0321]	275 280 285
[0322]	Asp Val Gln Ala Ala Phe Lys Ala Leu Ile Lys Asn Thr Asp Ile Lys
[0323]	290 295 300
[0324]	Asn Ser Gln Gln Tyr Lys Asp Ile Tyr Glu Asn Ser Ser Phe Thr Ala
[0325]	305 310 315 320
[0326]	Val Val Leu Gly Gly Asp Ala Gln Glu His Asn Lys Val Val Thr Lys
[0327]	325 330 335
[0328]	Asp Phe Asp Glu Ile Arg Lys Val Ile Lys Asp Asn Ala Thr Phe Ser
[0329]	340 345 350
[0330]	Thr Lys Asn Pro Ala Tyr Pro Ile Ser Tyr Thr Ser Val Phe Leu Lys
[0331]	355 360 365
[0332]	Asp Asn Ser Val Ala Ala Val His Asn Lys Thr Asp Tyr Ile Glu Thr
[0333]	370 375 380
[0334]	Thr Ser Thr Glu Tyr Ser Lys Gly Lys Ile Asn Leu Asp His Ser Gly
[0335]	385 390 395 400

[0336]	Ala Tyr Val Ala Gln Phe Glu Val Ala Trp Asp Glu Val Ser Tyr Asp
[0337]	405 410 415
[0338]	Lys Glu Gly Asn Glu Val Leu Thr His Lys Thr Trp Asp Gly Asn Tyr
[0339]	420 425 430
[0340]	Gln Asp Lys Thr Ala His Tyr Ser Thr Val Ile Pro Leu Glu Ala Asn
[0341]	435 440 445
[0342]	Ala Arg Asn Ile Arg Ile Lys Ala Arg Glu Cys Thr Gly Leu Ala Trp
[0343]	450 455 460
[0344]	Glu Trp Trp Arg Asp Val Ile Ser Glu Tyr Asp Val Pro Leu Thr Asn
[0345]	465 470 475 480
[0346]	Asn Ile Asn Val Ser Ile Trp Gly Thr Thr Leu Tyr Pro Gly Ser Ser
[0347]	485 490 495
[0348]	Ile Thr Tyr Asn
[0349]	500
[0350]	<210> 6
[0351]	<211> 501
[0352]	<212> PRT
[0353]	<213> Bacillus alvei
[0354]	<400> 6
[0355]	Met Lys Lys Lys Ser Asn His Leu Lys Gly Arg Lys Val Leu Val Ser
[0356]	1 5 10 15
[0357]	Leu Leu Val Ser Leu Gln Val Phe Ala Phe Ala Ser Ile Ser Ser Ala
[0358]	20 25 30
[0359]	Ala Pro Thr Glu Pro Asn Asp Ile Asp Met Gly Ile Ala Gly Leu Asn
[0360]	35 40 45
[0361]	Tyr Asn Arg Asn Glu Val Leu Ala Ile Gln Gly Asp Gln Ile Ser Ser
[0362]	50 55 60
[0363]	Phe Val Pro Lys Glu Gly Ile Gln Ser Asn Gly Lys Phe Ile Val Val
[0364]	65 70 75 80
[0365]	Glu Arg Asp Lys Lys Ser Leu Thr Thr Ser Pro Val Asp Ile Ser Ile
[0366]	85 90 95
[0367]	Val Asp Ser Ile Thr Asn Arg Thr Tyr Pro Gly Ala Ile Gln Leu Ala
[0368]	100 105 110
[0369]	Asn Lys Asp Phe Ala Asp Asn Gln Pro Ser Leu Val Met Ala Ala Arg
[0370]	115 120 125
[0371]	Lys Pro Leu Asp Ile Ser Ile Asp Leu Pro Gly Leu Lys Asn Glu Asn
[0372]	130 135 140
[0373]	Thr Ile Ser Val Gln Asn Pro Asn Tyr Gly Thr Val Ser Ser Ala Ile
[0374]	145 150 155 160
[0375]	Asp Gln Leu Val Ser Thr Trp Gly Glu Lys Tyr Ser Ser Thr His Thr
[0376]	165 170 175
[0377]	Leu Pro Ala Arg Leu Gln Tyr Ala Glu Ser Met Val Tyr Ser Gln Asn

[0378]	180	185	190
[0379]	Gln Ile Ser Ser Ala Leu Asn Val Asn Ala Lys Val Leu Asn Gly Thr		
[0380]	195	200	205
[0381]	Leu Gly Ile Asp Phe Asn Ala Val Ala Asn Gly Glu Lys Lys Val Met		
[0382]	210	215	220
[0383]	Val Ala Ala Tyr Lys Gln Ile Phe Tyr Thr Val Ser Ala Gly Leu Pro		
[0384]	225	230	235 240
[0385]	Asn Asn Pro Ser Asp Leu Phe Asp Asp Ser Val Thr Phe Ala Glu Leu		
[0386]	245	250	255
[0387]	Ala Arg Lys Gly Val Ser Asn Glu Ala Pro Pro Leu Met Val Ser Asn		
[0388]	260	265	270
[0389]	Val Ala Tyr Gly Arg Thr Ile Tyr Val Lys Leu Glu Thr Thr Ser Lys		
[0390]	275	280	285
[0391]	Ser Asn Asp Val Gln Thr Ala Phe Lys Leu Leu Leu Asn Asn Pro Ser		
[0392]	290	295	300
[0393]	Ile Gln Ala Ser Gly Gln Tyr Lys Asp Ile Tyr Glu Asn Ser Ser Phe		
[0394]	305	310	315 320
[0395]	Thr Ala Val Val Leu Gly Gly Asp Ala Gln Thr His Asn Gln Val Val		
[0396]	325	330	335
[0397]	Thr Lys Asp Phe Asn Val Ile Gln Ser Val Ile Lys Asp Asn Ala Gln		
[0398]	340	345	350
[0399]	Phe Ser Ser Lys Asn Pro Ala Tyr Pro Ile Ser Tyr Thr Ser Val Phe		
[0400]	355	360	365
[0401]	Leu Lys Asp Asn Ser Ile Ala Ala Val His Asn Asn Thr Glu Tyr Ile		
[0402]	370	375	380
[0403]	Glu Thr Lys Thr Thr Glu Tyr Ser Lys Gly Lys Ile Lys Leu Asp His		
[0404]	385	390	395 400
[0405]	Ser Gly Ala Tyr Val Ala Gln Phe Glu Val Tyr Trp Asp Glu Phe Ser		
[0406]	405	410	415
[0407]	Tyr Asp Ala Asp Gly Gln Glu Ile Val Thr Arg Lys Ser Trp Asp Gly		
[0408]	420	425	430
[0409]	Asn Trp Arg Asp Arg Ser Ala His Phe Ser Thr Glu Ile Pro Leu Pro		
[0410]	435	440	445
[0411]	Pro Asn Ala Lys Asn Ile Arg Ile Phe Ala Arg Glu Cys Thr Gly Leu		
[0412]	450	455	460
[0413]	Ala Trp Glu Trp Trp Arg Thr Val Val Asp Glu Tyr Asn Val Pro Leu		
[0414]	465	470	475 480
[0415]	Ala Ser Asp Ile Asn Val Ser Ile Trp Gly Thr Thr Leu Tyr Pro Lys		
[0416]	485	490	495
[0417]	Ser Ser Ile Thr His		
[0418]	500		
[0419]	<210> 7		

[0420]	<211>	574
[0421]	<212>	PRT
[0422]	<213>	Streptococcus canis
[0423]	<400>	7
[0424]	Met Lys Asp Met Ser Asn Lys Lys Ile Phe Lys Lys Tyr Ser Arg Val	
[0425]	1	5 10 15
[0426]	Ala Gly Leu Leu Thr Ala Ala Leu Ile Val Gly Asn Leu Val Thr Ala	
[0427]		20 25 30
[0428]	Asn Ala Asp Ser Asn Lys Gln Asn Thr Ala Asn Thr Glu Thr Thr Thr	
[0429]		35 40 45
[0430]	Thr Asn Glu Gln Pro Lys Pro Glu Ser Ser Glu Leu Thr Thr Glu Lys	
[0431]		50 55 60
[0432]	Ala Gly Gln Lys Met Asp Asp Met Leu Asn Ser Asn Asp Met Ile Lys	
[0433]		65 70 75 80
[0434]	Leu Ala Pro Lys Glu Met Pro Leu Glu Ser Ala Glu Lys Glu Glu Lys	
[0435]		85 90 95
[0436]	Lys Ser Glu Asp Asn Lys Lys Ser Glu Glu Asp His Thr Glu Glu Ile	
[0437]		100 105 110
[0438]	Asn Asp Lys Ile Tyr Ser Leu Asn Tyr Asn Glu Leu Glu Val Leu Ala	
[0439]		115 120 125
[0440]	Lys Asn Gly Glu Thr Ile Glu Asn Phe Val Pro Lys Glu Gly Val Lys	
[0441]		130 135 140
[0442]	Lys Ala Asp Lys Phe Ile Val Ile Glu Arg Lys Lys Lys Asn Ile Asn	
[0443]		145 150 155 160
[0444]	Thr Thr Pro Val Asp Ile Ser Ile Ile Asp Ser Val Thr Asp Arg Thr	
[0445]		165 170 175
[0446]	Tyr Pro Ala Ala Leu Gln Leu Ala Asn Lys Gly Phe Thr Glu Asn Lys	
[0447]		180 185 190
[0448]	Pro Asp Ala Val Val Thr Lys Arg Asn Pro Gln Lys Ile His Ile Asp	
[0449]		195 200 205
[0450]	Leu Pro Gly Met Gly Asp Lys Ala Thr Val Glu Val Asn Asp Pro Thr	
[0451]		210 215 220
[0452]	Tyr Ala Asn Val Ser Thr Ala Ile Asp Asn Leu Val Asn Gln Trp His	
[0453]		225 230 235 240
[0454]	Asp Asn Tyr Ser Gly Gly Asn Thr Leu Pro Ala Arg Thr Gln Tyr Thr	
[0455]		245 250 255
[0456]	Glu Ser Met Val Tyr Ser Lys Ser Gln Ile Glu Ala Ala Leu Asn Val	
[0457]		260 265 270
[0458]	Asn Ser Lys Ile Leu Asp Gly Thr Leu Gly Ile Asp Phe Lys Ser Ile	
[0459]		275 280 285
[0460]	Ser Lys Gly Glu Lys Lys Val Met Ile Ala Ala Tyr Lys Gln Ile Phe	
[0461]		290 295 300

[0504]	20	25	30
[0505]	Ser Asn Lys Gln Asn Thr Ala Asn Thr Glu Thr Thr Thr Thr Asn Glu		
[0506]	35	40	45
[0507]	Gln Pro Lys Pro Glu Ser Ser Glu Leu Thr Thr Glu Lys Ala Gly Gln		
[0508]	50	55	60
[0509]	Lys Met Asp Asp Met Leu Asn Ser Asn Asp Met Ile Lys Leu Ala Pro		
[0510]	65	70	75
[0511]	Lys Glu Met Pro Leu Glu Ser Ala Glu Lys Glu Glu Lys Lys Ser Glu		
[0512]	85	90	95
[0513]	Asp Asn Lys Lys Ser Glu Glu Asp His Thr Glu Glu Ile Asn Asp Lys		
[0514]	100	105	110
[0515]	Ile Tyr Ser Leu Asn Tyr Asn Glu Leu Glu Val Leu Ala Lys Asn Gly		
[0516]	115	120	125
[0517]	Glu Thr Ile Glu Asn Phe Val Pro Lys Glu Gly Val Lys Lys Ala Asp		
[0518]	130	135	140
[0519]	Lys Phe Ile Val Ile Glu Arg Lys Lys Lys Asn Ile Asn Thr Thr Pro		
[0520]	145	150	155
[0521]	Val Asp Ile Ser Ile Ile Asp Ser Val Thr Asp Arg Thr Tyr Pro Ala		
[0522]	165	170	175
[0523]	Ala Leu Gln Leu Ala Asn Lys Gly Phe Thr Glu Asn Lys Pro Asp Ala		
[0524]	180	185	190
[0525]	Val Val Thr Lys Arg Asn Pro Gln Lys Ile His Ile Asp Leu Pro Gly		
[0526]	195	200	205
[0527]	Met Gly Asp Lys Ala Thr Val Glu Val Asn Asp Pro Thr Tyr Ala Asn		
[0528]	210	215	220
[0529]	Val Ser Thr Ala Ile Asp Asn Leu Val Asn Gln Trp His Asp Asn Tyr		
[0530]	225	230	235
[0531]	Ser Gly Gly Asn Thr Leu Pro Ala Arg Thr Gln Tyr Thr Glu Ser Met		
[0532]	245	250	255
[0533]	Val Tyr Ser Lys Ser Gln Ile Glu Ala Ala Leu Asn Val Asn Ser Lys		
[0534]	260	265	270
[0535]	Ile Leu Asp Gly Thr Leu Gly Ile Asp Phe Lys Ser Ile Ser Lys Gly		
[0536]	275	280	285
[0537]	Glu Lys Lys Val Met Ile Ala Ala Tyr Lys Gln Ile Phe Tyr Thr Val		
[0538]	290	295	300
[0539]	Ser Ala Asn Leu Pro Asn Asn Pro Ala Asp Val Phe Asp Lys Ser Val		
[0540]	305	310	315
[0541]	Thr Phe Lys Glu Leu Gln Arg Lys Gly Val Ser Asn Glu Ala Pro Pro		
[0542]	325	330	335
[0543]	Leu Phe Val Ser Asn Val Ala Tyr Gly Arg Thr Val Phe Val Lys Leu		
[0544]	340	345	350
[0545]	Glu Thr Ser Ser Lys Ser Asn Asp Val Glu Ala Ala Phe Ser Ala Ala		

[0546]	355	360	365
[0547]	Leu Lys Gly Thr Asp Val Lys Thr Asn Gly Lys Tyr Ser Asp Ile Leu		
[0548]	370	375	380
[0549]	Glu Asn Ser Ser Phe Thr Ala Val Val Leu Gly Gly Asp Ala Ala Glu		
[0550]	385	390	395
[0551]	His Asn Lys Val Val Thr Lys Asp Phe Asp Val Ile Arg Asn Val Ile		
[0552]	405	410	415
[0553]	Lys Asp Asn Ala Thr Phe Ser Arg Lys Asn Pro Ala Tyr Pro Ile Ser		
[0554]	420	425	430
[0555]	Tyr Thr Ser Val Phe Leu Lys Asn Asn Lys Ile Ala Gly Val Asn Asn		
[0556]	435	440	445
[0557]	Arg Ser Glu Tyr Val Glu Thr Thr Ser Thr Glu Tyr Thr Ser Gly Lys		
[0558]	450	455	460
[0559]	Ile Asn Leu Ser His Gln Gly Ala Tyr Val Ala Gln Tyr Glu Ile Leu		
[0560]	465	470	475
[0561]	Trp Asp Glu Ile Asn Tyr Asp Asp Lys Gly Lys Glu Val Ile Thr Lys		
[0562]	485	490	495
[0563]	Arg Arg Trp Asp Asn Asn Trp Tyr Ser Lys Thr Ser Pro Phe Ser Thr		
[0564]	500	505	510
[0565]	Val Ile Pro Leu Gly Ala Asn Ser Arg Asn Ile Arg Ile Met Ala Arg		
[0566]	515	520	525
[0567]	Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Lys Val Ile Asp Glu		
[0568]	530	535	540
[0569]	Arg Asp Val Lys Leu Ser Lys Glu Ile Asn Val Asn Ile Ser Gly Ser		
[0570]	545	550	555
[0571]	Thr Leu Ser Pro Tyr Gly Ser Ile Thr Tyr Lys		
[0572]	565	570	
[0573]	<210> 9		
[0574]	<211> 571		
[0575]	<212> PRT		
[0576]	<213> Streptococcus pyogenes		
[0577]	<400> 9		
[0578]	Met Ser Asn Lys Lys Thr Phe Lys Lys Tyr Ser Arg Val Ala Gly Leu		
[0579]	1	5	10
[0580]	Leu Thr Ala Ala Leu Ile Ile Gly Asn Leu Val Thr Ala Asn Ala Glu		
[0581]	20	25	30
[0582]	Ser Asn Lys Gln Asn Thr Ala Ser Thr Glu Thr Thr Thr Ser Glu		
[0583]	35	40	45
[0584]	Gln Pro Lys Pro Glu Ser Ser Glu Leu Thr Ile Glu Lys Ala Gly Gln		
[0585]	50	55	60
[0586]	Lys Met Asp Asp Met Leu Asn Ser Asn Asp Met Ile Lys Leu Ala Pro		
[0587]	65	70	75
			80

[0588]	Lys Glu Met Pro Leu Glu Ser Ala Glu Lys Glu Glu Lys Lys Ser Glu
[0589]	85 90 95
[0590]	Asp Lys Lys Lys Ser Glu Glu Asp His Thr Glu Glu Ile Asn Asp Lys
[0591]	100 105 110
[0592]	Ile Tyr Ser Leu Asn Tyr Asn Glu Leu Glu Val Leu Ala Lys Asn Gly
[0593]	115 120 125
[0594]	Glu Thr Ile Glu Asn Phe Val Pro Lys Glu Gly Val Lys Lys Ala Asp
[0595]	130 135 140
[0596]	Lys Phe Ile Val Ile Glu Arg Lys Lys Lys Asn Ile Asn Thr Thr Pro
[0597]	145 150 155 160
[0598]	Val Asp Ile Ser Ile Ile Asp Ser Val Thr Asp Arg Thr Tyr Pro Ala
[0599]	165 170 175
[0600]	Ala Leu Gln Leu Ala Asn Lys Gly Phe Thr Glu Asn Lys Pro Asp Ala
[0601]	180 185 190
[0602]	Val Val Thr Lys Arg Asn Pro Gln Lys Ile His Ile Asp Leu Pro Gly
[0603]	195 200 205
[0604]	Met Gly Asp Lys Ala Thr Val Glu Val Asn Asp Pro Thr Tyr Ala Asn
[0605]	210 215 220
[0606]	Val Ser Thr Ala Ile Asp Asn Leu Val Asn Gln Trp His Asp Asn Tyr
[0607]	225 230 235 240
[0608]	Ser Gly Gly Asn Thr Leu Pro Ala Arg Thr Gln Tyr Thr Glu Ser Met
[0609]	245 250 255
[0610]	Val Tyr Ser Lys Ser Gln Ile Glu Ala Ala Leu Asn Val Asn Ser Lys
[0611]	260 265 270
[0612]	Ile Leu Asp Gly Thr Leu Gly Ile Asp Phe Lys Ser Ile Ser Lys Gly
[0613]	275 280 285
[0614]	Glu Lys Lys Val Met Ile Ala Ala Tyr Lys Gln Ile Phe Tyr Thr Val
[0615]	290 295 300
[0616]	Ser Ala Asn Leu Pro Asn Asn Pro Ala Asp Val Phe Asp Lys Ser Val
[0617]	305 310 315 320
[0618]	Thr Phe Lys Asp Leu Gln Arg Lys Gly Val Ser Asn Glu Ala Pro Pro
[0619]	325 330 335
[0620]	Leu Phe Val Ser Asn Val Ala Tyr Gly Arg Thr Val Phe Val Lys Leu
[0621]	340 345 350
[0622]	Glu Thr Ser Ser Lys Ser Asn Asp Val Glu Ala Ala Phe Ser Ala Ala
[0623]	355 360 365
[0624]	Leu Lys Gly Thr Asp Val Lys Thr Asn Gly Lys Tyr Ser Asp Ile Leu
[0625]	370 375 380
[0626]	Glu Asn Ser Ser Phe Thr Ala Val Val Leu Gly Gly Asp Ala Ala Glu
[0627]	385 390 395 400
[0628]	His Asn Lys Val Val Thr Lys Asp Phe Asp Val Ile Arg Asn Val Ile
[0629]	405 410 415

[0630]	Lys Asp Asn Ala Thr Phe Ser Arg Lys Asn Pro Ala Tyr Pro Ile Ser
[0631]	420 425 430
[0632]	Tyr Thr Ser Val Phe Leu Lys Asn Asn Lys Ile Ala Gly Val Asn Asn
[0633]	435 440 445
[0634]	Arg Thr Glu Tyr Val Glu Thr Thr Ser Thr Glu Tyr Thr Ser Gly Lys
[0635]	450 455 460
[0636]	Ile Asn Leu Ser His Gln Gly Ala Tyr Val Ala Gln Tyr Glu Ile Leu
[0637]	465 470 475 480
[0638]	Trp Asp Glu Ile Asn Tyr Asp Asp Lys Gly Lys Glu Val Ile Thr Lys
[0639]	485 490 495
[0640]	Arg Arg Trp Asp Asn Asn Trp Tyr Ser Lys Thr Ser Pro Phe Ser Thr
[0641]	500 505 510
[0642]	Val Ile Pro Leu Gly Ala Asn Ser Arg Asn Ile Arg Ile Met Ala Arg
[0643]	515 520 525
[0644]	Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Lys Val Ile Asp Glu
[0645]	530 535 540
[0646]	Arg Asp Val Lys Leu Ser Lys Glu Ile Asn Val Asn Ile Ser Gly Ser
[0647]	545 550 555 560
[0648]	Thr Leu Ser Pro Tyr Gly Ser Ile Thr Tyr Lys
[0649]	565 570
[0650]	<210> 10
[0651]	<211> 514
[0652]	<212> PRT
[0653]	<213> Clostridium novyi
[0654]	<400> 10
[0655]	Met Lys Lys Ser Leu Lys Thr Ile Ile Arg Ser Ile Ser Phe Leu Ser
[0656]	1 5 10 15
[0657]	Ile Leu Thr Leu Thr Cys Ser Cys Asn Phe Ile Thr Ser Thr Gln Lys
[0658]	20 25 30
[0659]	Asn Val Ser Leu Leu Ser Gly Pro Asn Lys Val Ile Lys Pro Lys Lys
[0660]	35 40 45
[0661]	Thr Lys Ser Leu Asp Asp Arg Ile Tyr Gly Leu Lys Tyr Asp Pro Asn
[0662]	50 55 60
[0663]	Lys Ile Leu Ser Phe Asn Gly Glu Lys Val Glu Asn Phe Val Pro Asn
[0664]	65 70 75 80
[0665]	Glu Gly Phe Ser Thr Pro Asp Lys Tyr Ile Val Ile Lys Arg Glu Lys
[0666]	85 90 95
[0667]	Lys Ser Ile Ser Asp Ser Thr Ala Asp Ile Ala Val Ile Asp Ser Met
[0668]	100 105 110
[0669]	Asn Asp Lys Thr Tyr Pro Gly Ala Ile Gln Leu Ala Asn Arg Asn Leu
[0670]	115 120 125
[0671]	Ile Glu Asn Lys Pro Asn Ile Val Ser Cys Glu Arg Lys Pro Ile Thr

