



(12) 发明专利申请

(10) 申请公布号 CN 102834716 A

(43) 申请公布日 2012. 12. 19

(21) 申请号 201180018451. 2

(51) Int. Cl.

(22) 申请日 2011. 02. 23

G01N 33/15(2006. 01)

(30) 优先权数据

G01N 33/48(2006. 01)

61/307, 441 2010. 02. 23 US

C12Q 1/68(2006. 01)

61/375, 707 2010. 08. 20 US

G07H 21/04(2006. 01)

(85) PCT申请进入国家阶段日

2012. 10. 11

(86) PCT申请的申请数据

PCT/US2011/025960 2011. 02. 23

(87) PCT申请的公布数据

W02011/106456 EN 2011. 09. 01

(71) 申请人 华盛顿大学

地址 美国华盛顿

(72) 发明人 J·H·古德拉彻 I·M·德灵顿

K·W·朗福德

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 罗菊华

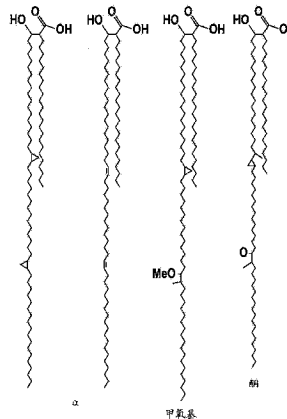
权利要求书 2 页 说明书 29 页 附图 6 页

(54) 发明名称

人工分枝菌酸膜

(57) 摘要

本文提供人工分枝菌酸膜。膜可为未支持的或栓系的。这些膜寿命长及高度抗电穿孔,展示它们的普遍强度。分枝菌酸膜适合于分枝杆菌外膜的控制的研究和可在其他实验,诸如纳米孔分析物转位实验中使用。



1. 人工膜,其包含分枝菌酸。
2. 权利要求 1 的膜,其包含多种分枝菌酸。
3. 权利要求 1 的膜,其还被定义为未支持的膜。
4. 权利要求 1 的膜,其还被定义为栓系的膜。
5. 权利要求 1 的膜,其中分枝菌酸还被定义为修饰的分枝菌酸。
6. 权利要求 5 的膜,其中修饰的分枝菌酸是交联的分枝菌酸。
7. 权利要求 1 的膜,其具有平均厚度约 5 ~ 约 22nm。
8. 权利要求 1 的膜,其在用 10mM HEPES 缓冲到 pH8.0 \pm 0.05 的去离子水中的 1.0M KCl 溶液存在下制备,其当施加穿过膜的电压以约 100mV/s 渐变时具有约 2.0V 的平均破裂电压。
9. 权利要求 1 的膜,其在用 10mM HEPES 缓冲到 pH8.0 \pm 0.05 的用去离子水制备的 1.0M KCl 溶液存在下具有抵御大于 1V 的电压大于 2 小时的能力。
10. 权利要求 1 的膜,其当除去顺式或反式侧上的缓冲剂时具有对破裂的抗性。
11. 权利要求 1 的膜,其还包含纳米孔。
12. 权利要求 11 的膜,其还包含多个纳米孔。
13. 权利要求 11 的膜,其中纳米孔是蛋白孔。
14. 权利要求 13 的膜,其中蛋白孔还被定义为 α -溶血素或其变体,耻垢分枝杆菌 (*Mycobacterium smegmatis*) 孔蛋白 (Msp) 孔蛋白,或 OmpATb。
15. 权利要求 14 的膜,其中蛋白孔还被定义为 α -溶血素或其变体。
16. 权利要求 13 的膜,其中纳米孔是突变 MspA 孔蛋白。
17. 权利要求 16 的膜,其中突变 MspA 孔蛋白的第 90,91 和 93 位的氨基酸各用天冬酰胺取代。
18. 权利要求 1 ~ 17 之任一项的膜,其还包含药物。
19. 权利要求 18 的膜,其中药物是结核治疗用抗菌性药物。
20. 权利要求 1 ~ 17 之任一项的膜,其还包含酶,分子发动机,纳米粒子,光珠,磁珠,或光。
21. 权利要求 1 的膜,其还被定义为双层膜。
22. 人工膜,其由多种分枝菌酸组成。
23. 人工膜,其由多种分枝菌酸和其他脂质的混合物组成。
24. 人工膜,其基本上由多种分枝菌酸和纳米孔组成。
25. 人工膜,其基本上由多种分枝菌酸,其他脂质的混合物,及纳米孔组成。
26. 权利要求 22 ~ 25 之任一项的膜,其还被定义为未支持的膜。
27. 权利要求 22 ~ 25 之任一项的膜,其还被定义为栓系的膜。
28. 系统,其包含人工膜,所述人工膜包含在第 1 液体导电介质和第 2 液体导电介质之间定位的分枝菌酸。
29. 权利要求 28 的系统,其中至少一个液体导电介质包含分析物。
30. 权利要求 29 的系统,其中分析物是核酸或蛋白。
31. 权利要求 30 的系统,其中膜还包含纳米孔。
32. 权利要求 31 的系统,其中纳米孔还被定义为插入膜后形成隧道的蛋白。

33. 权利要求 32 的系统,其中蛋白是突变耻垢分枝杆菌(*Mycobacterium smegmatis*)孔蛋白 A 孔蛋白(突变 MspA 孔蛋白)。

34. 权利要求 33 的系统,其中突变 MspA 孔蛋白的第 90,91 和 93 位的氨基酸各用天冬酰胺取代。

35. 权利要求 28 的系统,其中至少一个液体导电介质包含分析物,其中膜还包含纳米孔,且其中系统用于检测分析物的性质。

36. 权利要求 28 的系统,其中至少一个液体导电介质包含分析物,其中膜还包含具有隧道的蛋白孔,且其中系统用于通过隧道电泳转位分析物。

37. 权利要求 28 的系统,其还包含膜片钳放大器,数据获取装置,一个或更多与第 1 液体导电介质或第 2 液体导电介质,或其任何组合交通的温度调控装置。

38. 权利要求 28 ~ 37 之任一项的系统,其中膜还被定义为未支持的膜。

39. 权利要求 28 ~ 37 之任一项的系统,其中膜还被定义为栓系的膜。

40. 方法,其包括向权利要求 28 ~ 37 之任一项的系统施加电场。

41. 权利要求 40 的方法,其还包括以包括测量随着分析物与纳米孔的开口相互作用的离子电流的方法检测系统中的分析物以提供电流图案,其中电流图案中中断的呈现指示分析物的存在。

42. 权利要求 41 的方法,其中施加