[0672]	130	135	140
[0673]	Ile Ser Ile Asp Leu Pro Gly Met Gly Glu Glu Gly Lys Thr Thr Ile		
[0674]	145	150	155
[0675]	Thr Ser Pro Thr Tyr Ser Ser Val Lys Ala Gly Ile Asp Ser Leu Leu		
[0676]		165	170
[0677]	Asn Lys Trp Asn Ser His Tyr Ser Ser Ile Tyr Ser Ile Pro Thr Arg		
[0678]		180	185
[0679]	Phe Ser Tyr Ser Asp Ser Met Val Tyr Ser Lys Ser Gln Leu Ser Ala		
[0680]		195	200
[0681]	Lys Leu Gly Cys Asn Phe Lys Ala Leu Asn Lys Ala Leu Asp Ile Asp		
[0682]		210	215
[0683]	Phe Asp Ser Ile Tyr Lys Gly Gln Lys Lys Val Met Leu Leu Ala Tyr		
[0684]		225	230
[0685]	Lys Gln Ile Phe Tyr Thr Val Asn Val Asp Ala Pro Asn His Pro Ser		
[0686]		245	250
[0687]	Asp Phe Phe Gly Asp Lys Val Thr Phe Asn Asp Leu Ala Lys Lys Gly		
[0688]		260	265
[0689]	Val Asn Ser Lys Asn Pro Pro Val Tyr Val Ser Ser Val Ser Tyr Gly		
[0690]		275	280
[0691]	Arg Thr Ile Tyr Val Lys Leu Glu Thr Thr Ser Lys Ser Ala Asn Val		
[0692]		290	295
[0693]	Lys Ala Ala Phe Lys Ala Leu Ile Glu Asn Gln Asn Ile Ser Ser Asn		
[0694]		305	310
[0695]	Ser Glu Tyr Lys Asn Ile Leu Asn Gln Ser Ser Phe Thr Ala Thr Val		
[0696]		325	330
[0697]	Leu Gly Gly Gly Ala Lys Glu His Asn Lys Val Ile Thr Lys Asn Phe		
[0698]		340	345
[0699]	Asp Glu Ile Arg Asn Ile Ile Thr Asn Asn Ser Glu Tyr Ser Pro Arg		
[0700]		355	360
[0701]	Asn Pro Gly Tyr Pro Ile Ala Tyr Thr Thr Ser Phe Leu Lys Asp Asn		
[0702]		370	375
[0703]	Ser Val Ala Thr Val Asn Asn Lys Thr Asp Tyr Ile Glu Thr Thr Ser		
[0704]		385	390
[0705]	Thr Glu Tyr Thr Asn Gly Lys Ile Thr Leu Asp His Arg Gly Ala Tyr		
[0706]		405	410
[0707]	Val Ala Lys Phe Asn Ile Thr Trp Asp Glu Val Ser Tyr Asp Lys Asn		
[0708]		420	425
[0709]	Gly Lys Glu Ile Val Glu His Lys Ser Trp Glu Gly Asn Asp Phe Gly		
[0710]		435	440
[0711]	Arg Thr Ala His Phe Asn Thr Glu Leu Tyr Leu Lys Gly Asn Ala Arg		
[0712]		450	455
[0713]	Asn Ile Cys Ile Lys Ala Lys Glu Cys Thr Gly Leu Ala Trp Glu Trp		

[0714]	465	470	475	480
[0715]	Trp Arg Thr Ile Ile Asp Asp Lys Asn Val Pro Leu Val Lys Asn Arg			
[0716]		485	490	495
[0717]	Lys Val Tyr Ile Trp Gly Thr Thr Leu Tyr Pro Arg Thr Leu Thr Glu			
[0718]		500	505	510
[0719]	Ile Glu			
[0720]	<210> 11			
[0721]	<211> 527			
[0722]	<212> PRT			
[0723]	<213> Clostridium tetani			
[0724]	<400> 11			
[0725]	Met Asn Lys Asn Val Leu Lys Phe Val Ser Arg Ser Leu Leu Ile Phe			
[0726]	1	5	10	15
[0727]	Ser Met Thr Gly Leu Ile Ser Asn Tyr Asn Ser Ser Asn Val Leu Ala			
[0728]		20	25	30
[0729]	Lys Gly Asn Val Glu Glu His Ser Leu Ile Asn Asn Gly Gln Val Val			
[0730]		35	40	45
[0731]	Thr Ser Asn Thr Lys Cys Asn Leu Ala Lys Asp Asn Ser Ser Asp Ile			
[0732]		50	55	60
[0733]	Asp Lys Asn Ile Tyr Gly Leu Ser Tyr Asp Pro Arg Lys Ile Leu Ser			
[0734]		65	70	80
[0735]	Tyr Asn Gly Glu Gln Val Glu Asn Phe Val Pro Ala Glu Gly Phe Glu			
[0736]		85	90	95
[0737]	Asn Pro Asp Lys Phe Ile Val Val Lys Arg Glu Lys Lys Ser Ile Ser			
[0738]		100	105	110
[0739]	Asp Ser Thr Ala Asp Ile Ser Ile Ile Asp Ser Ile Asn Asp Arg Thr			
[0740]		115	120	125
[0741]	Tyr Pro Gly Ala Ile Gln Leu Ala Asn Arg Asn Leu Met Glu Asn Lys			
[0742]		130	135	140
[0743]	Pro Asp Ile Ile Ser Cys Glu Arg Lys Pro Ile Thr Ile Ser Val Asp			
[0744]		145	150	160
[0745]	Leu Pro Gly Met Ala Glu Asp Gly Lys Lys Val Val Asn Ser Pro Thr			
[0746]		165	170	175
[0747]	Tyr Ser Ser Val Asn Ser Ala Ile Asn Ser Ile Leu Asp Thr Trp Asn			
[0748]		180	185	190
[0749]	Ser Lys Tyr Ser Ser Lys Tyr Thr Ile Pro Thr Arg Met Ser Tyr Ser			
[0750]		195	200	205
[0751]	Asp Thr Met Val Tyr Ser Gln Ser Gln Leu Ser Ala Ala Val Gly Cys			
[0752]		210	215	220
[0753]	Asn Phe Lys Ala Leu Asn Lys Ala Leu Asn Ile Asp Phe Asp Ser Ile			
[0754]		225	230	240
[0755]	Phe Lys Gly Glu Lys Lys Val Met Leu Leu Ala Tyr Lys Gln Ile Phe			

[0756]		245		250		255
[0757]	Tyr Thr Val Ser Val Asp Pro Pro Asn Arg Pro Ser Asp Leu Phe Gly					
[0758]		260		265		270
[0759]	Asp Ser Val Thr Phe Asp Glu Leu Ala Leu Lys Gly Ile Asn Asn Asn					
[0760]		275		280		285
[0761]	Asn Pro Pro Ala Tyr Val Ser Asn Val Ala Tyr Gly Arg Thr Ile Tyr					
[0762]		290		295		300
[0763]	Val Lys Leu Glu Thr Thr Ser Lys Ser Ser His Val Lys Ala Ala Phe					
[0764]	305		310		315	320
[0765]	Lys Ala Leu Ile Asn Asn Gln Asp Ile Ser Ser Asn Ala Glu Tyr Lys					
[0766]		325		330		335
[0767]	Asp Ile Leu Asn Gln Ser Ser Phe Thr Ala Thr Val Leu Gly Gly Gly					
[0768]		340		345		350
[0769]	Ala Gln Glu His Asn Lys Ile Ile Thr Lys Asp Phe Asp Glu Ile Arg					
[0770]		355		360		365
[0771]	Asn Ile Ile Lys Asn Asn Ser Val Tyr Ser Pro Gln Asn Pro Gly Tyr					
[0772]		370		375		380
[0773]	Pro Ile Ser Tyr Thr Thr Thr Phe Leu Lys Asp Asn Ser Ile Ala Ser					
[0774]	385		390		395	400
[0775]	Val Asn Asn Lys Thr Glu Tyr Ile Glu Thr Thr Ala Thr Glu Tyr Thr					
[0776]		405		410		415
[0777]	Asn Gly Lys Ile Val Leu Asp His Ser Gly Ala Tyr Val Ala Gln Phe					
[0778]		420		425		430
[0779]	Gln Val Thr Trp Asp Glu Val Ser Tyr Asp Glu Lys Gly Asn Glu Ile					
[0780]		435		440		445
[0781]	Val Glu His Lys Ala Trp Glu Gly Asn Asn Arg Asp Arg Thr Ala His					
[0782]		450		455		460
[0783]	Phe Asn Thr Glu Ile Tyr Leu Lys Gly Asn Ala Arg Asn Ile Ser Val					
[0784]	465		470		475	480
[0785]	Lys Ile Arg Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Thr Ile					
[0786]		485		490		495
[0787]	Val Asp Val Lys Asn Ile Pro Leu Ala Lys Glu Arg Thr Phe Tyr Ile					
[0788]		500		505		510
[0789]	Trp Gly Thr Thr Leu Tyr Pro Lys Thr Ser Ile Glu Thr Lys Met					
[0790]		515		520		525
[0791]	<210> 12					
[0792]	<211> 528					
[0793]	<212> PRT					
[0794]	<213> <i>Listeria ivanovii</i>					
[0795]	<400> 12					
[0796]	Met Lys Lys Ile Met Leu Leu Leu Met Thr Leu Leu Leu Val Ser Leu					
[0797]	1	5		10		15

[0798]	Pro Leu Ala Gln Glu Ala Gln Ala Asp Ala Ser Val Tyr Ser Tyr Gln
[0799]	20 25 30
[0800]	Gly Ile Ile Ser His Met Ala Pro Pro Ala Ser Pro Pro Ala Lys Pro
[0801]	35 40 45
[0802]	Lys Thr Pro Val Glu Lys Lys Asn Ala Ala Gln Ile Asp Gln Tyr Ile
[0803]	50 55 60
[0804]	Gln Gly Leu Asp Tyr Asp Lys Asn Asn Ile Leu Val Tyr Asp Gly Glu
[0805]	65 70 75 80
[0806]	Ala Val Lys Asn Val Pro Pro Lys Ala Gly Tyr Lys Glu Gly Asn Gln
[0807]	85 90 95
[0808]	Tyr Ile Val Val Glu Lys Lys Lys Lys Ser Ile Asn Gln Asn Asn Ala
[0809]	100 105 110
[0810]	Asp Ile Gln Val Ile Asn Ser Leu Ala Ser Leu Thr Tyr Pro Gly Ala
[0811]	115 120 125
[0812]	Leu Val Lys Ala Asn Ser Glu Leu Val Glu Asn Gln Pro Asp Val Leu
[0813]	130 135 140
[0814]	Pro Val Lys Arg Asp Ser Val Thr Leu Ser Ile Asp Leu Pro Gly Met
[0815]	145 150 155 160
[0816]	Val Asn His Asp Asn Glu Ile Val Val Gln Asn Ala Thr Lys Ser Asn
[0817]	165 170 175
[0818]	Ile Asn Asp Gly Val Asn Thr Leu Val Asp Arg Trp Asn Asn Lys Tyr
[0819]	180 185 190
[0820]	Ser Glu Glu Tyr Pro Asn Ile Ser Ala Lys Ile Asp Tyr Asp Gln Glu
[0821]	195 200 205
[0822]	Met Ala Tyr Ser Glu Ser Gln Leu Val Ala Lys Phe Gly Ala Ala Phe
[0823]	210 215 220
[0824]	Lys Ala Val Asn Asn Ser Leu Asn Val Asn Phe Gly Ala Ile Ser Glu
[0825]	225 230 235 240
[0826]	Gly Lys Val Gln Glu Glu Val Ile Asn Phe Lys Gln Ile Tyr Tyr Thr
[0827]	245 250 255
[0828]	Val Asn Val Asn Glu Pro Thr Ser Pro Ser Arg Phe Phe Gly Lys Ser
[0829]	260 265 270
[0830]	Val Thr Lys Glu Asn Leu Gln Ala Leu Gly Val Asn Ala Glu Asn Pro
[0831]	275 280 285
[0832]	Pro Ala Tyr Ile Ser Ser Val Ala Tyr Gly Arg Asp Ile Phe Val Lys
[0833]	290 295 300
[0834]	Leu Ser Thr Ser Ser His Ser Thr Arg Val Lys Ala Ala Phe Asp Thr
[0835]	305 310 315 320
[0836]	Ala Phe Lys Gly Lys Ser Val Lys Gly Asp Thr Glu Leu Glu Asn Ile
[0837]	325 330 335
[0838]	Ile Gln Asn Ala Ser Phe Lys Ala Val Ile Tyr Gly Gly Ser Ala Lys
[0839]	340 345 350

[0840]	Asp Glu Val Glu Ile Ile Asp Gly Asp Leu Ser Lys Leu Arg Asp Ile
[0841]	355 360 365
[0842]	Leu Lys Gln Gly Ala Asn Phe Asp Lys Lys Asn Pro Gly Val Pro Ile
[0843]	370 375 380
[0844]	Ala Tyr Thr Thr Asn Phe Leu Lys Asp Asn Gln Leu Ala Val Val Lys
[0845]	385 390 395 400
[0846]	Asn Asn Ser Glu Tyr Ile Glu Thr Thr Ser Lys Ala Tyr Ser Asp Gly
[0847]	405 410 415
[0848]	Lys Ile Asn Leu Asp His Ser Gly Ala Tyr Val Ala Arg Phe Asn Val
[0849]	420 425 430
[0850]	Thr Trp Asp Glu Val Ser Tyr Asp Ala Asn Gly Asn Glu Val Val Glu
[0851]	435 440 445
[0852]	His Lys Lys Trp Ser Glu Asn Asp Lys Asp Lys Leu Ala His Phe Thr
[0853]	450 455 460
[0854]	Thr Ser Ile Tyr Leu Pro Gly Asn Ala Arg Asn Ile Asn Ile His Ala
[0855]	465 470 475 480
[0856]	Lys Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Thr Val Val Asp
[0857]	485 490 495
[0858]	Asp Arg Asn Leu Pro Leu Val Lys Asn Arg Asn Val Cys Ile Trp Gly
[0859]	500 505 510
[0860]	Thr Thr Leu Tyr Pro Ala Tyr Ser Asp Thr Val Asp Asn Pro Ile Lys
[0861]	515 520 525
[0862]	<210> 13
[0863]	<211> 529
[0864]	<212> PRT
[0865]	<213> <i>Listeria monocytogenes</i>
[0866]	<400> 13
[0867]	Met Lys Lys Ile Met Leu Val Phe Ile Thr Leu Ile Leu Val Ser Leu
[0868]	1 5 10 15
[0869]	Pro Ile Ala Gln Gln Thr Glu Ala Lys Asp Ala Ser Ala Phe Asn Lys
[0870]	20 25 30
[0871]	Glu Asn Leu Ile Ser Ser Met Ala Pro Pro Ala Ser Pro Pro Ala Ser
[0872]	35 40 45
[0873]	Pro Lys Thr Pro Ile Glu Lys Lys His Ala Asp Glu Ile Asp Lys Tyr
[0874]	50 55 60
[0875]	Ile Gln Gly Leu Asp Tyr Asn Lys Asn Asn Val Leu Val Tyr His Gly
[0876]	65 70 75 80
[0877]	Asp Ala Val Thr Asn Val Pro Pro Arg Lys Gly Tyr Lys Asp Gly Asn
[0878]	85 90 95
[0879]	Glu Tyr Ile Val Val Glu Lys Lys Lys Lys Ser Ile Asn Gln Asn Asn
[0880]	100 105 110
[0881]	Ala Asp Ile Gln Val Val Asn Ala Ile Ser Ser Leu Thr Tyr Pro Gly

[0882]	115	120	125
[0883]	Ala Leu Val Lys Ala Asn Ser Glu Leu Val Glu Asn Gln Pro Asp Val		
[0884]	130	135	140
[0885]	Leu Pro Val Lys Arg Asp Ser Leu Thr Leu Ser Ile Asp Leu Pro Gly		
[0886]	145	150	155
[0887]	Met Thr Asn Gln Asp Asn Lys Ile Val Val Lys Asn Ala Thr Lys Ser		
[0888]		165	170
[0889]	Asn Val Asn Asn Ala Val Asn Thr Leu Val Glu Arg Trp Asn Glu Lys		175
[0890]		180	185
[0891]	Tyr Ala Gln Ala Tyr Pro Asn Val Ser Ala Lys Ile Asp Tyr Asp Asp		190
[0892]		195	200
[0893]	Glu Met Ala Tyr Ser Glu Ser Gln Leu Ile Ala Lys Phe Gly Thr Ala		205
[0894]		210	215
[0895]	Phe Lys Ala Val Asn Asn Ser Leu Asn Val Asn Phe Gly Ala Ile Ser		220
[0896]		225	230
[0897]	Glu Gly Lys Met Gln Glu Glu Val Ile Ser Phe Lys Gln Ile Tyr Tyr		235
[0898]		245	250
[0899]	Asn Val Asn Val Asn Glu Pro Thr Arg Pro Ser Arg Phe Phe Gly Lys		255
[0900]		260	265
[0901]	Ala Val Thr Lys Glu Gln Leu Gln Ala Leu Gly Val Asn Ala Glu Asn		270
[0902]		275	280
[0903]	Pro Pro Ala Tyr Ile Ser Ser Val Ala Tyr Gly Arg Gln Val Tyr Leu		285
[0904]		290	295
[0905]	Lys Leu Ser Thr Asn Ser His Ser Thr Lys Val Lys Ala Ala Phe Asp		300
[0906]		305	310
[0907]	Ala Ala Val Ser Gly Lys Ser Val Ser Gly Asp Val Glu Leu Thr Asn		315
[0908]		325	330
[0909]	Ile Ile Lys Asn Ser Ser Phe Lys Ala Val Ile Tyr Gly Gly Ser Ala		335
[0910]		340	345
[0911]	Lys Asp Glu Val Gln Ile Ile Asp Gly Asn Leu Gly Asp Leu Arg Asp		350
[0912]		355	360
[0913]	Ile Leu Lys Lys Gly Ala Thr Phe Asn Arg Glu Thr Pro Gly Val Pro		365
[0914]		370	375
[0915]	Ile Ala Tyr Thr Thr Asn Phe Leu Lys Asp Asn Glu Leu Ala Val Ile		380
[0916]		385	390
[0917]	Lys Asn Asn Ser Glu Tyr Ile Glu Thr Thr Ser Lys Ala Tyr Thr Asp		395
[0918]		405	410
[0919]	Gly Lys Ile Asn Ile Asp His Ser Gly Gly Tyr Val Ala Gln Phe Asn		415
[0920]		420	425
[0921]	Ile Ser Trp Asp Glu Ile Asn Tyr Asp Pro Glu Gly Asn Glu Ile Val		430
[0922]		435	440
[0923]	Gln His Lys Asn Trp Ser Glu Asn Asn Lys Ser Lys Leu Ala His Phe		445

[0924]	450	455	460
[0925]	Thr Ser Ser Ile Tyr Leu Pro Gly Asn Ala Arg Asn Ile Asn Val Tyr		
[0926]	465	470	475 480
[0927]	Ala Lys Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg Thr Val Ile		
[0928]		485	490 495
[0929]	Asp Asp Arg Asn Leu Pro Leu Val Lys Asn Arg Asn Ile Ser Ile Trp		
[0930]		500	505 510
[0931]	Gly Thr Thr Leu Tyr Pro Lys Tyr Ser Asn Ser Val Asp Asn Pro Ile		
[0932]		515	520 525
[0933]	Glu		
[0934]	<210> 14		
[0935]	<211> 530		
[0936]	<212> PRT		
[0937]	<213> <i>Listeria seeligeri</i>		
[0938]	<400> 14		
[0939]	Met Lys Ile Phe Gly Leu Val Ile Met Ser Leu Leu Phe Val Ser Leu		
[0940]	1	5	10 15
[0941]	Pro Ile Thr Gln Gln Pro Glu Ala Arg Asp Val Pro Ala Tyr Asp Arg		
[0942]		20	25 30
[0943]	Ser Glu Val Thr Ile Ser Pro Ala Glu Thr Pro Glu Ser Pro Pro Ala		
[0944]		35	40 45
[0945]	Thr Pro Lys Thr Pro Val Glu Lys Lys His Ala Glu Glu Ile Asn Lys		
[0946]		50	55 60
[0947]	Tyr Ile Trp Gly Leu Asn Tyr Asp Lys Asn Ser Ile Leu Val Tyr Gln		
[0948]	65	70	75 80
[0949]	Gly Glu Ala Val Thr Asn Val Pro Pro Lys Lys Gly Tyr Lys Asp Gly		
[0950]		85	90 95
[0951]	Ser Glu Tyr Ile Val Val Glu Lys Lys Lys Lys Gly Ile Asn Gln Asn		
[0952]		100	105 110
[0953]	Asn Ala Asp Ile Ser Val Ile Asn Ala Ile Ser Ser Leu Thr Tyr Pro		
[0954]		115	120 125
[0955]	Gly Ala Leu Val Lys Ala Asn Arg Glu Leu Val Glu Asn Gln Pro Asn		
[0956]		130	135 140
[0957]	Val Leu Pro Val Lys Arg Asp Ser Leu Thr Leu Ser Val Asp Leu Pro		
[0958]	145	150	155 160
[0959]	Gly Met Thr Lys Lys Asp Asn Lys Ile Phe Val Lys Asn Pro Thr Lys		
[0960]		165	170 175
[0961]	Ser Asn Val Asn Asn Ala Val Asn Thr Leu Val Glu Arg Trp Asn Asp		
[0962]		180	185 190
[0963]	Lys Tyr Ser Lys Ala Tyr Pro Asn Ile Asn Ala Lys Ile Asp Tyr Ser		
[0964]		195	200 205
[0965]	Asp Glu Met Ala Tyr Ser Glu Ser Gln Leu Ile Ala Lys Phe Gly Thr		

[0966]	210	215	220
[0967]	Ala Phe Lys Ala Val Asn Asn Ser Leu Asn Val Asn Phe Glu Ala Ile		
[0968]	225	230	235 240
[0969]	Ser Asp Gly Lys Val Gln Glu Glu Val Ile Ser Phe Lys Gln Ile Tyr		
[0970]		245	250 255
[0971]	Tyr Asn Ile Asn Val Asn Glu Pro Thr Ser Pro Ser Lys Phe Phe Gly		
[0972]		260	265 270
[0973]	Gly Ser Val Thr Lys Glu Gln Leu Asp Ala Leu Gly Val Asn Ala Glu		
[0974]		275	280 285
[0975]	Asn Pro Pro Ala Tyr Ile Ser Ser Val Ala Tyr Gly Arg Gln Val Tyr		
[0976]	290	295	300
[0977]	Val Lys Leu Ser Ser Ser Ser His Ser Asn Lys Val Lys Thr Ala Phe		
[0978]	305	310	315 320
[0979]	Glu Ala Ala Met Ser Gly Lys Ser Val Lys Gly Asp Val Glu Leu Thr		
[0980]		325	330 335
[0981]	Asn Ile Ile Lys Asn Ser Ser Phe Lys Ala Val Ile Tyr Gly Gly Ser		
[0982]		340	345 350
[0983]	Ala Lys Glu Glu Val Glu Ile Ile Asp Gly Asn Leu Gly Glu Leu Arg		
[0984]		355	360 365
[0985]	Asp Ile Leu Lys Lys Gly Ser Thr Tyr Asp Arg Glu Asn Pro Gly Val		
[0986]	370	375	380
[0987]	Pro Ile Ser Tyr Thr Thr Asn Phe Leu Lys Asp Asn Asp Leu Ala Val		
[0988]	385	390	395 400
[0989]	Val Lys Asn Asn Ser Glu Tyr Ile Glu Thr Thr Ser Lys Ser Tyr Thr		
[0990]		405	410 415
[0991]	Asp Gly Lys Ile Asn Ile Asp His Ser Gly Gly Tyr Val Ala Gln Phe		
[0992]		420	425 430
[0993]	Asn Ile Ser Trp Asp Glu Val Ser Tyr Asp Glu Asn Gly Asn Glu Ile		
[0994]	435	440	445
[0995]	Lys Val His Lys Lys Trp Gly Glu Asn Tyr Lys Ser Lys Leu Ala His		
[0996]	450	455	460
[0997]	Phe Thr Ser Ser Ile Tyr Leu Pro Gly Asn Ala Arg Asn Ile Asn Ile		
[0998]	465	470	475 480
[0999]	Tyr Ala Arg Glu Cys Thr Gly Leu Phe Trp Glu Trp Trp Arg Thr Val		
[1000]		485	490 495
[1001]	Ile Asp Asp Arg Asn Leu Pro Leu Val Lys Asn Arg Asn Val Ser Ile		
[1002]		500	505 510
[1003]	Trp Gly Thr Thr Leu Tyr Pro Arg His Ser Asn Asn Val Asp Asn Pro		
[1004]		515	520 525
[1005]	Ile Gln		
[1006]	530		
[1007]	<210> 15		

[1008]	<211>	497
[1009]	<212>	PRT
[1010]	<213>	Streptococcus suis
[1011]	<400>	15
[1012]	Met Arg Lys Ser Ser His Leu Ile Leu Ser Ser Ile Val Ser Leu Ala	
[1013]	1	5 10 15
[1014]	Leu Val Gly Val Thr Pro Leu Ser Val Leu Ala Asp Ser Lys Gln Asp	
[1015]		20 25 30
[1016]	Ile Asn Gln Tyr Phe Gln Ser Leu Thr Tyr Glu Pro Gln Glu Ile Leu	
[1017]		35 40 45
[1018]	Thr Asn Glu Gly Glu Tyr Ile Asp Asn Pro Pro Ala Thr Thr Gly Met	
[1019]		50 55 60
[1020]	Leu Glu Asn Gly Arg Phe Val Val Leu Arg Arg Glu Lys Lys Asn Ile	
[1021]		65 70 75 80
[1022]	Thr Asn Asn Ser Ala Asp Ile Ala Val Ile Asp Ala Lys Ala Ala Asn	
[1023]		85 90 95
[1024]	Ile Tyr Pro Gly Ala Leu Leu Arg Ala Asp Gln Asn Leu Leu Asp Asn	
[1025]		100 105 110
[1026]	Asn Pro Thr Leu Ile Ser Ile Ala Arg Gly Asp Leu Thr Leu Ser Leu	
[1027]		115 120 125
[1028]	Asn Leu Pro Gly Leu Ala Asn Gly Asp Ser His Thr Val Val Asn Ser	
[1029]		130 135 140
[1030]	Pro Thr Arg Ser Thr Val Arg Thr Gly Val Asn Asn Leu Leu Ser Lys	
[1031]		145 150 155 160
[1032]	Trp Asn Asn Thr Tyr Ala Gly Glu Tyr Gly Asn Thr Gln Ala Glu Leu	
[1033]		165 170 175
[1034]	Gln Tyr Asp Glu Thr Met Ala Tyr Ser Met Ser Gln Leu Lys Thr Lys	
[1035]		180 185 190
[1036]	Phe Gly Thr Ser Phe Glu Lys Ile Ala Val Pro Leu Asp Ile Asn Phe	
[1037]		195 200 205
[1038]	Asp Ala Val Asn Ser Gly Glu Lys Gln Val Gln Ile Val Asn Phe Lys	
[1039]		210 215 220
[1040]	Gln Ile Tyr Tyr Thr Val Ser Val Asp Glu Pro Glu Ser Pro Ser Lys	
[1041]		225 230 235 240
[1042]	Leu Phe Ala Glu Gly Thr Thr Val Glu Asp Leu Lys Arg Asn Gly Ile	
[1043]		245 250 255
[1044]	Thr Asp Glu Val Pro Pro Val Tyr Val Ser Ser Val Ser Tyr Gly Arg	
[1045]		260 265 270
[1046]	Ser Met Phe Ile Lys Leu Glu Thr Ser Ser Arg Ser Thr Gln Val Gln	
[1047]		275 280 285
[1048]	Ala Ala Phe Lys Ala Ala Ile Lys Gly Val Asp Ile Ser Gly Asn Ala	
[1049]		290 295 300

[1092]	Met Thr Tyr Ser Ile Asp Leu Pro Gly Leu Ala Ser Ser Asp Ser Phe
[1093]	100 105 110
[1094]	Leu Gln Val Glu Asp Pro Ser Asn Ser Ser Val Arg Gly Ala Val Asn
[1095]	115 120 125
[1096]	Asp Leu Leu Ala Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val
[1097]	130 135 140
[1098]	Pro Ala Arg Met Gln Tyr Glu Lys Ile Thr Ala His Ser Met Glu Gln
[1099]	145 150 155 160
[1100]	Leu Lys Val Lys Phe Gly Ser Asp Phe Glu Lys Thr Gly Asn Ser Leu
[1101]	165 170 175
[1102]	Asp Ile Asp Phe Asn Ser Val His Ser Gly Glu Lys Gln Ile Gln Ile
[1103]	180 185 190
[1104]	Val Asn Phe Lys Gln Ile Tyr Tyr Thr Val Ser Val Asp Ala Val Lys
[1105]	195 200 205
[1106]	Asn Pro Gly Asp Val Phe Gln Asp Thr Val Thr Val Glu Asp Leu Arg
[1107]	210 215 220
[1108]	Gln Arg Gly Ile Ser Ala Asp Arg Pro Leu Val Tyr Ile Ser Ser Val
[1109]	225 230 235 240
[1110]	Ala Tyr Gly Arg Gln Val Tyr Leu Lys Leu Glu Thr Thr Ser Lys Ser
[1111]	245 250 255
[1112]	Asp Glu Val Glu Ala Ala Phe Glu Ala Leu Ile Lys Gly Val Lys Val
[1113]	260 265 270
[1114]	Ala Pro Gln Thr Glu Trp Lys Gln Ile Leu Asp Asn Thr Glu Val Lys
[1115]	275 280 285
[1116]	Ala Val Ile Leu Gly Gly Asp Pro Ser Ser Gly Ala Arg Val Val Thr
[1117]	290 295 300
[1118]	Gly Lys Val Asp Met Val Glu Asp Leu Ile Gln Glu Gly Ser Arg Phe
[1119]	305 310 315 320
[1120]	Thr Ala Asp His Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
[1121]	325 330 335
[1122]	Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr Asp Tyr Val Glu
[1123]	340 345 350
[1124]	Thr Lys Val Thr Ala Tyr Arg Asn Gly Asp Leu Leu Leu Asp His Ser
[1125]	355 360 365
[1126]	Gly Ala Tyr Val Ala Gln Tyr Tyr Ile Thr Trp Asp Glu Leu Ser Tyr
[1127]	370 375 380
[1128]	Asp Tyr Gln Gly Lys Glu Val Leu Thr Pro Lys Ala Trp Asn Arg Asn
[1129]	385 390 395 400
[1130]	Gly Gln Asp Leu Thr Ala His Phe Thr Thr Ser Ile Pro Leu Lys Gly
[1131]	405 410 415
[1132]	Asn Val Arg Asn Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
[1133]	420 425 430

[1134]	Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Asn Asp Leu Pro Leu Val
[1135]	435 440 445
[1136]	Arg Lys Arg Thr Ile Ser Ile Trp Gly Thr Thr Leu Tyr Pro Gln Val
[1137]	450 455 460
[1138]	Glu Asp
[1139]	465
[1140]	<210> 17
[1141]	<211> 532
[1142]	<212> PRT
[1143]	<213> Streptococcus intermedius
[1144]	<400> 17
[1145]	Met Lys Thr Lys Gln Asn Ile Ala Arg Lys Leu Ser Arg Val Val Leu
[1146]	1 5 10 15
[1147]	Leu Ser Thr Leu Val Leu Ser Ser Ala Ala Pro Ile Ser Ala Ala Phe
[1148]	20 25 30
[1149]	Ala Glu Thr Pro Thr Lys Pro Lys Ala Ala Gln Thr Glu Lys Lys Pro
[1150]	35 40 45
[1151]	Glu Lys Lys Pro Glu Asn Ser Asn Ser Glu Ala Ala Lys Lys Ala Leu
[1152]	50 55 60
[1153]	Asn Asp Tyr Ile Trp Gly Leu Gln Tyr Asp Lys Leu Asn Ile Leu Thr
[1154]	65 70 75 80
[1155]	His Gln Gly Glu Lys Leu Lys Asn His Ser Ser Arg Glu Ala Phe His
[1156]	85 90 95
[1157]	Arg Pro Gly Glu Tyr Val Val Ile Glu Lys Lys Lys Gln Ser Ile Ser
[1158]	100 105 110
[1159]	Asn Ala Thr Ser Lys Leu Ser Val Ser Ser Ala Asn Asp Asp Arg Ile
[1160]	115 120 125
[1161]	Phe Pro Gly Ala Leu Leu Lys Ala Asp Gln Ser Leu Leu Glu Asn Leu
[1162]	130 135 140
[1163]	Pro Thr Leu Ile Pro Val Asn Arg Gly Lys Thr Thr Ile Ser Val Asn
[1164]	145 150 155 160
[1165]	Leu Pro Gly Leu Lys Asn Gly Glu Ser Asn Leu Thr Val Glu Asn Pro
[1166]	165 170 175
[1167]	Ser Asn Ser Thr Val Arg Thr Ala Val Asn Asn Leu Val Glu Lys Trp
[1168]	180 185 190
[1169]	Ile Gln Lys Tyr Ser Lys Thr His Ala Val Pro Ala Arg Met Gln Tyr
[1170]	195 200 205
[1171]	Glu Ser Ile Ser Ala Gln Ser Met Ser Gln Leu Gln Ala Lys Phe Gly
[1172]	210 215 220
[1173]	Ala Asp Phe Ser Lys Val Gly Ala Pro Leu Asn Val Asp Phe Ser Ser
[1174]	225 230 235 240
[1175]	Val His Lys Gly Glu Lys Gln Val Phe Ile Ala Asn Phe Arg Gln Val

[1176]		245		250		255
[1177]	Tyr Tyr Thr Ala Ser Val Asp Ser Pro Asn Ser Pro Ser Ala Leu Phe					
[1178]		260		265		270
[1179]	Gly Ser Gly Ile Thr Pro Thr Asp Leu Ile Asn Arg Gly Val Asn Ser					
[1180]		275		280		285
[1181]	Lys Thr Pro Pro Val Tyr Val Ser Asn Val Ser Tyr Gly Arg Ala Met					
[1182]		290		295		300
[1183]	Tyr Val Lys Phe Glu Thr Thr Ser Lys Ser Thr Lys Val Gln Ala Ala					
[1184]	305		310		315	320
[1185]	Ile Asp Ala Val Val Lys Gly Ala Lys Leu Lys Ala Gly Thr Glu Tyr					
[1186]		325		330		335
[1187]	Glu Asn Ile Leu Lys Asn Thr Lys Ile Thr Ala Val Val Leu Gly Gly					
[1188]		340		345		350
[1189]	Asn Pro Gly Glu Ala Ser Lys Val Ile Thr Gly Asn Ile Asp Thr Leu					
[1190]		355		360		365
[1191]	Lys Asp Leu Ile Gln Lys Gly Ser Asn Phe Ser Ala Gln Ser Pro Ala					
[1192]		370		375		380
[1193]	Val Pro Ile Ser Tyr Thr Thr Ser Phe Val Lys Asp Asn Ser Ile Ala					
[1194]	385		390		395	400
[1195]	Thr Ile Gln Asn Asn Thr Asp Tyr Ile Glu Thr Lys Val Thr Ser Tyr					
[1196]		405		410		415
[1197]	Lys Asp Gly Ala Leu Thr Leu Asn His Asp Gly Ala Phe Val Ala Arg					
[1198]		420		425		430
[1199]	Phe Tyr Val Tyr Trp Glu Glu Leu Gly His Asp Ala Asp Gly Tyr Glu					
[1200]		435		440		445
[1201]	Thr Ile Arg Ser Arg Ser Trp Ser Gly Asn Gly Tyr Asn Arg Gly Ala					
[1202]		450		455		460
[1203]	His Tyr Ser Thr Thr Leu Arg Phe Lys Gly Asn Val Arg Asn Ile Arg					
[1204]	465		470		475	480
[1205]	Val Lys Val Leu Gly Ala Thr Gly Leu Ala Trp Glu Pro Trp Arg Leu					
[1206]		485		490		495
[1207]	Ile Tyr Ser Lys Asn Asp Leu Pro Leu Val Pro Gln Arg Asn Ile Ser					
[1208]		500		505		510
[1209]	Thr Trp Gly Thr Thr Leu His Pro Gln Phe Glu Asp Lys Val Val Lys					
[1210]		515		520		525
[1211]	Asp Asn Thr Asp					
[1212]		530				
[1213]	<210> 18					
[1214]	<211> 665					
[1215]	<212> PRT					
[1216]	<213> Streptococcus mitis					
[1217]	<400> 18					

[1218]	Met Asn Gln Glu Lys Arg Leu His Arg Phe Val Lys Lys Cys Gly Leu
[1219]	1 5 10 15
[1220]	Gly Val Cys Ser Ala Val Val Ala Ala Phe Leu Leu Asn Ala Gln Gly
[1221]	20 25 30
[1222]	Val Ala Leu Ala Thr Glu Gln Gly Asn Arg Pro Val Glu Thr Glu Asn
[1223]	35 40 45
[1224]	Ile Ala Arg Gly Lys Gln Ala Ser Gln Ser Ser Thr Ala Tyr Gly Gly
[1225]	50 55 60
[1226]	Ala Ala Thr Arg Ala Val Asp Gly Asn Val Asp Ser Asp Tyr Gly His
[1227]	65 70 75 80
[1228]	His Ser Val Thr His Thr Asn Phe Glu Asp Asn Ala Trp Trp Gln Val
[1229]	85 90 95
[1230]	Asp Leu Gly Lys Thr Glu Asn Val Gly Lys Val Lys Leu Tyr Asn Arg
[1231]	100 105 110
[1232]	Gly Asp Gly Asn Val Ala Asn Arg Leu Ser Asn Phe Asp Val Val Leu
[1233]	115 120 125
[1234]	Leu Asn Glu Ala Lys Gln Glu Val Ala Arg Gln His Phe Asp Ser Leu
[1235]	130 135 140
[1236]	Asn Gly Lys Ala Glu Leu Glu Val Phe Phe Thr Ala Lys Asp Ala Arg
[1237]	145 150 155 160
[1238]	Tyr Val Lys Val Glu Leu Lys Thr Lys Asn Thr Pro Leu Ser Leu Ala
[1239]	165 170 175
[1240]	Glu Val Glu Val Phe Arg Ser Ala Thr Thr Gln Val Gly Gln Asp Arg
[1241]	180 185 190
[1242]	Thr Ala Pro Val Val Asp Gln Thr Ser Ala Leu Lys Asp Tyr Leu Phe
[1243]	195 200 205
[1244]	Gly Leu Thr Tyr Asn Pro Leu Asp Ile Leu Thr Arg Lys Gly Glu Thr
[1245]	210 215 220
[1246]	Leu Glu Asn Arg Tyr Asn Thr Ser Ala Lys Glu Gln Asn Gly Glu Phe
[1247]	225 230 235 240
[1248]	Val Val Val Glu Lys Ile Lys Lys Thr Leu Ser Thr Gly Thr Ala Asp
[1249]	245 250 255
[1250]	Val Ser Ile Asn Gly Asn Gln Asn Val Phe Leu Gly Gly Leu Tyr Lys
[1251]	260 265 270
[1252]	Ala Asn Gln Asn Leu Leu Glu Asn Gln Pro Glu Leu Ile Ser Leu Ala
[1253]	275 280 285
[1254]	Arg Ala Lys Gly Thr Val Ser Val Asp Leu Pro Gly Met Ile His Ser
[1255]	290 295 300
[1256]	Glu Asn Lys Ile Glu Ala Asn Pro Thr Thr Ser Gly Met Gln Glu Ala
[1257]	305 310 315 320
[1258]	Met Asn Thr Leu Val Glu Lys Trp Thr Lys Asn Tyr Ser Ser Ser His
[1259]	325 330 335

[1260]	Ser Val Pro Ala Arg Val Gln Tyr Glu Ser Thr Thr Ala Tyr Ser Met
[1261]	340 345 350
[1262]	Asn Gln Leu Lys Ala Lys Phe Gly Ala Asp Phe Glu Lys Ala Gly Ala
[1263]	355 360 365
[1264]	Pro Leu Lys Ile Asp Phe Glu Ala Val Gln Lys Gly Glu Lys Gln Ile
[1265]	370 375 380
[1266]	Glu Val Val Asn Phe Lys Gln Ile Tyr Tyr Thr Ala Thr Phe Asp Ala
[1267]	385 390 395 400
[1268]	Pro Thr Asn Pro Ala Ala Val Phe Asp Lys Ser Val Thr Pro Glu Asp
[1269]	405 410 415
[1270]	Leu Lys Gln Arg Gly Val Asp Ser Gln Thr Pro Pro Val Tyr Val Ser
[1271]	420 425 430
[1272]	Asn Val Ser Tyr Gly Arg Gln Ile Tyr Val Lys Phe Glu Ser Thr Ser
[1273]	435 440 445
[1274]	Lys Ser Thr Glu Leu Lys Ala Ala Ile Asn Ala Val Ile Lys Gly Ala
[1275]	450 455 460
[1276]	Thr Ile Ala Pro Asn Ser Glu Trp Ser Arg Leu Leu Lys Asn Thr Ser
[1277]	465 470 475 480
[1278]	Val Thr Ala Val Ile Val Gly Gly Asn Ala Ser Gly Ala Ala Lys Val
[1279]	485 490 495
[1280]	Val Thr Gly Thr Val Glu Asn Leu Lys Glu Leu Ile Arg Glu Gly Ala
[1281]	500 505 510
[1282]	Asn Phe Ser Ala Gln Ser Pro Ala Val Pro Ile Ser Tyr Lys Thr Ala
[1283]	515 520 525
[1284]	Phe Leu Lys Asp Asn Ala Gln Ala Thr Leu Gln Asn Ser Thr Asp Tyr
[1285]	530 535 540
[1286]	Ile Glu Thr Lys Val Thr Ser Tyr Lys Asn Gly Phe Leu Lys Leu His
[1287]	545 550 555 560
[1288]	His Lys Gly Ala Tyr Ile Ala Arg Tyr Tyr Ile Tyr Trp Asp Glu Ile
[1289]	565 570 575
[1290]	Thr Tyr Asp Glu Gln Gly Asn Pro Glu Ile Arg Ser Arg Gln Trp Glu
[1291]	580 585 590
[1292]	Asp Asn Gly Lys Asn Arg Thr Ser Gly Phe Gln Thr Glu Ile Gln Phe
[1293]	595 600 605
[1294]	Arg Gly Asn Val Arg Asn Leu Arg Ile Lys Val Gln Glu Lys Thr Gly
[1295]	610 615 620
[1296]	Leu Val Trp Glu Pro Trp Arg Thr Val Tyr Asn Arg Thr Asp Leu Pro
[1297]	625 630 635 640
[1298]	Leu Val Gln Gln Arg Thr Ile Thr His Trp Gly Thr Thr Leu Asn Pro
[1299]	645 650 655
[1300]	Lys Val Asp Glu Lys Ile Val Asn Glu
[1301]	660 665

[1302] <210> 19
 [1303] <211> 534
 [1304] <212> PRT
 [1305] <213> Arcanobacterium pyogenes
 [1306] <400> 19
 [1307] Met Lys Arg Lys Ala Phe Ala Ser Leu Val Ala Ser Val Val Ala Ala
 [1308] 1 5 10 15
 [1309] Ala Thr Val Thr Met Pro Thr Ala Ser Phe Ala Ala Gly Leu Gly Asn
 [1310] 20 25 30
 [1311] Ser Ser Gly Leu Thr Asp Gly Leu Ser Ala Pro Arg Ala Ser Ile Ser
 [1312] 35 40 45
 [1313] Pro Thr Asp Lys Val Asp Leu Lys Ser Ala Gln Glu Thr Asp Glu Thr
 [1314] 50 55 60
 [1315] Gly Val Asp Lys Tyr Ile Arg Gly Leu Lys Tyr Asp Pro Ser Gly Val
 [1316] 65 70 75 80
 [1317] Leu Ala Val Lys Gly Glu Ser Ile Glu Asn Val Pro Val Thr Lys Asp
 [1318] 85 90 95
 [1319] Gln Leu Lys Asp Gly Thr Tyr Thr Val Phe Lys His Glu Arg Lys Ser
 [1320] 100 105 110
 [1321] Phe Asn Asn Leu Arg Ser Asp Ile Ser Ala Phe Asp Ala Asn Asn Ala
 [1322] 115 120 125
 [1323] His Val Tyr Pro Gly Ala Leu Val Leu Ala Asn Lys Asp Leu Ala Lys
 [1324] 130 135 140
 [1325] Gly Ser Pro Thr Ser Ile Gly Ile Ala Arg Ala Pro Gln Thr Val Ser
 [1326] 145 150 155 160
 [1327] Val Asp Leu Pro Gly Leu Val Asp Gly Lys Asn Lys Val Val Ile Asn
 [1328] 165 170 175
 [1329] Asn Pro Thr Lys Ser Ser Val Thr Gln Gly Leu Asn Gly Leu Leu Asp
 [1330] 180 185 190
 [1331] Gly Trp Ile Gln Arg Asn Ser Lys Tyr Pro Asp His Ala Ala Lys Ile
 [1332] 195 200 205
 [1333] Ser Tyr Asp Glu Thr Met Val Thr Ser Lys Arg Gln Leu Glu Ala Lys
 [1334] 210 215 220
 [1335] Leu Gly Leu Gly Phe Glu Lys Val Ser Ala Lys Leu Asn Val Asp Phe
 [1336] 225 230 235 240
 [1337] Asp Ala Ile His Lys Arg Glu Arg Gln Val Ala Ile Ala Ser Phe Lys
 [1338] 245 250 255
 [1339] Gln Ile Tyr Tyr Thr Ala Ser Val Asp Thr Pro Thr Ser Pro His Ser
 [1340] 260 265 270
 [1341] Val Phe Gly Pro Asn Val Thr Ala Gln Asp Leu Lys Asp Arg Gly Val
 [1342] 275 280 285
 [1343] Asn Asn Lys Asn Pro Leu Gly Tyr Ile Ser Ser Val Ser Tyr Gly Arg

[1344]	290	295	300
[1345]	Gln Ile Phe Val Lys Leu Glu Thr Thr Ser Thr Ser Asn Asp Val Gln		
[1346]	305	310	315
[1347]	Ala Ala Phe Ser Gly Leu Phe Lys Ala Lys Phe Gly Asn Leu Ser Thr		
[1348]		325	330
[1349]	Glu Phe Lys Ala Lys Tyr Ala Asp Ile Leu Asn Lys Thr Arg Ala Thr		
[1350]		340	345
[1351]	Val Tyr Ala Val Gly Gly Ser Ala Arg Gly Gly Val Glu Val Ala Thr		
[1352]		355	360
[1353]	Gly Asn Ile Asp Ala Leu Lys Lys Ile Ile Lys Glu Glu Ser Thr Tyr		
[1354]	370	375	380
[1355]	Ser Thr Lys Val Pro Ala Val Pro Val Ser Tyr Ala Val Asn Phe Leu		
[1356]	385	390	395
[1357]	Lys Asp Asn Gln Leu Ala Ala Val Arg Ser Ser Gly Asp Tyr Ile Glu		
[1358]		405	410
[1359]	Thr Thr Ala Thr Thr Tyr Lys Ser Gly Glu Ile Thr Phe Arg His Gly		
[1360]		420	425
[1361]	Gly Gly Tyr Val Ala Lys Phe Arg Leu Lys Trp Asp Glu Ile Ser Tyr		
[1362]		435	440
[1363]	Asp Pro Gln Gly Lys Glu Ile Arg Thr Pro Lys Thr Trp Ser Gly Asn		
[1364]	450	455	460
[1365]	Trp Ala Ala Arg Thr Leu Gly Phe Arg Glu Thr Ile Gln Leu Pro Ala		
[1366]	465	470	475
[1367]	Asn Ala Arg Asn Ile His Val Glu Ala Gly Glu Ala Thr Gly Leu Ala		
[1368]		485	490
[1369]	Trp Asp Pro Trp Trp Thr Val Ile Asn Lys Lys Asn Leu Pro Leu Val		
[1370]		500	505
[1371]	Pro His Arg Glu Ile Val Leu Lys Gly Thr Thr Leu Asn Pro Trp Val		
[1372]		515	520
[1373]	Glu Asp Asn Val Lys Ser		
[1374]	530		
[1375]	<210> 20		
[1376]	<211> 1416		
[1377]	<212> DNA		
[1378]	<213> Streptococcus pneumoniae		
[1379]	<400> 20		
[1380]	ctagtcattt tctacattat cctctacctg aggatagaga gttgttcccc aaatagaaat 60		
[1381]	cgtcgctta cgcaactagtg gcaaatcggg tttttcataa accgtacgcc accattceca 120		
[1382]	ggcaagcccg gtacactctc taattttgac agagagatta cgaacattcc cttttaaagg 180		
[1383]	aatactagtg gtaaagtgag ccgtcaaadc ctgcccattt ctgtccaag ccttaggagt 240		
[1384]	caagacttcc ttaccttgat gatcatagga taattcatcc caagtaatat aatattgggc 300		
[1385]	aacataggca ccactatgat ccagcagtaa atctccgttt ctgtaagctg taaccttagt 360		

[1386] ctcaacatag tctgtactgt tttgaaaggt cgcaactaca ttgtcacgta aaaaagaagt 420
 [1387] tgtataggaa atcggcaagc ctggatgata tgctgtaaag cgactgcctt ctgaaatcaa 480
 [1388] gtctctacc atatccacct tgctgttac aactcgggca cccgaacttg ggtegccecc 540
 [1389] taaaataacc gccttcactt ctgtattgtc caaaatctgc ttccactctg tctgaggagc 600
 [1390] taccttgact cttttatca aagcttcaaa agcagcctct acttcatcac tcttactcgt 660
 [1391] ggtttccaac ttgagataga cttggcgcce ataagcaaca ctcgaaatat agaccaaaag 720
 [1392] acgctctgca gaaattcctc tctgttttaa atcctctacc gttacagtat ctgaaacac 780
 [1393] atctctgga tttttaacag cgtctacgct gactgtataa taaatctgct taaaattaac 840
 [1394] aatctgaatc tgctttcac ctgaatggac agagttaaaa tcaatatcaa gagaattccc 900
 [1395] tgtcttttca aagtcagaac caaactgac cttgagttgt tccatgctgt gagccgttat 960
 [1396] ttttccatac tgcaattctag ctgggacatt attgacctga ccataatctt gatgccactt 1020
 [1397] agccaacaaa tcgtttaccg ctccgcgaac acttgaattg ctggggctct ccacttggag 1080
 [1398] aaagctatcg ctacttgcca aaccaggcaa atcaactata taagtcacg gagcacgac 1140
 [1399] aaccgcaaga agagtgggat tattctctaa caaggtctca tccactacga gaagtgtcc 1200
 [1400] aggatagagg cgactgtcgt tggtagctgt tacagaaata tcacttgtat ttgtcgacaa 1260
 [1401] gctccgcttc tttcttctga taacaacaaa ctcatcgggt agctgattac cctctttgat 1320
 [1402] gaaacgattt tcaactctt ctccctgatg ggtcaagagt ttctttttat cgtaattcat 1380
 [1403] agctagtata aagtcattta ctgctttatt tgccat 1416
 [1404] <210> 21
 [1405] <211> 1530
 [1406] <212> DNA
 [1407] <213> Bacillus cereus
 [1408] <400> 21
 [1409] ctgaatatta agaaaaacac taaaagaaga aagttccttg catgtttatt agttagttaa 60
 [1410] tgcactatta attattcacc tatttccttc gcagaaacac aagcaagtaa tgcaactgat 120
 [1411] gtaacacaaa atgctagtgg cattgatact ggtatagcaa atcttaaata taataatcaa 180
 [1412] gaggttttag ctgtaaatgg tgataaagta gaaagtttg ttccgaaaga aagtatcaat 240
 [1413] tcaaatggta aattttagt agtggaacgc gagaaaaat cacttacaac gtcaccagtc 300
 [1414] gatatttcaa ttattgatc ttagtgaat cgcacgatac caggagctgt acaacttga 360
 [1415] aataaagctt ttgcagacaa tcaaccaagt ttattagtgg ctaagagaaa gcctttgaat 420
 [1416] attagtatag acttaccggg catgagaaaa gaaatacaa tcaactgtca aaatccgaca 480
 [1417] tatgtaaatg tagctggagc agtagacgat ttagtatcta cttggaatga aaagtattct 540
 [1418] acaacacata cgttacctgc aagaatgcag tatacggaat ctatggtta tagtaaatca 600
 [1419] caaatagcaa gtgctcttaa ttttaacgct aaatatcttg ataacagtct aaacattgat 660
 [1420] tttaatgcgg ttgcaaatgg agagaaaaaa gtgatgtag cggcgtataa gcaaatattt 720
 [1421] tatacggtaa gtgctgaact acctaacaat ccatccgacc tttttgataa tagcgttact 780
 [1422] tttgacgagt taactcgaag aggcgtaagt aattcggctc cacctgttat ggtttcaaat 840
 [1423] gtagcttatg gtagaacgat ttatgtaaaa ttagaaacaa catctaagag caaagatgta 900
 [1424] caagctgctt ttaaagcctt acttaagaat aacagcgttg aaacaagtgg acagtataaa 960
 [1425] gatatttttg aagaaagtac ctttactgct gtagtattag gcggagatgc gaaagagcat 1020
 [1426] aacaaggttg ttactaaaga tttcaatgaa atccgaaata tcattaaaga taatgctgaa 1080
 [1427] ttaagtctta aaaatccagc ataccaat tcatatacaa gcactttctt aaaagataat 1140

[1428]	tcaactgctg ctgttcataa caatacagat tatattgaga cgacaactac agaattattca	1200
[1429]	agtgctaaaa tgacacttga tcattacggg gcttacgttg ctcaatttga tgtatcttgg	1260
[1430]	gatgaattca catttgacca aaagggtaac gaagtactaa cacataaaac gtgggatggt	1320
[1431]	agcggaaaag acaaaacggc tcattactct acagttatcc cactcccacc aaattcaaaa	1380
[1432]	aatataaaaa ttgtagcaag agaattgaca ggtcttgcac gggaatggtg gagaacaatt	1440
[1433]	attaatgaac aaaatgtcc attaacaaat gaaataaaag tttcaattgg aggaacaaca	1500
[1434]	ttatacccaa cagctagtat tagtcattaa	1530
[1435]	<210> 22	
[1436]	<211> 1539	
[1437]	<212> DNA	
[1438]	<213> Bacillus anthracis	
[1439]	<400> 22	
[1440]	ctaagtacta atagtagcag ttggatataa tgttgttctt ccaattgaaa cttttatttc	60
[1441]	atgtgtaaat ggaacatddd gttcattaat gatagttctc caccattccc atgcaagacc	120
[1442]	tgtacattct cttgcaacga tttttatatt ttttgaattt ggtggaagag ggataactgt	180
[1443]	agagtaatgg gctgtdttgt ctttgccact accttcccat gttttatgtg ttagtacttc	240
[1444]	tttcccattt tgatcaaatg taaattcatc ccaagataca tcaaattgag caacgtaagc	300
[1445]	accgtaatga tcgagtgtca ttttagcact tgaatattct gtagttgtcg tctcgatata	360
[1446]	atctgtattg ttatgaacag cagcagttgc attatctttt aagaacgtgc ttgtatatga	420
[1447]	aattgggtat gctggatddd taaagcttaa ttcagcatta tctttaataa tatttcggat	480
[1448]	ttcattgaaa tcttttgtaa caacctgtt atgctcttcc gcactctcgc ctaatactac	540
[1449]	agcagtaaag gtactttctt caaaaatatac tttgtactgt ccaactcgtt cgacgctgtt	600
[1450]	attcttaagt agagctttaa atgcagcttg tacatcttta ctcttagatg ttgtttctaa	660
[1451]	ttttacataa accgttctac cataagctac atttgaacc ataacagtg gagccgaatt	720
[1452]	acttactcct ttacaggtta actcatcaaa agtaacacta ttatcaaaga gatcagatgg	780
[1453]	attattaggt agttcagcac ttacagtata aaatatttgt ttatatgccg ccaccatcac	840
[1454]	tttcttttct ccatttcaa ctgcattaaa gtcaatattt agactgttat caagatattt	900
[1455]	agcattaaca ttaagagcac ttgcgatttg tgatttacta taaacatag attctgtata	960
[1456]	ctgcattctt gcaggtaacg tatgtgttgt ggaatacttt tcattccaag tagatactaa	1020
[1457]	atcatctact gctccagcca cattaccata tgtcggattt tggacagtga ttgtattttc	1080
[1458]	ttttctcatg ccaggttaagt ctatactaat attcaaaggc tttctcttag ctactaataa	1140
[1459]	actcggttga ttgtctgcaa atgctttatt ggcaagttgt acagctcctg gatacgtacg	1200
[1460]	attcactaca gaatcaataa tcaaaaatatac gactgggtgac gttgtaagtg attttttctc	1260
[1461]	gcgttccact actacaaatt taccgttga attgatactc tctttcggaa taaaactctc	1320
[1462]	tactttatca ccatttactg ctaaaaatatac tctactatca tacttttagat ttgctatacc	1380
[1463]	agtattaata tcaactagcat tttttattgc accagttgca ttaccggctt gtgtttctgc	1440
[1464]	aaaagaaata gacgaataat gaatgtgaca tagactaact aataaacatg caaggaactt	1500
[1465]	tcttctttta gtgtttttct taatattcag aaaaatcac	1539
[1466]	<210> 23	
[1467]	<211> 1530	
[1468]	<212> DNA	
[1469]	<213> Bacillus thuringiensis	

[1470] <400> 23
 [1471] ctgaatatta agaaaaacac taaaagaaga aagttccttg catgtttatt agttagtcta 60
 [1472] tgcaccatta attattcatc tatttccttc gcagaaacac aagcaagtaa tgcgactgat 120
 [1473] gtaacacaaa atgctagtgg cattgatact ggtatagcaa atcttaaata taatattcaa 180
 [1474] gaggttttag ctgtaaatgg tgataaagta gaaagtttg ttccgaaaga aagtatcaat 240
 [1475] tcaaatggta aattttagt agtggaacgc gagaaaaaat cacttacaac gtcaccagtc 300
 [1476] gatatttcaa ttattgattc tgtggatgaat cgcacgtatc caggagctgt acaacttget 360
 [1477] aataaagctt ttgcagacaa tcaaccgagt ttattagtgg ctaagagaaa gcctttgaat 420
 [1478] attagtatag acttaccggg catgagaaaa gaaaatacaa tcaactgtcca aaatccgaca 480
 [1479] tatggtaatg tagctggagc agtagacgat ttagtatcta cttggaatga aaagtattct 540
 [1480] acaacacata cgttacctgc aagaatgcag tatacggaat ctatggttta tagtaaatca 600
 [1481] caaattgcaa gtgctcttaa tgttaacgct aaatatcttg ataacagtct aaatattggt 660
 [1482] tttaatgchg ttgcaaatgg agagaaaaaa gtgatgtag cggcgtataa gcaaatattt 720
 [1483] tatacggtaa gtgctgaact acctaacaat ccatccgacc tttttgataa tagcgttact 780
 [1484] tttgacgagt taactcgaag aggcgtaaat aattcggctc cacctgttat ggtttcaaat 840
 [1485] gtagcttatg gtagaacgat ttatgtaaaa ttagaaacaa catctaagag caaagatgta 900
 [1486] caagctgctt ttaaagcctt acttaagaat aacagcgttg aaacaagtgg acagtataaa 960
 [1487] gatatttttg aagaaagtac ctttactgct gtagtattag gcggagatgc gaaagaacat 1020
 [1488] aacaaggttg ttactaaaga tttcaatgaa atccgaaata ttattaaaga taatgctgaa 1080
 [1489] ttaagtctta aaaatccagc ataccaatt tcatatacaa gtactttctt aaaagataat 1140
 [1490] gcaactgctg ctgttcataa caatacagat tatattgaga cgacaactac agaattattca 1200
 [1491] agtgctaaaa tgacacttga tcattacggt gcttacgttg ctcaatttga tgtatcttgg 1260
 [1492] gatgaattca catttgatca aaagggtaac gaagtactaa cacataaac gtgggatggt 1320
 [1493] agtggaaaag acaaaaacgc tcattactct acagttatcc ctcttccacc gaattcaaaa 1380
 [1494] aatataaaaa ttgtagcaag agaattgaca ggtcttgcac gggaatggtg gagaacaatt 1440
 [1495] attaatgaac aaaacgtcc attaacaat gaaataaaag tttcaattgg aggaacaaca 1500
 [1496] ttatacccaa cagctagtat tagtcattaa 1530
 [1497] <210> 24
 [1498] <211> 1503
 [1499] <212> DNA
 [1500] <213> Clostridium perfringens
 [1501] <400> 24
 [1502] atgataagat ttaagaaaac aaaattaata gcaagtattg caatggcttt atgtctgttt 60
 [1503] tctcaaccag taatcagttt ctcaaaggat ataacagata aaaatcaaag tattgattct 120
 [1504] ggaatatcaa gtttaagtta caatagaaat gaagtttttag ctagtaatgg agataaaatt 180
 [1505] gaaagctttg ttcaaagga aggtaaaaaa gctggtaata aatttatagt ttagaacgt 240
 [1506] caaaaaagat cccttacaac atcaccagta gatatatcaa taattgattc tgtaaatgac 300
 [1507] cgtacatatc caggagcatt acaacttga gataaagcat ttgtggaaaa tagacctaca 360
 [1508] atcttaatgg taaaaagaaa gcctattaac attaatatag atttaccagg attaaagggc 420
 [1509] gaaaatagta taaaggttga tgatccaacc tatggaaaag tttctggagc aattgatgaa 480
 [1510] ttagtatcta agtggatga aaagtattca tctacacata ctttaccagc aagaactcaa 540
 [1511] tattcagaat ctatggttta tagtaaatca caaatatcaa gtgcccttaa tgttaattget 600

[1512] aaagtccttg aaaactcact tggagtagac tttaatgcag tagcaacaa tgagaaaaa 660
 [1513] gttatgattt tagcatataa acaaatattc tatacagtaa gtgcagattt acctaagaat 720
 [1514] ccatcagatc tttttgatga cagtgttaca tttaatgatt taaaacaaaa gggagtaagt 780
 [1515] aatgaagcac ctccacttat gttttcaaat gtagcttatg gaagaacaat atatgttaag 840
 [1516] ttagaaacta cttctagtag taaagatgta caagctgctt tcaaagctct tataaagaac 900
 [1517] actgatataa aaaatagtc acaatataaa gatattttatg aaaatagttc cttcacagca 960
 [1518] gtagttttag gaggagatgc acaagaacat acaaaagttg taactaaaga ctttgatgaa 1020
 [1519] ataagaaaag taattaaaga caatgcaact tttagtacaa aaaaccagc atatccaata 1080
 [1520] tcttatacta gtgttttctt aaaagataac tcagttgctg ctgttcacaa taaaacagat 1140
 [1521] tatatagaaa caacttctac agagtattct aagggaaaaa taaacttaga tcatagtgga 1200
 [1522] gcctatgttg cacagtttga agtagcctgg gatgaagttt catatgacaa agaaggaaat 1260
 [1523] gaagtttta ctcataaac atgggatgga aattatcaag ataaaacagc tcaactattca 1320
 [1524] acagtaatac ctcttgaagc taatgcaaga aatataagaa taaaagcaag agagtgtaca 1380
 [1525] ggccttgctt gggaatggtg gagagatgtt ataagtgaat atgatgttcc attaacaaat 1440
 [1526] aatataaatg tttcaatatg gggaacaact ttataccctg gatctagtat tacttacaat 1500
 [1527] taa 1503
 [1528] <210> 25
 [1529] <211> 1506
 [1530] <212> DNA
 [1531] <213> *Bacillus alvei*
 [1532] <400> 25
 [1533] atgaagaaaa aatcaaacca cttgaaagga aggaaagtac tcgtaagttt gttagtaagc 60
 [1534] ttacaagtgt tcgcttttgc gagtatttcc tccgcagcac caaccgaacc caatgatatt 120
 [1535] gatatgggga tcgcaggact gaactataat cgcaatgagg ttttgctat tcaaggggat 180
 [1536] caaatcagta gttttgttcc caaagaagc attcagtcca atggcaaatt tategtggtt 240
 [1537] gaacgggaca aaaaatcact cacaacgtca cccgtagata tttccatcgt tgattcaatc 300
 [1538] acgaatcgca cgtatccagg cgcaatacag cttgcaaata aggattttgc cgataatcag 360
 [1539] cctagtctgg ttatggctgc gagaaaacca ttagatatta gcatcgatct gcctggtttg 420
 [1540] aagaacgaga atacaatttc cgttcaaaat ccaattacg gcaactgtatc tagtgccatt 480
 [1541] gatcagctcg tgtcgacttg gggcgagaag tattccagca cgcatacact gcctgcaaga 540
 [1542] ttacaatacag ctgaatccat gttttacagc caaaatcaaa tttccagcgc cctgaatgtg 600
 [1543] aacgctaaag tactgaacgg taaactcgga atcgacttta acgcggttgc aaatggcgag 660
 [1544] aaaaaagtga tggttgctgc ttacaagcaa atcttttata ccgtaagtgc aggactgcc 720
 [1545] aacaatccat cagacttgtt cgatgacagt gtgacatttg ctgagttagc tcgcaaggga 780
 [1546] gtaagcaatg aggcaccgcc gctgatggta tctaactgg cttacggcag aactatttac 840
 [1547] gtaaaattgg aaacaacttc taagagcaat gacgtacaaa cggcatttaa attgttgctc 900
 [1548] aataatccta gcatacaagc tagcggacag tacaagata tttatgagaa cagctcgttt 960
 [1549] actgccgttg tactaggcgg cgacgcgcaa acccataacc aagtcgttac gaaagacttc 1020
 [1550] aatgttatcc aaagtgaat caaggacaat gcacaattta gcagcaagaa ccttgcttac 1080
 [1551] ccgatttcat atacaagtgt cttcttgaaa gacaattcca ttgctgctgt tcacaacaat 1140
 [1552] accgagtata tcgagacgaa aacgacgaa tattcgaagg gtaaaattaa gcttgatcat 1200
 [1553] agtgggtgcat atgtagctca gtttgaagta tattgggatg aattttcata tgatgcagat 1260

[1554]	ggacaagaaa tcgtgactcg taaaagttgg gatggaaatt ggcgcgatag atctgctcat	1320
[1555]	ttctcaacag aatccact tcctcctaata gccaagaaca taagaatttt tgcgagagaa	1380
[1556]	tgcacaggtc ttgcttggga atgggtggaga acagttgttg atgaatataa cgttccgta	1440
[1557]	gcaagcgata ttaatgttc gatttgggga acaacgttat atccgaaatc atccattact	1500
[1558]	cactaa	1506
[1559]	<210>	26
[1560]	<211>	1725
[1561]	<212>	DNA
[1562]	<213>	Streptococcus canis
[1563]	<400>	26
[1564]	atgaaggaca tgtctaataa aaaaatattt aaaaaataca gtcgcgtcgc tgggctatta	60
[1565]	acggcagctc ttatcgttg taatcttgtt actgctaata ctgactcgaa caaacaaaac	120
[1566]	actgccaata cagaaaccac aacgacaaat gaacaaccaa aaccagaaag tagtgagcta	180
[1567]	actacagaaa aagcaggta gaaaatggat gatatgctta actctaacga tatgattaag	240
[1568]	cttgctccca aagaaatgcc actagaatct gcagaaaaag aagaaaaaaa gtcagaagac	300
[1569]	aataaaaaaa gcgaagaaga tcatactgaa gaaatcaatg acaagattta ttcactaat	360
[1570]	tataatgagc ttgaagtact tgctaaaaat ggtgaaacca ttgaaaactt tgttcctaaa	420
[1571]	gaaggcgta agaaagctga caaatttatt gtcattgaaa gaaagaaaaa aaatatcaac	480
[1572]	actacaccgg tcgatatttc catcattgac tctgtcactg ataggaccta tccagcagcc	540
[1573]	cttcagctgg ctaataaagg ttttacgaa aacaaaccag acgcagtagt caccaagcga	600
[1574]	aaccacaaa aatccatat tgatttacca ggtatgggag acaaagcaac ggttgagtc	660
[1575]	aatgacccta cctatgcaa tgttcaaca gctattgata atcttgtaa ccaatggcat	720
[1576]	gataattatt ctggtgtaaa tacgcttctt gccagaacac aatatactga atcaatggta	780
[1577]	tattctaat cacagattga agcagctcta aatgttaata gtaaaatctt agatggtagt	840
[1578]	ttaggcattg atttcaagtc gatttcaaaa ggtgaaaaga aggtgatgat tgcagcatac	900
[1579]	aagcaattt tttacaccgt atcagcaaac ctctctaata atcctgcgga tgtgtttgat	960
[1580]	aatcagtgca ctttaaaaga gttgcaagca aaagggtgca gcaatgaagc cccgccactc	1020
[1581]	tttgtagta acgtagctta tggctgaact gttttgtca aactagaaac aagttctaaa	1080
[1582]	agtaatgatg ttgaagcggc ctttagtgca gctctaaaag gaacagatgt taaaactaat	1140
[1583]	ggaaaatact ctgatatttt agaaaatagt tcatttacag ctgtcgtttt aggagcagat	1200
[1584]	gctgcagagc acaataaggt agtcacaaaa gactttgatg ttattagaaa cgttatcaaa	1260
[1585]	gctaattgcta cttcagtag aaaaaacca gcttatccta tttcatacac cagtgttttc	1320
[1586]	cttaaaaata ataaaattgc ggggtgtaaat aacagaagtg aatacgttga aacaacatct	1380
[1587]	accgagtaca cgagtggaaa aattaacctg tctcatcaag gtgcctatgt tgctcaatat	1440
[1588]	gaaatccttt gggatgaaat caattatgat gacaaaggaa aagaagtgat tactaaacga	1500
[1589]	cgttgggaca acaactggta tagtaagaca tcaccattta gcacagttat cccactagga	1560
[1590]	gctaattcac gaaatatccg tatcatggct agagagtgca ccggcttagc ttgggaatgg	1620
[1591]	tggcggaaaag tgatcgacga aagagatgtg aaactgtcta aagaaatcaa tgtcaacatc	1680
[1592]	tcaggatcaa ccctgagccc atatggttcg attacttata agtag	1725
[1593]	<210>	27
[1594]	<211>	1715
[1595]	<212>	DNA

[1596]	<213>	Streptococcus equisimilis	
[1597]	<400>	27	
[1598]		atgtctaata aaaaaatatt taaaaaatac agtcgcgtcg ctgggctatt aacggcagct	60
[1599]		cttatcgttg gtaatcttgt tactgctaata ctgactcgaa caaacaaaaac actgccaata	120
[1600]		cagaaaccac aacgacaaat gaacaaccaa aaccagaaaag tagtgagcta actacagaaa	180
[1601]		aagcaggtca gaaaatggat gatatgctta actctaacga tatgattaag cttgctccca	240
[1602]		aagaaatgcc actagaatct gcagaaaaag aagaaaaaaa gtcagaagac aataaaaaaa	300
[1603]		gcgaagaaga tcatactgaa gaaatcaatg acaagattta ttcactaaat tataatgagc	360
[1604]		ttgaagtact tgctaaaaat ggtgaaacca ttgaaaactt tgttcctaaa gaaggcgta	420
[1605]		agaaagctga caaatttatt gtcattgaaa gaaagaaaaa aaatatcaac actacaccgg	480
[1606]		tcgatatttc catcattgac tctgtcactg ataggaccta tccagcagcc cttcagctgg	540
[1607]		ctaataaagg ttttaccgaa aacaaccag acgcagtagt caccaagcga aaccacaaa	600
[1608]		aatccatat tgattacca ggtatgggag ataaagcaac ggttgaggtc aatgacccta	660
[1609]		cctatgccaa tgttcaaca gctattgata atcttgtaa ccaatggcat gataattatt	720
[1610]		ctggtggtaa tacgcttct gccagaacac aatatactga atcaatggta tattctaaat	780
[1611]		cacagattga agcagctcta aatgtaata gtaaatctt agatggtagt ttaggcattg	840
[1612]		atttcaagtc gatttcaaaa ggtgaaaaga aggtgatgat tgcagcatac aagcaaattt	900
[1613]		tttacaccgt atcagcaaac cttcctaata atcctgcgga tgtgtttgat aaatcagtga	960
[1614]		cctttaaaga gttgcaacga aaaggtgtca gcaatgaagc cccgccactc tttgtgagta	1020
[1615]		acgtagctta tggctgaact gttttgtca aactagaaac aagttctaaa agtaatgatg	1080
[1616]		ttgaagcggc ctttagtga gctctaaaag gaacagatgt taaaactaat ggaaaatact	1140
[1617]		ctgatatttt agaaaatagt tcatttacag ctgtcgtttt aggaggagat gctgcagagc	1200
[1618]		acaataaggt agtcacaaaa gactttgatg ttattagaaa cgttatcaaa gataatgcta	1260
[1619]		cattcagtag aaaaaacca gcttatccta tttcatacac cagtgttttc cttaaaaata	1320
[1620]		ataaaattgc ggggtgtcaat aacagaagtg aatacgttga aacaacatct accgagtaca	1380
[1621]		cgagtggaaa aattaacctg tctcatcaag gtgcctatgt tgctcaatat gaaatccttt	1440
[1622]		gggatgaaat caattatgat gacaaaggaa aagaagtgat taaaaacga cgttgggaca	1500
[1623]		acaactggta tagtaagaca tcaccattta gcacagttat cccactagga gctaattcac	1560
[1624]		gaaatatccg tatcatggct agagagtga cgggattagc ttgggaatgg tggcgaaaag	1620
[1625]		tgatcgacga aagagatgtg aaactgtcta aagaaatcaa tgtcaacatc tcaggatcaa	1680
[1626]		ccttgagccc atatggttcg attacttata agtag	1715
[1627]	<210>	28	
[1628]	<211>	1716	
[1629]	<212>	DNA	
[1630]	<213>	Streptococcus pyogenes	
[1631]	<400>	28	
[1632]		atgtctaata aaaaaacatt taaaaaatac agtcgcgtcg ctgggctact aacggcagct	60
[1633]		cttatcattg gtaatcttgt tactgctaata gctgaatcga acaaacaaaa cactgctagt	120
[1634]		acagaaacca caacgacaag tgagcaacca aaaccagaaa gtagtgagct aactatcgaa	180
[1635]		aaagcaggtc agaaaatgga tgatatgctt aactctaacg atatgattaa gcttgctccc	240
[1636]		aaagaaatgc cactagaatc tgcaaaaaa gaagaaaaaa agtcagaaga caaaaaaaag	300
[1637]		agcgaagaag atcacactga agaaatcaat gacaagattt attcactaaa ttataatgag	360

[1638]	cttgaagtac ttgctaaaaa tggatgaaacc attgaaaatt ttgttcctaa agaaggcggt 420
[1639]	aagaaagctg ataaatttat tgtcattgaa agaaagaaaa aaaatatcaa cactacacca 480
[1640]	gtcgatattt ccattattga ctctgtcact gataggacct atccagcagc ccttcagctg 540
[1641]	gctaataaag gttttaccga aaacaaacca gacgcggtag tcaccaagcg aaaccacaaa 600
[1642]	aaaatccata ttgatttacc aggtatggga gacaaagcaa cggttgaggt caatgaccct 660
[1643]	acctatgcca atgtttcaac agctattgat aatcttgta accaatggca tgataattat 720
[1644]	tctggtggta atacgettcc tgccagaaca caatatactg aatcaatggt atatttctaag 780
[1645]	tcacagattg aagcagctct aaatgttaat agcaaaatct tagatggtac tttaggcatt 840
[1646]	gatttcaagt cgatttcaaa aggtgaaaag aaggatgata ttgcagcata caagcaaatt 900
[1647]	ttttacaccg tatcagcaaa ctttccctaat aatcctgcgg atgtgtttga taaatcagtg 960
[1648]	acctttaaag atttgcaacg aaaagggtgc agcaatgaag ctccgccact ctttgtgagt 1020
[1649]	aacgtagcct atggtcgaac tgttttgtc aaactagaaa caagttctaa aagtaatgat 1080
[1650]	gttgaagcgg cttttagtgc agctctaaaa ggaacagatg ttaaaacgaa tggaaaatac 1140
[1651]	tctgatattc tagaaaatag ctcatctaca gctgtcgttt taggaggaga tgctgcagag 1200
[1652]	cacaataagg tggtcacaaa agactttgat gttattagaa acgttatcaa agacaatgct 1260
[1653]	accttcagta gaaaaaaccc agcttatcct atttcataca ccagtgtttt ccttaaaaat 1320
[1654]	aataaaattg cgggtgtcaa taacagaact gaatacgttg aaacaacatc taccgagtac 1380
[1655]	actagtggaa aaattaacct gtctcatcaa ggcgcgtagt ttgctcaata tgaaatcctt 1440
[1656]	tgggatgaaa tcaattatga tgacaaagga aaagaagtga ttacaaaacg acgttgggac 1500
[1657]	aacaactggt atagtaagac atcaccattt agcacagtta tcccactagg agctaattca 1560
[1658]	cgaaatatcc gtatcatggc tagagagtgc accggcttag cttgggaatg gtggcgaaaa 1620
[1659]	gtgatcgacg aaagagatgt gaaactatct aaagaaatca atgtcaacat ctcaggatca 1680
[1660]	acctgagcc catatggttc gattacttat aagtag 1716
[1661]	<210> 29
[1662]	<211> 1545
[1663]	<212> DNA
[1664]	<213> Clostridium novyi
[1665]	<400> 29
[1666]	atgaagaaat ctttaaaaac tataattcgt agcatatctt ttctctcaat attaacatta 60
[1667]	actttagtgg gcaactttat aacaagcacc cagaaaaatg taagcttatt atcgggacct 120
[1668]	aataaagtta ttaaacctaa gaaaactaaa tccttagatg ataggattta tggattaaaa 180
[1669]	tatgatccta ataaaatact atcatttaat ggagaaaaag ttgaaaattt tgtacctaac 240
[1670]	gaaggttttt caacaccgga taagtacatc gttataaaac gtgaaaagaa aagtatatca 300
[1671]	gattctacag cagatatagc tgttatagac tcgatgaatg acaaaactta tcttgggtgca 360
[1672]	atacaacttg caaatagaaa tcttatagaa aacaagccta atatagtatc ttgtgaaaga 420
[1673]	aagccaataa caataagtat agatttacct gggatgggtg aagaagggaa gacaactata 480
[1674]	acttctccta cttattcttc tgttaaagca ggaattgatt cattgctaaa taagtggat 540
[1675]	tcacattatt cgtcaatata tagcattcca actagattta gctattcaga ttctatggtt 600
[1676]	tatagtaagt ctcaattatc agctaaatta gtttgtaatt tcaaagcttt aaataaagca 660
[1677]	ttggatattg attttgattc tatttataaa ggacaaaaga aagtaatgct tcttgcatat 720
[1678]	aagcaaattt tttatacagt aaatgtagat gccccaaatc atccatcaga cttctttggg 780
[1679]	gataaagtaa catttaatga cttagcaaaa aaaggagtta atagtaagaa tctctctgta 840

[1680] tacgtttcaa gtgtatctta tggtagaact atttatgtaa aacttgaaac tactttctaaa 900
 [1681] agtgccaatg taaaagctgc atttaaagca ctaatagaaa atcaaaatat aagcagtaat 960
 [1682] tctgagtaca aaaatatttt aatcaaagc tcatttacag ctacagtatt aggtgggtgga 1020
 [1683] gctaaagaac ataacaaagt aataactaaa aactttgatg aaataagaaa tataataact 1080
 [1684] aacaactcag aatatagtc tagaaatcca gggtatccta tagcatatac tactttctttt 1140
 [1685] ctaaaaagata atagtgttgc aacagtgaac aataaaacag attatataga gacaacctct 1200
 [1686] acagaatata ctaatggaaa aattactctt gatcatagag gtgcatatgt agctaaatc 1260
 [1687] aatattactt gggatgaagt aagctatgat aagaatggaa aagaaatagt agaacacaaa 1320
 [1688] tcatgggaag gaaatgactt cggtagaaca gctcatttca atacagaatt atacctaaaa 1380
 [1689] ggtaacgctc gaaatatttg cataaaagct aaagaatgca ctggcttgc ttgggaatgg 1440
 [1690] tggagaacta taatagatga taaaaatgt ctttagtta aaaacagaaa agtctatatt 1500
 [1691] tggggaacaa cattatatcc tagaacatta acagaaatag aataa 1545
 [1692] <210> 30
 [1693] <211> 1584
 [1694] <212> DNA
 [1695] <213> Clostridium tetani
 [1696] <400> 30
 [1697] ttacatttta gtttcaattg atgtctttgg atataatggt gtaccccata tatagaaagt 60
 [1698] tctttctttt gctagtggtta tatttttaac atctacaatg gttctccacc attcccatgc 120
 [1699] aagacctgta cattctctta tttttacaga tatatttctt gcatttcctt ttaaataat 180
 [1700] ttctgtatta aatgagccg ttctatctct attatttctt tccaagctt tatgtttcaac 240
 [1701] tatttcattt cccttttctt catagctaac ctcatcccat gttacttggga attgagcaac 300
 [1702] atacgtcca ctatgatcaa gtactatttt accattagta tattctgttg cagttgtttc 360
 [1703] tatatattct gttttattat ttacggatgc tatactattg tcttttaaaa atgtagtagt 420
 [1704] atatgaaatt ggatatcctg gattttgtgg actatacaact gaattatttt taataatatt 480
 [1705] tcttatttca tcaaaatcct tagtgattat cttattatgt tcttgtgctc ctccccctaa 540
 [1706] aacagtagct gtaaatgaac tttgatttaa tatatcctta tattctgcat tactacttat 600
 [1707] atctgggtta ttgattagtg ctttaaaagc tgctttaaca tgtgaactct tagaggtagt 660
 [1708] ttcaagtttt acataaattg ttctaccata tgcaacattt gaaacatag caggaggatt 720
 [1709] attattattt atccccctta aagccaatc atcaaaagt acactatcac caaataaatc 780
 [1710] tgatggacga tttgggtggg ctacacttac tgtataaaaa atttgtttgt atgcaagaag 840
 [1711] cataaccttt tttcacctt taaatatgga atcaaaatct atatttaatg ctttgtttaa 900
 [1712] agctttaaaa ttgcatccaa ctgctgcaga taactgtgat tggctataca ccatagtatc 960
 [1713] agaataactc attctttagt gtatagtata tttagatgaa tactttgaa tccatgtatc 1020
 [1714] taatatagaa tttattgcag aattaactga ggagtaggta ggtgaattaa caactttctt 1080
 [1715] accatcttca gccatacccg gtaagtcaac acttatagta ataggttttc tctcatatga 1140
 [1716] aattatatca ggtttgtttt ccataagatt tctgttagca agttgtatag ctccaggata 1200
 [1717] agttctatca tttattgagt ctataattga aatgtctgct gtggaatctg aaatactctt 1260
 [1718] cttctcaagt tttaccacaa tgaacttate tggattttca aatccttcag caggtacaaa 1320
 [1719] atttctacc tgttcacat tatatgataa tatcttacgt ggatcatagg ataatacata 1380
 [1720] aatatttttg tctatgtcac tactgttgc ttttgccaaa ttacattttg tatttgacgt 1440
 [1721] tactacttgt ccattattga ttaatgaatg ttctctaca tttcctttag ctaatacatt 1500

[1722]	actggaatta taattagata ttaagccagt cattgaaaat attagcaatg aacgtgatac	1560
[1723]	aaattttaat acatttttgt tcat	1584
[1724]	<210>	31
[1725]	<211>	1587
[1726]	<212>	DNA
[1727]	<213>	<i>Listeria ivanovii</i>
[1728]	<400>	31
[1729]	atgaaaaaaa taatgctact tttaatgaca ttgttactag taagtttacc gttagcacia	60
[1730]	gaagctcaag cagatgcctc agtatatagt taccaaggca taatttcaca catggcacca	120
[1731]	ccagcgtctc cgcctgcaaa gcctaagacg ccggttghaa agaaaaatgc agctcaaadc	180
[1732]	gatcaatata tacaagggtt ggattatgat aaaaacaata tattagtgtg cgatggagaa	240
[1733]	gctgttaaaa atgttccacc aaaagcagga tacaagaag gaaatcaata tattgtagtg	300
[1734]	gagaaaaaga aaaaatctat caatcaaat aacgcagata ttcaagtat taactcgctt	360
[1735]	gcaagcctta cttatccagg agcttttagt aaggcgaatt cagagttagt cgaaaatcaa	420
[1736]	cccgatgtcc tccctgtgaa acgagattca gttacactta gtattgattt gcctggaatg	480
[1737]	gttaaccatg acaatgaaat agtcgttcaa aatgcaacta aatccaatat aatgacgga	540
[1738]	gtgaatactt tagtagatcg ttggaataat aaatactccg aagaatacc aaatattagt	600
[1739]	gcaaaaattg actatgatca agaaatggcc tatagcgaat cgcaattagt tgcaaaattt	660
[1740]	ggtgcagcat ttaaagctgt taataatagt ttgaatghaa actttggagc gattagtgha	720
[1741]	ggtaaggtgc aagaagaagt tattaattc aaacaaattt attatactgt taatgttaat	780
[1742]	gaacctaca gcccttcag attccttggc aaaagtgha ctaaagaaaa cttgcaagcg	840
[1743]	ctggcggha atgcgghaaa tccaccgca tacatctcta gtgttgcata tggctgtgac	900
[1744]	atttctgtga aattatcgac tagttcacac agcaccagag tgaaggctgc attcgatact	960
[1745]	gcatttaagg gtaaatcagt taaagggtat acagaattag aaaatattat tcaaatgct	1020
[1746]	tcatttaaaag cgggtattta tgggtgttca gccaaagatg aagtagaaat aattgatgha	1080
[1747]	gatttaagca aattacgaga tattttaaaa caaggggcta attttgataa gaaaaatccg	1140
[1748]	ggcgtaccga ttgcgtatac aactaattc ttgaaagata atcagttagc agttgttaaa	1200
[1749]	aataattcgg aatatacga aacaacttct aaggcttact cggatgghaa aattaaccta	1260
[1750]	gatcattccg gtgcctatgt tgcgagattc aatgttactt gggatgaagt tagctatgat	1320
[1751]	gctaattgha atgaagttgt tgaacataaa aatgggtccg aaaatgataa agataagtha	1380
[1752]	gctcatttta cgacatcaat ctatttgcca gggaatgha ggaatattaa tattcatgca	1440
[1753]	aaagaatgha ctggcttggc ttgggaatgg tggagaacgg ttgtggatgha tagaaactg	1500
[1754]	ccattagtha aaaaagaaaa gtgtttgatc tgggghaaca cgctttatcc agcgtatagt	1560
[1755]	gatactgtag ataatccaat taagtha	1587
[1756]	<210>	32
[1757]	<211>	1590
[1758]	<212>	DNA
[1759]	<213>	<i>Listeria monocytogenes</i>
[1760]	<400>	32
[1761]	atgaaaaaaa taatgctagt ttttattaca cttatattag ttagtctacc aattgcaaa	60
[1762]	caaaactgaag caaaggatgc atctgcattc aataaagaaa atttaattc atccatgha	120
[1763]	ccaccagcat ctccgctgc aagtccctaa acgccaatcg aaaaghaaca cgcgatgha	180

[1764]	atcgataagt atatacaagg attggattac aataaaaaca atgtattagt ataccacgga	240
[1765]	gatgcagtga caaatgtgcc gccaaagaaa ggttataaag atggaaatga atatatcgtt	300
[1766]	gtggagaaaa agaagaaatc catcaatcaa aataatgcag atatccaagt tgtgaatgca	360
[1767]	atttcgagcc taacatatcc aggtgctctc gtgaaagcga attcgggaatt agtagaaaat	420
[1768]	caaccgatg ttcttctgt caaacgtgat tcattaacac ttagcattga tttgccagga	480
[1769]	atgactaatc aagacaataa aattgttgta aaaaatgcta ctaaatecga cgtaacaac	540
[1770]	gcagtaaata cattagtga aagatggaat gaaaaatag ctcaagctta tccaaatgta	600
[1771]	agtgcaaaaa ttgattatga tgacgaaatg gcttacagt aatcacaatt aattgcaaaa	660
[1772]	tttggtagcg catttaaagc tgtaaataat agcttgaatg taaacttcgg cgcaatcagt	720
[1773]	gaaggaaaa tgcaagaaga agtcattagt tttaaacaaa tttactataa cgtgaatggt	780
[1774]	aatgaacctc caagacctc cagattttc ggcaaagctg ttactaaaga gcagttgca	840
[1775]	gcgcttgag tgaatgcaga aaatcctct gcataatct caagtgtgce atatggccgt	900
[1776]	caagtttatt tgaattatc aactaattc catagtacta aagtaaaagc tgcttttgac	960
[1777]	gctgccgtaa gtgggaaatc tgtctcaggt gatgtagaac tgacaaatc catcaaaaat	1020
[1778]	tcttcttca aagccgtaat ttacggtagc tccgcaaaag atgaagttca aatcatcgac	1080
[1779]	ggtaacctcg gagacttac agatattttg aaaaaagtg ctacttttaa ccgggaaaca	1140
[1780]	ccaggagttc ccattgccta tacaacaaac ttcttaaaag acaatgaatt agctgttatt	1200
[1781]	aaaaacaact cagaatatat tgaacaact tcaaaagctt atacagatgg aaaaatcaac	1260
[1782]	atcgatcact ctggaggata cgttgtcaa ttcaacatct cttgggatga aataaattat	1320
[1783]	gatcctgaag gtaacgaaat tgttcaacat aaaaactgga gcgaaaaca taaaagtaag	1380
[1784]	ctagctcatt tcacatcgtc catctatttg ccaggtaac caagaaatat taatgtttac	1440
[1785]	gctaaagaat gcaactggtt agcttgggaa tggtaggagaa cggtaatgta tgaccggaac	1500
[1786]	ctaccgcttg tgaaaaatag aaatatctcc atctggggca ctacacttta tccgaaatat	1560
[1787]	agtaatagtg tagataatcc aatcgaataa	1590
[1788]	<210>	33
[1789]	<211>	1593
[1790]	<212>	DNA
[1791]	<213>	Listeria seeligeri
[1792]	<400>	33
[1793]	atgaaaatat ttggtttagt tatcatgtcg ttgctatttg ttagtttgcc aataacaca	60
[1794]	caacctgaag cgagggatgt ccccgctac gatagaagcg aggtgactat atctctgct	120
[1795]	gaaacaccag agtccccacc ggcaacacca aaaacacctg tagagaaaa gcatgcggaa	180
[1796]	gaaattaata aatataattg gggattaaac tatgataaaa atagtattct ggtctatca	240
[1797]	ggtgaagcag ttacaaactg tccaccgaaa aagggtaca aagatggcag tgaatatatt	300
[1798]	gtcgttga aaagaaaa aggtatcaat caaaacaatg cagacattc tgtcataat	360
[1799]	gcaatttga gccttactta tcctggagcg ttggtaaaag caaatagaga attagtagaa	420
[1800]	aatcaacctc atgtactacc agtaaacga gattcactta cattaagtgt agatttacc	480
[1801]	ggaatgacta aaaaagataa taaaatattc gttaaaaacc ctacaaagtc aaacgtaaat	540
[1802]	aatgccgtga atacattagt agagcgttg aatgataagt attcaaaagc gtatcctaat	600
[1803]	attaatgcaa aaattgatta ttccgatgaa atggcttata gtgaatcaca attaattgcc	660
[1804]	aaatttggga ctgcctttaa agctgttaat aatagtttga atgtaattt tgaggcaatt	720
[1805]	agtgatggga aagtacaaga agaagtcatt agttttaagc aaatttatta taatattaac	780

[1806]	gttaatgaac ctacaagtcc ttccaaatc tttggggta gtgttaccaa agaacaacta	840
[1807]	gatgcttttag gtgtaaagtc cgaaaatcct cctgcttaca tttctagtgt tgcttacggt	900
[1808]	cgccaagttt atgtgaaact atcctctagc tcgcatagta acaaagttaa aactgctttc	960
[1809]	gaggcggcga tgagtggcaa atcagtgaaa ggggatgtag aattaacaaa tattataaaa	1020
[1810]	aattcttctt ttaaagcagt catttatggt ggctcagcga aagaagaggt tgaattatt	1080
[1811]	gatggcaatt taggcgaact tcgagatatt ttgaaaaaag ggtccactta tgatagagaa	1140
[1812]	aacctggcg ttccgatctc gtacacaact aactttttga aagataatga cttagcggtt	1200
[1813]	gttaaaaaca actcagaata tatcgaaca acttcgaaat cttatacaga tggaaaaatt	1260
[1814]	aatattgatc attctggtgg ttatgtagcc caattcaata tatcttggga tgaagtaagt	1320
[1815]	tatgacgaga acggaaatga aataaaagtt cataagaaat ggggcgaaaa ttataagagt	1380
[1816]	aagttagctc atttacttc ttctatctat ttgccaggaa atgcgagaaa tattaacatc	1440
[1817]	tatgcaagag aatgtaccgg cttgttttgg gaatggtgga gaactgttat agatgacaga	1500
[1818]	aacttaccat tagtaaaaa tagaaatgta tctatttggg gtacaacgct ttaccaaga	1560
[1819]	cattctaata atgtagataa tccgattcag tag	1593
[1820]	<210>	34
[1821]	<211>	1494
[1822]	<212>	DNA
[1823]	<213>	Streptococcus suis
[1824]	<400>	34
[1825]	atgagaaaaa gttcgcaact gattttaagc tcaatagtca gtttggcaact cgtaggggtc	60
[1826]	acaccattga gtgttcttgc agattccaaa caagatatta atcagtattt tcaaagcttg	120
[1827]	acttacgagc cacaagagat tcttacaat gagggagaat acattgataa tccgccagca	180
[1828]	acaactggta tgtagaaaa cggacgttt gtagtactc gcagagaaaa gaagaatatt	240
[1829]	acgaacaata gtgcagatat tgctgttatt gatgctaagg ctgcaaatat ttatccaggt	300
[1830]	gctttattgc gtgctgacca aaatctctg gataataatc caacgcttat cagtattgag	360
[1831]	cggggagatc tgacgcttag tttgaattta cctggtttgg ccaatgggga tagccacact	420
[1832]	gttgtaaatt ctccaacaag aagtactgtt cgaacagggg tgaataacct tctgtctaaa	480
[1833]	tggaataata cgtatgctgg agagtatggc aataccaag cagagcttca atatgatgaa	540
[1834]	acaatggcat acagtatgtc acaattgaaa acgaagtctg gaacctctt tgaaaaaatt	600
[1835]	gctgtaccat tagatatcaa ttttgatgcc gtgaattcgg gtgaaaaaca ggttcagatt	660
[1836]	gttaacttta acaaattta ttatacagtt agtgttgatg aaccagaatc tccaagcaag	720
[1837]	cttttgcag aagggacaac tgtagaagat ttgaaacgaa atgggataac agatgaggt	780
[1838]	cctctgttt atgtttccag cgtttcttat ggacgcteta tgttcatcaa gttagaaact	840
[1839]	agcagtagga gtaccaagt tcaagccgca tttaaagcag ccatcaaagg cgttgatatt	900
[1840]	agtggcaatg ctgagtatca agacattctg aaaaatactt cattctctgc ttatattttt	960
[1841]	ggtgggatg caggtagcgc ggctactgtt gtgagcggaa atattgaaac actgaagaag	1020
[1842]	attattgaag aaggtgcaag atacggaaaa ctcaatccag gtgttccgat ttcgtattca	1080
[1843]	accaactttg tcaaagacaa tagacctgct cagattttga gcaattcaga gtacatagaa	1140
[1844]	acaacttcaa cagtccataa tagcagtgca ttgacattgg atcattcagg tgcttatggt	1200
[1845]	gcgaaataca acattacttg ggaagaagta tcttacaatg aagctggaga agaagtttgg	1260
[1846]	gaacccaaaag cttgggataa gaatggtgta aatctgacct cacactggag tgaaccatt	1320
[1847]	caaattccag gaaatgctcg caatcttcat gtcaatattc aagaatgtac aggattagca	1380

[1848] tgggagtggg ggagaacagt ttatgacaaa gatttaccac ttgttggta acgtaaaata 1440
 [1849] accatctggg gaacaacggt ataccacag tatgcggatg aggtgataga gtaa 1494
 [1850] <210> 35
 [1851] <211> 1398
 [1852] <212> DNA
 [1853] <213> Streptococcus mitis
 [1854] <400> 35
 [1855] atggcaata aagcagtaa tgactttata ctagctatgg attacgataa aaagaaactc 60
 [1856] ttgaccatc agggagaaag tattgaaaat cgtttcatca aagaggggaa tcagctacc 120
 [1857] gatgagttt ttgttatcga aagaaagaag cggagcctgt cgacaaatac aagtgatatt 180
 [1858] tctgtgacag ctaccaacga cagtcgctc tatcctggag cacttctcgt agtggatgag 240
 [1859] acctgttag agaataatcc cactcttctt gcggtcgatc gtgctccgat gacttatagt 300
 [1860] attgatttgc ctggtttggc aagtagtgat agctttctcc aagtagaaga tcccagcaat 360
 [1861] tcaagtgttc gcggagcggg aaacgatttg ttggctaagt ggcataaga ttatggtcag 420
 [1862] gtcaataatg tcccagctag aatgcagtat gaaaaaatca cggctcacag catggaacaa 480
 [1863] ctcaaggtca agtttggttc tgactttgaa aagacagga attctcttga tattgatttt 540
 [1864] aactctgtcc attcgggcca aaagcagatt cagattgta attttaagca gatttattat 600
 [1865] acagtcagcg tagatgctgt taaaaatcca ggagatgat ttcaagatac tgtaacggta 660
 [1866] gaggatttga ggcagagagg aatttctgcc gatcgtcctt tggctctatat ttcgagtgtt 720
 [1867] gcttatgggc gccaggttta tctcaagttg gagaccacga gtaagagtga tgaagtcgag 780
 [1868] gctgcttttg aagctttgat aaaaggagtc aaggtagctc ctacagacaga gtggaagcag 840
 [1869] attttgaca atacagaagt gaagcgggtt attttagggg gcgaccagag ttcgggtgcc 900
 [1870] cgagttgtaa caggcaaggt ggatatgta gaggactga ttcaagaagg cagtcgcttt 960
 [1871] acagccgatc atccaggttt gccgatttct tatacaactt cttttttacg ggacaatgta 1020
 [1872] gttgcgacct ttcaaacag tacagactat gttgagacta aggtgacagc ctacagaaac 1080
 [1873] ggagatttac tgctggatca tagtgggtcc tatgttgctc aatattacat tacttgggat 1140
 [1874] gaattatcct atgattatca aggtaaggaa gttttgactc ctaaggcttg gaacagaaat 1200
 [1875] gggcaggatt tgacggctca ttttaccact agtattcctt taaaagggaa tgttcgaat 1260
 [1876] ctctctgtca aaattagaga gtgtaccgga cttgectggg aatggtggcg tacggtttat 1320
 [1877] gaaaaaacg atttgccct agtgcgtaag cggacgattt ctatttgggg aacgactctt 1380
 [1878] tatcctcagg tagaggat 1398
 [1879] <210> 36
 [1880] <211> 1599
 [1881] <212> DNA
 [1882] <213> Streptococcus intermedius
 [1883] <400> 36
 [1884] atgaaaacta agcagaatat tgctcgaaa ttgtcaagag ttgttttatt aagcactctc 60
 [1885] gttctctctt cggcagcacc gatttcagct gcattcgtg aaacacctac caaaccaaaa 120
 [1886] gcagctcaa cagagaaaa acccgaaaag aaaccgaaa acagcaactc tgaagctgca 180
 [1887] aaaaaagctc tgaatgatta tatttgggga ttgcagtatg ataaactaa cattttaaca 240
 [1888] caccaaggtg aaaaattgaa aaaccactct agccgcgagg catttcatcg cccaggtgag 300
 [1889] tatgttgta tcgaaaagaa aaaacaaagc atttcaaagc caacatctaa gttatctgta 360

[1890]	agttcagcaa atgatgaccg catcttccca ggtgcattgc taaaagcggc ccaaagtttg	420
[1891]	ttagaaaatc ttccaaccct aatcccagtt aatcgcgca aaacaactat tagtgtaaac	480
[1892]	ttaccgggat tgaaaaatgg cgaaagtaat cttacagttg aaaatccatc caacagtaca	540
[1893]	gttcgaacag ctgttaacaa tttagttgaa aatggattc aaaaatactc taaaactcat	600
[1894]	gctgtgccag ctagaatgca atatgaatct attagcgccc aaagcatgag ccaattacaa	660
[1895]	gcaaaatttg gtgctgattt ctcaaaagtc ggtgcaccac ttaatgttga cttctcatct	720
[1896]	gttcacaaaag gtgaaaaaca agtatttatt gcgaacttta gacaagttta ctacacagct	780
[1897]	agcgtagact ctccaaatag tccttctgca ctctttggct ctggtatcac accaactgat	840
[1898]	ttaatcaatc gtggagttaa ttctaaaacc ccaccagttt atgtttcaaa tgtatcatat	900
[1899]	ggccgtgcaa tgtatgtgaa atttgaaact acaagcaaga gtacaaaagt acaagccgct	960
[1900]	attgatgctg ttgttaaagg agcaaaactt aaagctggaa cagaatatga aaatattcta	1020
[1901]	aaaaatacta aatcaactgc tgttgttctc ggtggtaacc caggtgaage ttctaaagtc	1080
[1902]	atcacaggta atattgatac tttgaaagat ttgatccaaa aaggtagcaa tttcagtgct	1140
[1903]	caaagtccag ctgtaccaat ctcttacct acttcttttg taaaagacaa ttctattgca	1200
[1904]	actatccaaa acaacacaga ctacatcgaa acaaaagtaa catcctataa agatgggtct	1260
[1905]	ctcacctca atcatgatgg tgctttcgtt gcacgcttct atgtttattg ggaagaacte	1320
[1906]	ggacatgatg ctgatggcta cgaaactatt cgctcaagat cttggagtgg aaatggctac	1380
[1907]	aatcgcggtg cacactatc tacaactctc cgtttcaagg gaaatgtag aaacattcgc	1440
[1908]	gtaaaagtac taggagccac tggactagct tgggagcctt ggagactgat ctatagcaag	1500
[1909]	aacgatcttc ctttggttcc acaagaaac attagcactt ggggaacaac cttcatcca	1560
[1910]	cagtttgaag ataaagttgt gaaagataac actgattaa	1599
[1911]	<210>	37
[1912]	<211>	1998
[1913]	<212>	DNA
[1914]	<213>	<i>Streptococcus mitis</i>
[1915]	<400>	37
[1916]	atgaatcaag aaaaacgttt gcatcgcttt gtcaaaaagt gtggactcgg tgtgtgtagt	60
[1917]	gctgtgttg cggccttttt attgaacgct cagggagtag ctttggctac agagcaaggg	120
[1918]	aatcgtccag ttgagaccga aaacatcgct cgtggaaaac aagctagtca aagttctact	180
[1919]	gcttatggag gagctgctgc acgagcagtg gacggtaatg ttgacagtga ctatgggcac	240
[1920]	cattctgtaa cgcacacaaa ctttgaagac aatgcttggg ggcaagtga tcttggaaaa	300
[1921]	acagagaatg ttggaaaagt taaactctac aatcgtggag atgggaatgt agctaactgt	360
[1922]	ctttccaatt ttgatgttgt tttgttaaat gaagcaaac aagaagttgc tcgtcaacac	420
[1923]	tttgatagtt tgaatgggaa ggcagaactt gaagttttct tcaccgcaa agccgctcgt	480
[1924]	tatgtaaaag tggagttgaa aactaaaaat acccattga gcttggcaga agtggaaagta	540
[1925]	ttcgttcag caacaactca agttggacag gatagaactg caccagtagt agatcaacaa	600
[1926]	tcagcattga aagactacct ttttggctta gcttataatc ctttggatat tttactcgc	660
[1927]	aaggagaaa ccttagaaaa tcgctacaac acaagtgcta aggaacaaaa tggagagttt	720
[1928]	gtcgtttag aaaaaatcaa gaaaaccctc tctacaggca cagcagatgt ttccatcaat	780
[1929]	ggaaatcaaa atgtcttct aggtggcttg tataaggcaa accaaaatct gctagaaaat	840
[1930]	cagccagaat tgattagtct tgcccgtgca aaggggacag tcagcgtcga tttacctggt	900
[1931]	atgattcagt ctgacagccg aattgaagca gatcctacaa ctagtggtat gcagtctgcg	960

[1932] atgaataacct tggttgaaag atggacaaag aattactcat ctagccattc cgttcctgcc 1020
 [1933] cgtgttcagt atgaatcaac tacagcctat agcatgaatc aattgaaagc aaaatttggg 1080
 [1934] gcagactttg aaaaagcagg tgcaccgctc aagattgact ttgaggctgt gcaaaaagggt 1140
 [1935] gaaaagcaaa ttgaagttgt aaactttaa caaatctact atacagcgac atttgatgca 1200
 [1936] ccgaccaatc cagcggctgt atttgacaag agtgtgacac ctgaagattt aaaacaaaga 1260
 [1937] ggcgttgatt cacaaactcc acctgtatat gtatcaaatg tttcttacgg acgtcaaate 1320
 [1938] tatgttaagt ttgagtcagc aagtaagtct actgaattaa aagcagctat taatgcgggt 1380
 [1939] atcaaaggcg caacaattgc tccaaattct gaatggagcc gtctattgaa gaatacttct 1440
 [1940] gtaacagcgg taattgtagg aggtaatgct agcggtgccg ccaaagttgt cacaggaaca 1500
 [1941] gtcgaaaact tgaaggaact catcagagaa ggagctaact ttagcgcctca aagtccagct 1560
 [1942] gtgccaatct catataagac tgccttccca aaagataatg cccaagcaac ttacaaaaat 1620
 [1943] agtacagact atatcgaaac gaaggttact tcttacaaaa atggtttctt gaaacttcat 1680
 [1944] cataagggtg cttatgttgc gcgttactac atctattggg atgaaattac atatgatgaa 1740
 [1945] caaggaaatc cagaaatccg ttcacgtcaa tgggaagata acggaaaaaa cagaacttca 1800
 [1946] ggcttccaaa cagagattca atttagagga aatgtccgta atcttcgcat caaggttcaa 1860
 [1947] gaaaagacgg gtcttgtatg ggaacatgg cgtacagttt acaaccgcac agacttacca 1920
 [1948] ctagtacaac agcgtacaat tacacattgg ggaacaactc taaaccctaa agttgatgaa 1980
 [1949] aaaattgtga atgagtaa 1998
 [1950] <210> 38
 [1951] <211> 1605
 [1952] <212> DNA
 [1953] <213> Arcanobacterium pyogenes
 [1954] <400> 38
 [1955] atgaaacgaa aggcctttgc atcgtatgtag gcgagtgtag tgcagcagc aactgtcacg 60
 [1956] atgcccacag catcttttgc tgccggattg gaaacagct cgggattgac ggacggcttg 120
 [1957] tcagcgcgc gagcctccat ctccccgac gataaagttg accttaagtc ggcgcaagag 180
 [1958] accgacgaga cggcgtcga taagtacatt cgtggctga aatacgate cctctggtga 240
 [1959] cttgcagtcagggtgagtc tattgaaaat gtgccggtta ccaaggatca gctcaaggac 300
 [1960] ggcacctaca cggatattta gcatgaacgc aagagtttta acaatttgcg ttccgacatc 360
 [1961] tctgcgttcg atgcgaaca cccccagtc tatcctggcg cgtcgtgtt agcaaataaa 420
 [1962] gatcttcaa aaggtagtc gacttcgac ggaattgac gtgctccgca aactgtcagc 480
 [1963] gtcgacttgc caggattagt tgacggtaag aataagtcg tcatcaaca tcccacgaag 540
 [1964] agttccgtga ctcaaggact gaacggcctt ctcgacggtt ggattcagcg caatagcaag 600
 [1965] tatcctgacc atgctgcaaa gatctctac gatgagacta tgggtgacgc aaagcgtcaa 660
 [1966] ctggaggcaa agcttggcct cggattttaa aaggtctcag ccaagctcaa cgtggacttc 720
 [1967] gatgcaattc ataagcgtga acggcagtg gctatcgctt cttcaaca gatttactac 780
 [1968] acggctagcg tagatacacc gacatctcca catagcgtt tcggcccga tgctaccgca 840
 [1969] caggatttga aagatcgggg agtcaataac aagaatctc taggatacat ttctcggtc 900
 [1970] agctatggac gccagatttt tgtcaagctg gaaacgacct cgacttcaa tgatgtaca 960
 [1971] gcggctttta gcggcctgtt caaagctaag ttcggcaatc tttccacaga gttcaaggct 1020
 [1972] aagtatgccg atatctgaa caagaccgca gctactgtgt acgctgttgg tggcagcgt 1080
 [1973] agaggcggag ttgaagttgc aactggcaat atcgatec tcaagaagat catcaaggag 1140

[1974]	gaaagcacct actctacgaa ggttctgcc gtgcccgttt cctatgccgt caatttcttg										1200
[1975]	aaggataatc agttggcagc tgtaggagc agcggtgatt acattgaaac cactgcaacg										1260
[1976]	acttacaagt ctggtagat caccttccgc catggcggtg gctacgtcgc aaagttcagg										1320
[1977]	ctgaagtggg acgagatcag ctacgaccg cagggaagg aaatccgtac acccaagacg										1380
[1978]	tggagcggga attgggcagc ccgaccctt ggcttccgtg agactattca acttccagca										1440
[1979]	aacgcgcga acatccatgt ggaagcaggc gaggcaactg gcctagcgtg ggatccgtgg										1500
[1980]	tggaccgta tcaataagaa gaatctcccc ttgggtccac atcgagagat cgctcttaag										1560
[1981]	ggcagcagc tcaatccctg ggtcaggac aatgtcaaat cctag										1605
[1982]	<210> 39										
[1983]	<211> 11										
[1984]	<212> PRT										
[1985]	<213> Streptococcus pneumoniae										
[1986]	<400> 39										
[1987]	Glu Cys Thr Gly Leu Ala Trp Glu Trp Trp Arg										
[1988]	1				5					10	
[1989]	<210> 40										
[1990]	<211> 471										
[1991]	<212> PRT										
[1992]	<213> Streptococcus pneumoniae										
[1993]	<400> 40										
[1994]	Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala Met Asn Tyr Asp										
[1995]	1				5					10	15
[1996]	Lys Lys Lys Leu Leu Thr His Gln Gly Glu Ser Ile Glu Asn Arg Phe										
[1997]					20					25	30
[1998]	Ile Lys Glu Gly Asn Gln Leu Pro Asp Glu Phe Val Val Ile Glu Arg										
[1999]					35					40	45
[2000]	Lys Lys Arg Ser Leu Ser Thr Asn Thr Ser Asp Ile Ser Val Thr Ala										
[2001]					50					55	60
[2002]	Thr Asn Asp Ser Arg Leu Tyr Pro Gly Ala Leu Leu Val Val Asp Glu										
[2003]					65					70	75
[2004]	Thr Leu Leu Glu Asn Asn Pro Thr Leu Leu Ala Val Asp Arg Ala Pro										
[2005]					85					90	95
[2006]	Met Thr Tyr Ser Ile Asp Leu Pro Gly Leu Ala Ser Ser Asp Ser Phe										
[2007]					100					105	110
[2008]	Leu Gln Val Glu Asp Pro Ser Asn Ser Ser Val Arg Gly Ala Val Asn										
[2009]					115					120	125
[2010]	Asp Leu Leu Ala Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val										
[2011]					130					135	140
[2012]	Pro Ala Arg Met Gln Tyr Glu Lys Ile Thr Ala His Ser Met Glu Gln										
[2013]					145					150	155
[2014]	Leu Lys Val Lys Phe Gly Ser Asp Phe Glu Lys Thr Gly Asn Ser Leu										
[2015]					165					170	175

[2016]	Asp Ile Asp Phe Asn Ser Val His Ser Gly Glu Lys Gln Ile Gln Ile
[2017]	180 185 190
[2018]	Val Asn Phe Lys Gln Ile Tyr Tyr Thr Val Ser Val Asp Ala Val Lys
[2019]	195 200 205
[2020]	Asn Pro Gly Asp Val Phe Gln Asp Thr Val Thr Val Glu Asp Leu Lys
[2021]	210 215 220
[2022]	Gln Arg Gly Ile Ser Ala Glu Arg Pro Leu Val Tyr Ile Ser Ser Val
[2023]	225 230 235 240
[2024]	Ala Tyr Gly Arg Gln Val Tyr Leu Lys Leu Glu Thr Thr Ser Lys Ser
[2025]	245 250 255
[2026]	Asp Glu Val Glu Ala Ala Phe Glu Ala Leu Ile Lys Gly Val Lys Val
[2027]	260 265 270
[2028]	Ala Pro Gln Thr Glu Trp Lys Gln Ile Leu Asp Asn Thr Glu Val Lys
[2029]	275 280 285
[2030]	Ala Val Ile Leu Ser Gly Asp Pro Ser Ser Gly Ala Arg Val Val Thr
[2031]	290 295 300
[2032]	Gly Lys Val Asp Met Val Glu Asp Leu Ile Gln Glu Gly Ser Arg Phe
[2033]	305 310 315 320
[2034]	Thr Ala Asp His Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
[2035]	325 330 335
[2036]	Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr Asp Tyr Val Glu
[2037]	340 345 350
[2038]	Thr Lys Val Thr Ala Tyr Arg Asn Gly Asp Leu Leu Leu Asp His Ser
[2039]	355 360 365
[2040]	Gly Ala Tyr Val Ala Gln Tyr Tyr Ile Thr Trp Asp Glu Leu Ser Tyr
[2041]	370 375 380
[2042]	Asp His Gln Gly Lys Glu Val Leu Thr Pro Lys Ala Trp Asp Arg Asn
[2043]	385 390 395 400
[2044]	Gly Gln Asp Leu Thr Ala His Phe Thr Thr Ser Ile Pro Leu Lys Gly
[2045]	405 410 415
[2046]	Asn Val Arg Asn Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
[2047]	420 425 430
[2048]	Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Thr Asp Leu Pro Leu Val
[2049]	435 440 445
[2050]	Arg Lys Arg Thr Ile Ser Ile Trp Gly Thr Thr Asp Tyr Pro Gln Val
[2051]	450 455 460
[2052]	Glu Asp Lys Val Glu Asn Asp
[2053]	465 470

腊状芽孢杆菌溶解素	VNRTPGAVQ	LANKAFADNQ	PSLLVAKRKP	LNISIDLPGM	-RKENTITVQ	NPTYGNVAGA	VDDLSTWNE	177
芽孢杆菌溶解素	VNRTPGAVQ	LANKAFADNQ	PSLLVAKRKP	LNISIDLPGM	-RKENTITVQ	NPTYGNVAGA	VDDLSTWNE	180
苏云金芽孢杆菌溶解素	ANRTPGAVQ	LANKAFADNQ	PSLLVAKRKP	LNISIDLPGM	-RKENTITVQ	NPTYGNVAGA	VDDLSTWNE	180
产气荚膜杆菌溶解素	NDRTPGALQ	LADKAFVENR	PTILMVKRKP	ININIDLPLG	-KGENSIKVD	DPTYGKVSQA	IDELVSKWNE	167
蜂房芽孢杆菌溶解素	TNRTPGAIQ	LANKDFADNQ	PSLVMAARKP	LDISIDLPLG	-KNENTISVQ	NPNYGTVSSA	IDQLVSTWGE	169
犬链球菌溶解素	TDRTPAALQ	LANKGFTENK	PDAVVTKRNP	QKIHIDLPGM	-GDKATVEVN	DPTYANVSTA	IDNLVNQWHD	241
类马链球菌溶解素	TDRTPAALQ	LANKGFTENK	PDAVVTKRNP	QKIHIDLPGM	-GDKATVEVN	DPTYANVSTA	IDNLVNQWHD	238
链球菌溶解素0	TDRTPAALQ	LANKGFTENK	PDAVVTKRNP	QKIHIDLPGM	-GDKATVEVN	DPTYANVSTA	IDNLVNQWHD	238
诺维氏梭菌溶解素	NDKTPGAIQ	LANRNLLENK	PNIIVSCERKP	ITISIDLPGM	-GEEGKTTIT	SPTYASVVKAG	IDSLLNKWNS	181
破伤风溶解素	NDRTPGAIQ	LANRNLLENK	PDIIVSCERKP	ITISIDLPGM	-GEEGKTTIT	SPTYASVVKAG	IDSLLNKWNS	181
伊万诺夫李斯特菌溶解素	ASLTPPGALV	KANSELVENQ	PDVLPVKRDS	VTLISIDLPGM	VNHDNEIVVQ	NATKSNINDG	VNTLVDRWNN	190
李斯特菌溶解素	SSLTPPGALV	KANSELVENQ	PDVLPVKRDS	VTLISIDLPGM	VNHDNEIVVQ	NATKSNINDG	VNTLVDRWNN	190
李斯特菌溶解素	SSLTPPGALV	KANSELVENQ	PDVLPVKRDS	VTLISIDLPGM	VNHDNEIVVQ	NATKSNINDG	VNTLVDRWNN	190
李斯特菌溶解素	SSLTPPGALV	KANRELVENQ	PNVLPVKRDS	LTLSIDLPGM	TNQQDNKIVVK	NATKSNVNA	VNTLVERWNE	191
猪溶解素	AANIYPGALL	RADQNLLDNN	PTLISIARGD	LTLSIDLPGM	TKKDNKIFVK	NPTKSNVNA	VNTLVERWND	192
肺炎球菌溶解素	DSRLYPGALL	VVDETLLEN	PTLLAVDRAP	MTYSIDLPLG	ANGDSHTVVN	SPTRSTVRTG	VNNLLSKWNN	163
链球菌溶解素 (PLY)	dsrlypgall	vvdetlenn	ptllavdrap	mtysidplgl	assdsflqve	dpsnssvrga	vndllakwhq	136
缓症链球菌溶解素	DDRIFPGALL	kvdetlenn	ptllavdrap	mtysidplgl	assdsflqve	dpsnssvrga	vndllakwhq	136
中间链球菌溶解素	NONVFLGGLY	KADQSLENN	PTLIPVNRGK	TTISVNLPLG	KNGESNLTVE	NPSNSTVRTA	VNNLVEKWIQ	194
草绿色链球菌溶解素	NAHVYPGALV	KANQNLENN	PELISLARAK	GTIVSVDLPGM	IHSENKIEA-	NPTTSGMQEA	MNTLVEKWTK	330
化脓隐杆		LANKDLAKGS	PTSIGIARAP	QTVSVDLPLG	VDGKNKVVIN	NPTKSSVTQG	LNGLLDGIWQ	196
腊状芽孢杆菌溶解素	KYSTTH-TLP	ARMQYTESMV	YSKSQIASAL	NVNAKYLDNS	LNIDFNAPAN	GEKKVMVAAY	KQIFYTVSAE	246
芽孢杆菌溶解素	KYSTTH-TLP	ARMQYTESMV	YSKSQIASAL	NVNAKYLDNS	LNIDFNAPAN	GEKKVMVAAY	KQIFYTVSAE	249
苏云金芽孢杆菌溶解素	KYSETH-TLP	ARMQYTESMV	YSKSQIASAL	NVNAKYLDNS	LNIDFNAPAN	GEKKVMVAAY	KQIFYTVSAE	249
产气荚膜杆菌溶解素	KYSSSTH-TLP	ARTQYTESMV	YSKSQISSAL	NVNAKVLENS	LGVDFNAPAN	NEKKVMILAY	KQIFYTVSAD	236
蜂房芽孢杆菌溶解素	KYSSSTH-TLP	ARLQYAESMV	YSQNISSAL	NVNAKVLNGT	LGIDFNAPAN	GEKKVMVAAY	KQIFYTVSAG	238
犬链球菌溶解素	NYSGGN-TLP	ARTQYTESMV	YSKSQIEAAL	NVNSKILDGT	LGIDFKSISK	GEKKVMIAAY	KQIFYTVSAN	310
类马链球菌溶解素	NYSGGN-TLP	ARTQYTESMV	YSKSQIEAAL	NVNSKILDGT	LGIDFKSISK	GEKKVMIAAY	KQIFYTVSAN	307
链球菌溶解素0	NYSGGN-TLP	ARTQYTESMV	YSKSQIEAAL	NVNSKILDGT	LGIDFKSISK	GEKKVMIAAY	KQIFYTVSAN	307
诺维氏梭菌溶解素	HYSSIIY-SIP	TRFSYSDSMV	YSKSQLSAKL	GCNFKALNKA	LDIDFDSIYK	GQKKVMLLAY	KQIFYTVNVD	250
破伤风溶解素	KYSSKY-TIP	TRMSYSDTMV	YSQSLSAAV	GCNFKALNKA	LDIDFDSIFK	GEKKVMLLAY	KQIFYTVSVD	262
伊万诺夫李斯特菌溶解素	KYSEYPNIS	AKIDYDQEMA	YSESQLVAKF	GAAFKAVNNS	LVNDFGAISE	GKQVEEVINF	KQIYTVNUN	260
李斯特菌溶解素	KYAQAYPNVS	AKIDYDDEMA	YSESQLIAKF	GTAFKAVNNS	LVNDFGAISE	GKQVEEVISF	KQIYNNVUN	261
李斯特菌溶解素	KYSKAYPNIN	AKIDYSDEMA	YSESQLIAKF	GTAFKAVNNS	LVNDFEAI SD	GKQVEEVISF	KQIYNNINUN	262
猪溶解素	TYAGEYGNTO	AELQYDEMTA	YSMSQLKTKF	GTSFEKI AVP	LDINFDVNS	GEKKQIVNIF	KQIYTVSVD	233
链球菌溶解素 (PLY)	DYGOVN-NVP	ARMQYKITA	HSMEQLKVKF	GSDFEKTTGNS	LDIDFNSVHS	GEKQIQIVNF	KQIYTVSVD	205
缓症链球菌溶解素	dygqvn-nvp	armqyekita	hsmeqlkvkf	gsdfektgns	ldidfnsvhs	gekqiqivnf	kqiytvsvd	205
中间链球菌溶解素	KYSKTH-AVP	ARMQYESI SA	QSMSQLQAKF	GADFSKVGAP	LVNDFSSVHK	GEKQVFI ANF	RQVYVYATASVD	263
草绿色链球菌溶解素	NYSSSH-SVP	ARVQYESTTA	YSMNQLKAKF	GADFEKAGAP	LKIDFEAVQK	GEKQIEVNF	KQIYVYATFD	399
化脓隐杆	RNSK-YPDHA	AKISYDETMV	TSKRQLEAKL	GLGFEKVS AK	LVNDFDAIHK	RERQVAIASF	KQIYVYATASVD	265

图1C

芽胞杆菌溶素	LPNNPSDLFD	NSVTFDELTR	KGVSNSAPPV	MVSNVAYGRT	IYVKLETTSK	SKDVQAAAFKA	LLK----	NNS 312
芽胞杆菌溶素	LPNNPSDLFD	NSVTFDELTR	KGVSNSAPPV	MVSNVAYGRT	YVVKLETTSK	SKDVQAAAFKA	LLK----	NNS 315
芽胞杆菌溶素	LPNNPSDLFD	NSVTFDELTR	KGVSNSAPPV	MVSNVAYGRT	VYVKLETTSK	SKDVQAAAFKA	LLK----	NNS 315
产气荚膜杆菌溶素	LPKNPSDLFD	DSVTFNDLKR	KGVSNEAPPV	MVSNVAYGRT	IYVKLETTSS	SKDVQAAAFKA	L1K----	NTD 302
蜂房芽胞杆菌溶素	LPNNPSDLFD	DSVTFDELAR	KGVSNEAPPV	MVSNVAYGRT	IYVKLETTSS	SNDVQTAFLK	LLN----	NPS 304
犬链球菌溶素	LPNNPADVFD	KS VTFKELQA	KGVSNEAPPV	FVSNVAYGRT	VFVKLETTSSK	SNDVEAAFS A	ALK----	GTD 376
类马链球菌溶素	LPNNPADVFD	KS VTFKELQR	KGVSNEAPPV	FVSNVAYGRT	VFVKLETTSSK	SNDVEAAFS A	ALK----	GTD 373
链球菌溶素0	LPNNPADVFD	KS VTFKDLQR	KGVSNEAPPV	FVSNVAYGRT	VFVKLETTSSK	SNDVEAAFS A	ALK----	GTD 373
诺维氏梭菌溶素	APNHPSDLFFG	DKVTFNDLAK	KGVSNKPPV	YVSSVAYGRT	IYVKLETTSSK	SANVKA AFKA	L1E----	NQN 316
破伤风溶素	PPNRPDLFFG	DSVTFDELAL	KGINNPPPA	YVSSVAYGRT	IYVKLETTSSK	SHVKA AFKA	L1N----	NQD 328
伊万诺夫李斯特菌溶素	EPTSPSRFFG	KS VTKENLQA	LGVAENPPA	YISSVAYGRD	IFVKLSTSSH	STRVKA AFDT	AFK----	GKS 326
李斯特菌溶素	EPTSPSRFFG	KAVTKEQLQA	LGVAENPPA	YISSVAYGRQ	VYVKLSTNSH	STKVKA AFD A	AVS----	GKS 327
李斯特菌溶素	EPTSPSKFFG	GSVTKEQLDA	LGVAENPPA	YISSVAYGRQ	VYVKLSSSSH	SNKVKA TAFE A	AMS----	GKS 328
李斯特菌溶素	EPEPSKLF A	EGTTVEDLKR	NGITDEVPPV	YVSSVSYGRS	MF1KLETTSSR	STQVQA AFKA	A1K----	GVD 299
猪溶素	AVKNPGDV FQ	DTVTVEDLKQ	RGISAERPLV	YISSVAYGRQ	VYVKLETTSSK	SDEVEAAFE A	L1K----	GVK 271
肺炎球菌溶素	avknpqdvfq	dtvtvedlrq	rglsadrplv	YISSVAYGRQ	VYVKLETTSSK	SDEVEAAFE A	L1K----	GVK 271
链球菌溶素 (PLY)	SPNPSALFG	SGITPTDLIN	RGVNSKT PPV	YVSNVSYGRA	MYVKFETTSSK	STKVQA AIDA	VVK----	GAK 329
中间链球菌溶素	APTNPAAVFD	KS VTPEDLKQ	RGVDSQT PPV	YVSNVSYGRQ	IYVKFETTSSK	STELKAAINA	V1K----	GAT 465
草绿色链球菌溶素	TPTSPHSVFG	PNVTAQDLKD	RGVNNKNPLG	YISSVSYGRQ	IYVKLETTST	SNDVQA AFSG	L1FKAKFGNLS	335
化脓性链球菌溶素								
芽胞杆菌溶素	VETSGQYKDI	FEESTFTAVV	LGDDAKEHNK	VVTKDFNEIR	NI IKDNAELS	LKNPAYPISY	TSTFLKDNST	382
芽胞杆菌溶素	VETSGQYKDI	FEESTFTAVV	LGDDAKEHNK	VVTKDFNEIR	NI IKDNAELS	FKNPAYPISY	TSTFLKDNAT	385
芽胞杆菌溶素	VETSGQYKDI	FEESTFTAVV	LGDDAKEHNK	VVTKDFNEIR	NI IKDNAELS	FKNPAYPISY	TSTFLKDNAT	385
产气荚膜杆菌溶素	IKNSQYKDI	YENSSFTAVV	LGDDAQEHNK	VVTKDFDEIR	KV1KDNATFS	TKNPAYPISY	TSVFLKDNVS	372
蜂房芽胞杆菌溶素	IQASGQYKDI	YENSSFTAVV	LGDDAQEHNK	VVTKDFNVIQ	SV1KDNAQFS	SKNPAYPISY	TSVFLKDNIS	374
犬链球菌溶素	VKTNGKYS DI	LENS SFTAVV	LGDDAAEHNK	VVTKDFDVI R	NV1KANATFS	RKNPAYPISY	TSVFLKNNK1	446
类马链球菌溶素	VKTNGKYS DI	LENS SFTAVV	LGDDAAEHNK	VVTKDFDVI R	NV1KANATFS	RKNPAYPISY	TSVFLKNNK1	443
链球菌溶素0	VKTNGKYS DI	LENS SFTAVV	LGDDAAEHNK	VVTKDFDVI R	NV1KANATFS	RKNPAYPISY	TSVFLKNNK1	443
诺维氏梭菌溶素	ISSNSEYKNI	LNQSSFTATV	LGDDAKEHNK	VITKNFDEIR	NI ITNNS EYS	PRNP GYPIAY	TTSFLKDNVS	386
破伤风溶素	ISSNAEYKDI	LNQSSFTATV	LGDDAQEHNK	IITKDFDEIR	NI IKNSVYS	PQNP GYPIAY	TTTTFLKDNIS	398
伊万诺夫李斯特菌溶素	VKGDTELENI	IQNASFKAVI	YGGS AKDEVE	IDGDL SKLR	DILKQGANFD	KKNPGVPIAY	TTNFKDNQL	396
李斯特菌溶素	VSGDVELTNI	IKNS SFKAVI	YGGS AKDEVQ	IDGNLGLDR	DILKKGATFN	RETPGVPIAY	TTNFKDNEL	397
李斯特菌溶素	VKGDVELTNI	IKNS SFKAVI	YGGS AKEEVE	IDGNLGLDR	DILKKGATFN	RETPGVPIAY	TTNFKDNEL	398
李斯特菌溶素	ISGNAEYQDI	LKNTSFSAYI	YGGDSAGSAAT	VVSGNIETLK	K1IEEGARYD	KLNP GVPISY	STNFKVKNRP	369
猪溶素	VAPQTEWKQI	LDNTEVKAVI	LGDDPSSGAR	VVTGKVDME	DL1QEGRFT	ADHPGLPISY	TTSFLRDNV	341
链球菌溶素	vapqtewkqi	ldntevkavi	lgddpsgar	vvtgkvdmve	dl1qegsrft	adhpglpisy	ttstflrdnv	341
链球菌溶素 (PLY)	LKAGTEYENI	LKNTKITAVV	LGGNPGEASK	VITGNIDTLK	DL1QKGSNFS	AQSPAVPISY	TTSFVKDNIS	399
中间链球菌溶素	IAPNSEWSRL	LKNTSVTAVI	VGGNASGAAK	VVTGTVENLK	EL1REGANFS	AQSPAVPISY	KTAF LKDNAQ	535
草绿色链球菌溶素	TEFKAKYADI	LNKTRATVYA	VGGSARGGVE	VATGNIDALK	K11KEESTYS	TKVPAVPVSY	AVNFKDNQL	405
化脓性链球菌溶素								

图1D

腊状芽孢杆菌溶解素	AAVHNNTDYI	ETTTTEYSSA	KMTLDHYGAY	VAFQFDVSWDE	FTFDQKGNEV	LTHKTTWEGSG	KDKTAAHYSITV	452
芽孢杆菌溶解素	AAVHNNTDYI	ETTTTEYSSA	KMTLDHYGAY	VAFQFDVSWDE	FTFDQNGKEV	LTHKTTWEGSG	KDKTAAHYSITV	455
产气荚膜杆菌溶解素	AAVHNNTDYI	ETTTTEYSSA	KMTLDHYGAY	VAFQFDVSWDE	FTFDQNGKEV	LTHKTTWEGSG	KDKTAAHYSITV	455
蜂房芽孢杆菌溶解素	AAVHNKTDYI	ETTSTEYSKG	KINLDHSGAY	VAFQFEVAWDE	VSYDKEGNEV	LTHKTTWEGSG	CDKTAHYSITV	442
尖链球杆菌溶解素	AAVHNKTDYI	ETTSTEYSKG	KINLDHSGAY	VAFQFEVAWDE	FSYDADGQEI	VTRKTTWEGNW	CDKTAHYSITV	444
尖链球杆菌溶解素0	AGVNNRSEYV	ETTSTEYTSK	KINLSHQGAY	VAFQYEILWDE	INYYDDKGKEV	ITKRRWDDNNW	YSKTSFFSITV	516
链球菌溶解素	AGVNNRSEYV	ETTSTEYTSK	KINLSHQGAY	VAFQYEILWDE	INYYDDKGKEV	ITKRRWDDNNW	YSKTSFFSITV	513
诺氏链球菌溶解素	ATVNNKTDYI	ETTSTEYTSK	KINLSHQGAY	VAFQYEILWDE	INYYDDKGKEV	ITKRRWDDNNW	YSKTSFFSITV	513
破伤风溶解素	AVNNKTDYI	ETTATEYTSK	KINLDHSGAY	VAFQFQVTWDE	VSYDKNGKEI	VEHKWEGND	RDRTAHFNTE	456
伊万诺夫链球菌溶解素	AVVKNNSYI	ETTSKAYS DG	KINLDHSGAY	VAFQFQVTWDE	VSYDANGNEV	VEHKWSEND	KDKLAHFTTS	466
李斯特菌溶解素	AVIKNNSEYI	ETTSKAYTDG	KINLDHSGAY	VAFQFNI SWDE	INYPDEGNEI	VQHKNWSENN	KSKLAHFTTS	467
李斯特菌溶解素	AVVKNNSYI	ETTSKAYTDG	KINLDHSGAY	VAFQFNI SWDE	VSYDANGNEV	VEHKWSENN	KSKLAHFTTS	468
李斯特菌溶解素	AVVKNNSYI	ETTSKAYTDG	KINLDHSGAY	VAFQFNI SWDE	VSYDANGNEV	VEHKWSENN	KSKLAHFTTS	468
粘多糖	AQILSNSEYI	ETTSVHNSS	ALLDHSYGAY	VAFQYNI TWEE	VSYNEAGEEV	WEPKAWDKNG	VNLTSHWSET	439
肺炎球菌溶解素	ATFQNSTDYV	ETKVTAYRNG	DLLDHSGAY	VAFQYNI TWEE	LSYDHQKKEV	LTPKAWDRNG	CDLTAHFTTS	411
链球菌溶解素(PLY)	atfqnstdyv	etkvtayrng	dllldhsgay	vafqyнитwee	lsydyqgkev	ltpkawdrng	cdltahtts	411
中间链球菌溶解素	ATIQNNTDYI	ETKVTYSKDG	ALLDHSYGAY	VAFQRFVYWEE	LGHADAGYET	IRSRWSWGNG	YNRGAHYSITV	469
草绿色链球菌溶解素	ATLQNSTDYI	ETKVTYSKDG	FLKLHKGAY	VAFQRYIYWDE	ITYDEQGNPE	IRSRQWEDNG	KNRTSGFQTE	605
化脓链球菌溶解素	AAVRSNGDYI	ETTATTYKSG	EITFRHGGAY	VAFKFLK WDE	ISYDPQKKEI	RTPKTWSGNW	AAARTLGFRET	475
链球菌溶解素	KIVARECTGL	KIVARECTGL	AWEWRTIIN	EQNVPLTNEI	KVSI G	PTTLY	PTASIS-H-	509
芽孢杆菌溶解素	KIVARECTGL	KIVARECTGL	AWEWRTIIN	EQNVPLTNEI	KVSI G	PTTLY	PTASIS-H-	512
产气荚膜杆菌溶解素	KIVARECTGL	KIVARECTGL	AWEWRTIIN	EQNVPLTNEI	KVSI G	PTTLY	PTASIS-H-	512
蜂房芽孢杆菌溶解素	RIKARECTGL	RIKARECTGL	AWEWRTDVIS	EYDVPLTNNI	NVSIW	PTTLY	PCSSIT-VN-	500
尖链球杆菌溶解素	RIKARECTGL	RIKARECTGL	AWEWRTDVIS	EYDVPLTNNI	NVSIW	PTTLY	PCSSIT-VN-	500
尖链球杆菌溶解素	RIFARECTGL	RIFARECTGL	AWEWRTVVD	EYNVPLASDI	NVSIW	PTTLY	PCSSIT-YK-	501
尖链球杆菌溶解素0	RIFARECTGL	RIFARECTGL	AWEWRTVVD	EYNVPLASDI	NVSIW	PTTLY	PCSSIT-YK-	501
链球菌溶解素	RIMARECTGL	RIMARECTGL	AWEWRTKVID	ERDVKLSKEI	NVSIW	PTTLY	PCSSIT-YK-	574
链球菌溶解素	RIMARECTGL	RIMARECTGL	AWEWRTKVID	ERDVKLSKEI	NVSIW	PTTLY	PCSSIT-YK-	571
诺氏链球菌溶解素	CIKARECTGL	CIKARECTGL	AWEWRTIID	ERDVKLSKEI	NVSIW	PTTLY	PCSSIT-YK-	571
破伤风溶解素	CIKARECTGL	CIKARECTGL	AWEWRTIID	ERDVKLSKEI	NVSIW	PTTLY	PCSSIT-YK-	571
链球菌溶解素	SVKIRECTGL	SVKIRECTGL	AWEWRTIIVD	DRNLPLVKNR	KVYIWE	PTTLY	PRTLTE-IE-	514
破伤风溶解素	SVKIRECTGL	SVKIRECTGL	AWEWRTIIVD	DRNLPLVKNR	KVYIWE	PTTLY	PRTLTE-IE-	514
伊万诺夫链球菌溶解素	NIHARECTGL	NIHARECTGL	AWEWRTIIVD	VKNIP LAKER	TFYIWE	PTTLY	PKTSIE-TK-	527
李斯特菌溶解素	NIHARECTGL	NIHARECTGL	AWEWRTIIVD	VKNIP LAKER	TFYIWE	PTTLY	PKTSIE-TK-	527
李斯特菌溶解素	NYAKARECTGL	NYAKARECTGL	AWEWRTIIVD	DRNLPLVKNR	NVSIW	PTTLY	PAYSIT-VDN	PIK 528
李斯特菌溶解素	NYAKARECTGL	NYAKARECTGL	AWEWRTIIVD	DRNLPLVKNR	NVSIW	PTTLY	PAYSIT-VDN	PIK 528
李斯特菌溶解素	NIYARECTGL	NIYARECTGL	FWEWRTVID	DRNLPLVKNR	NVSIW	PTTLY	PHSNN-VDN	PIE 529
链球菌溶解素	HVNIQECTGL	HVNIQECTGL	AWEWRTVVD	-KDLP L VQR	KITIWE	PTTLY	PCYADE-V-	PIQ 530
链球菌溶解素	HVNIQECTGL	HVNIQECTGL	AWEWRTVVD	-KDLP L VQR	KITIWE	PTTLY	PCYADE-V-	PIQ 530
链球菌溶解素(PLY)	SVKIRECTGL	SVKIRECTGL	AWEWRTIIVD	KTDLPLVQR	TISIWE	PTTLY	PQVEDK-VEN	-D 471
链球菌溶解素	svkirectgl	svkirectgl	awewrtiivd	ktdlplvqr	tisiwe	pttly	paved----	466
链球菌溶解素	rvkvlgatgl	rvkvlgatgl	awewrllivs	knldplvpqr	nistw	pttly	pcfedkvvkd	NTD 532
链球菌溶解素	rvkvlgatgl	rvkvlgatgl	awewrllivs	knldplvpqr	nistw	pttly	pcfedkvvkd	NTD 532
链球菌溶解素	RIKVQECTGL	RIKVQECTGL	WVEPRTVYN	RTDLP L VQR	TITHW	PTTLY	PKVDEKIV--	NE 665
化脓链球菌溶解素	HVEAGEATGL	HVEAGEATGL	AWDPWTVIN	KKNLPLVPHR	EIVLK	PTTLY	PKVEDNVKS-	--- 534

环3

环2

环1

图1E

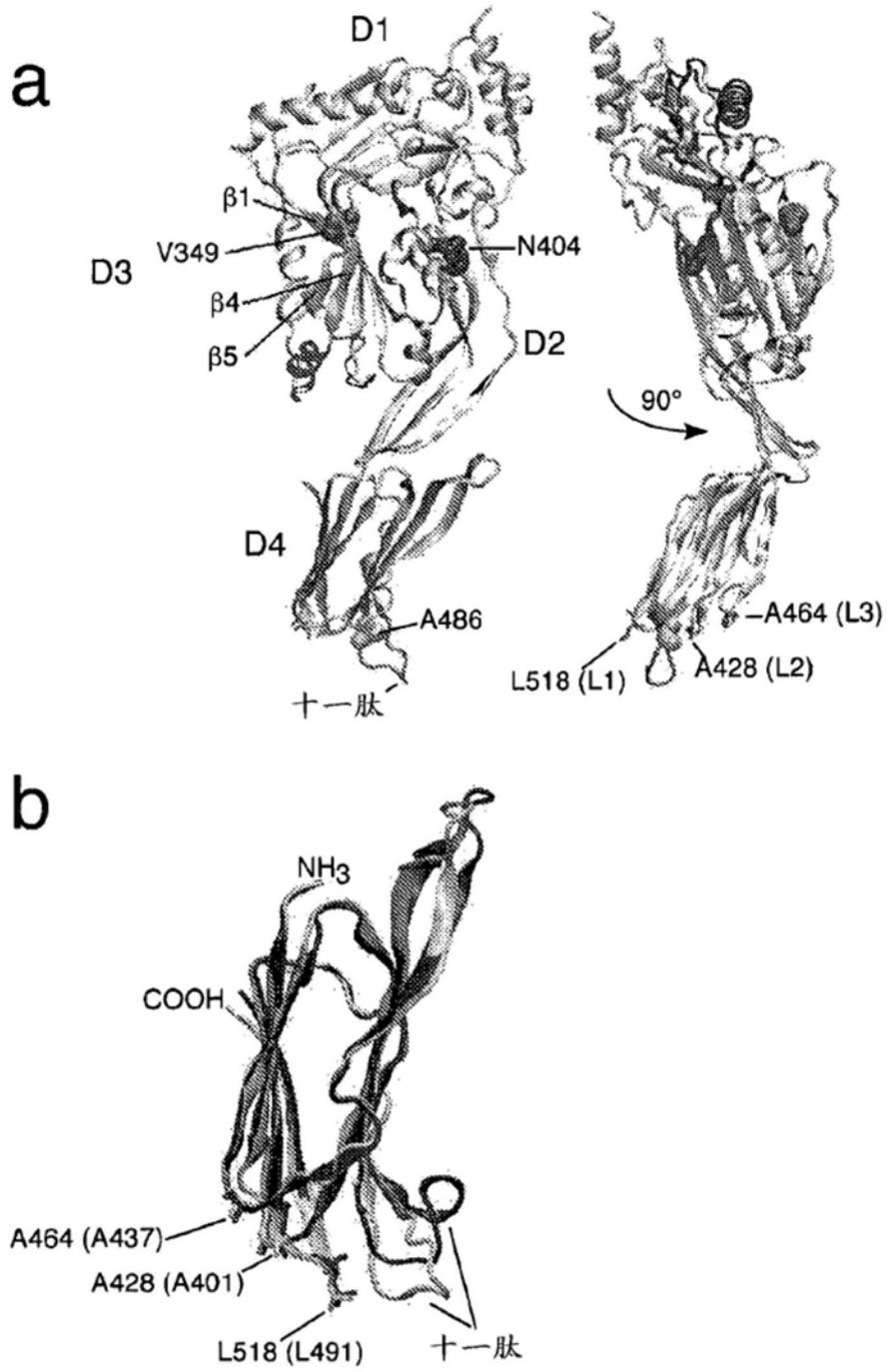


图2

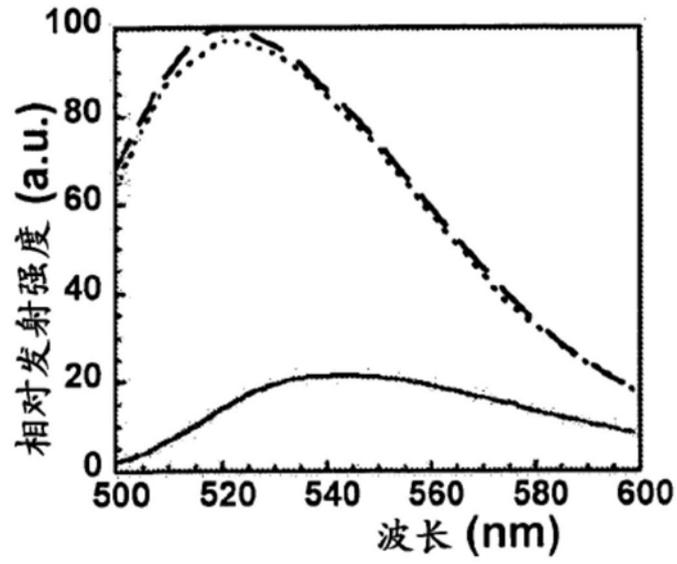


图3

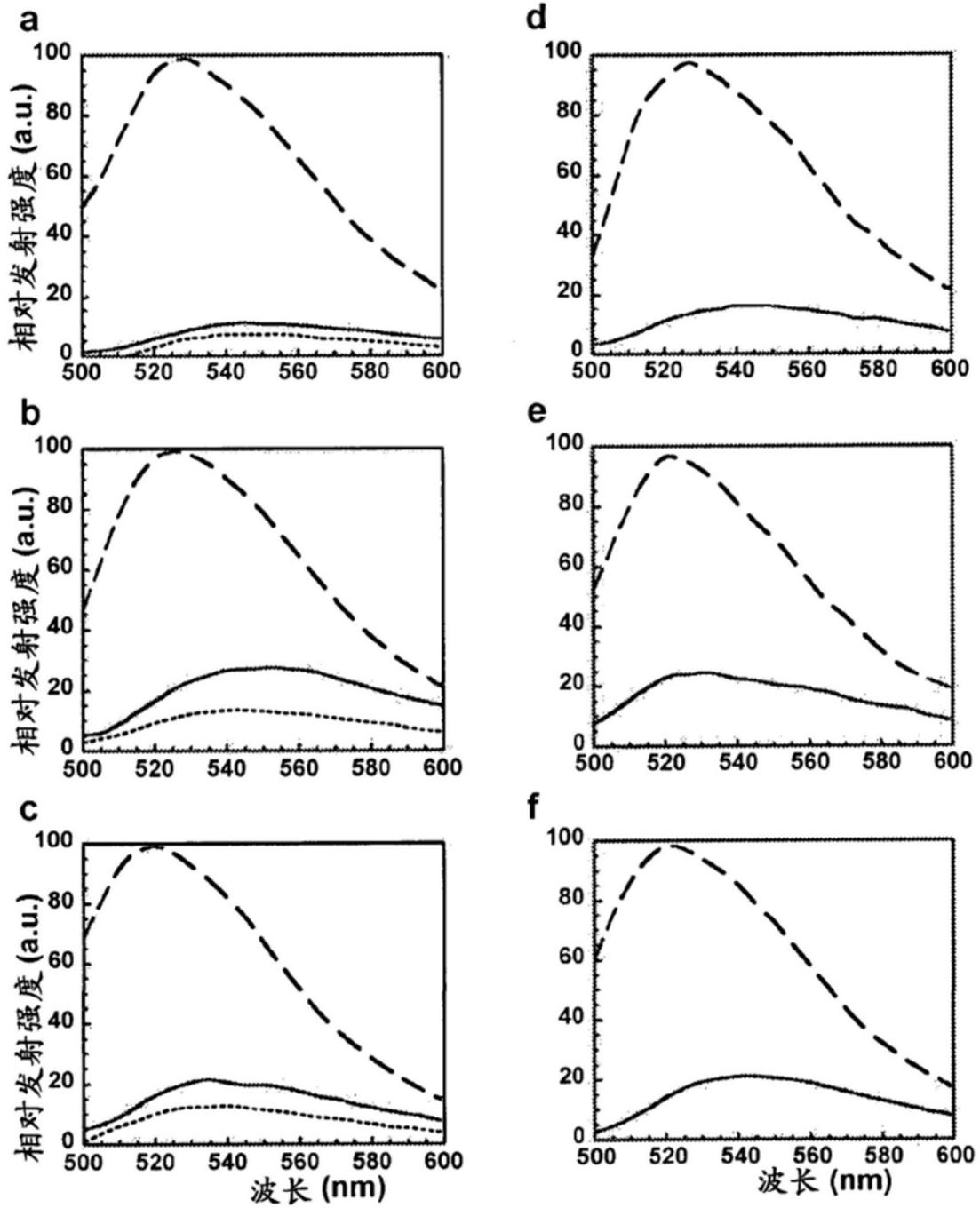


图4

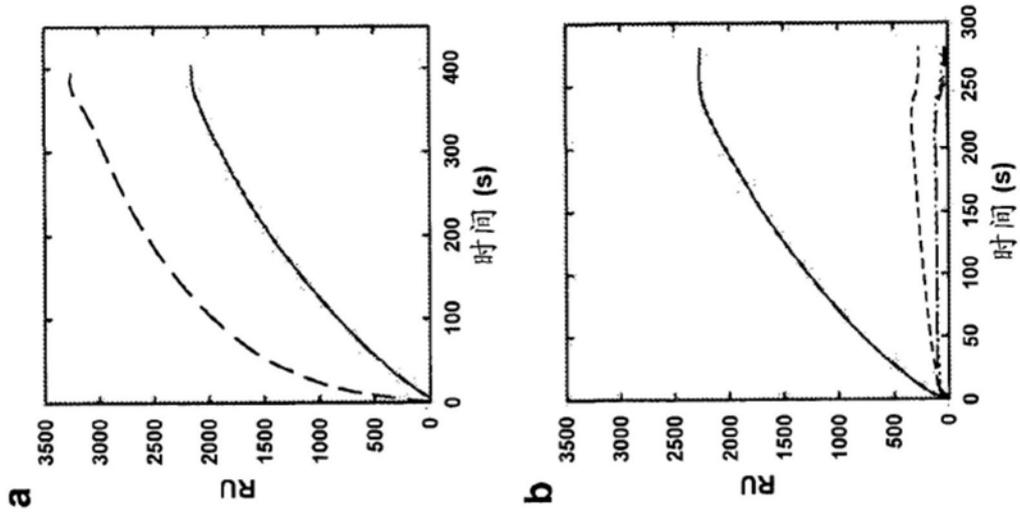


图5

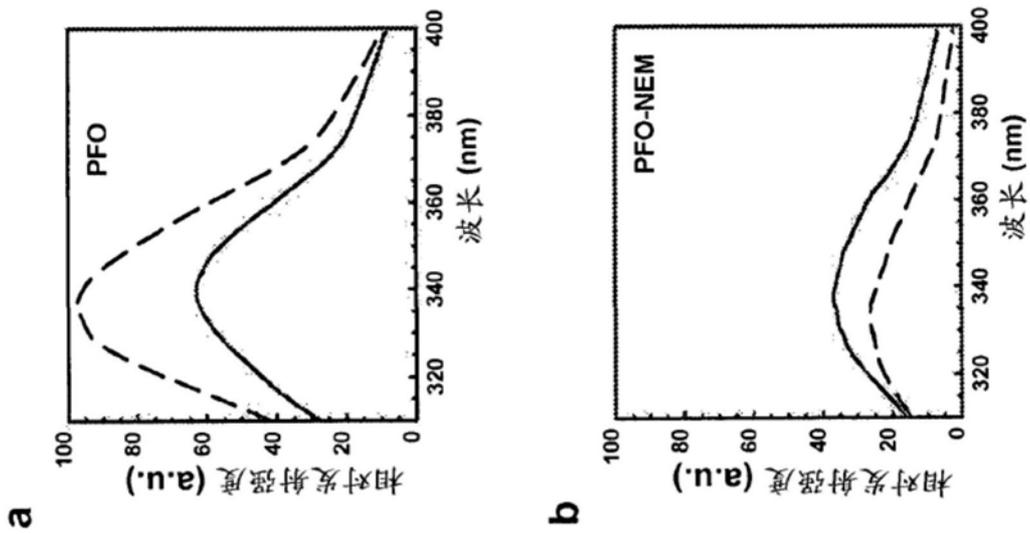


图6

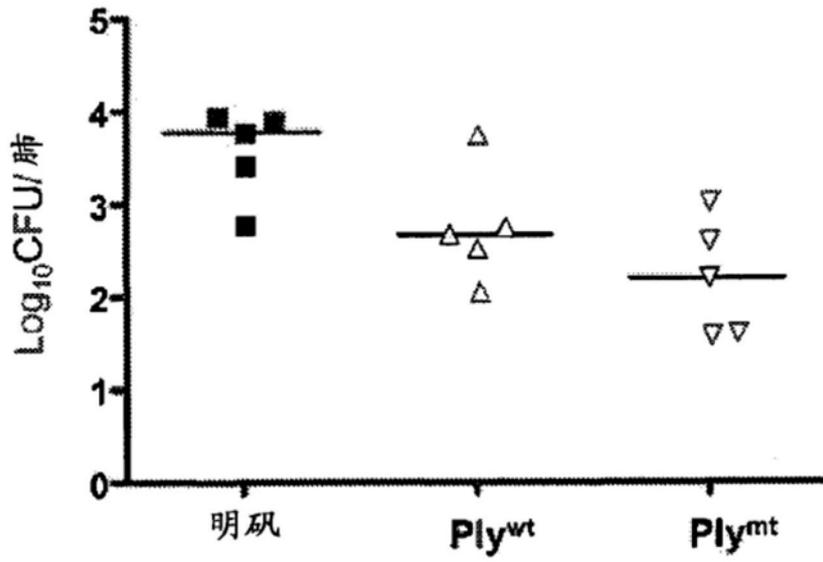


图7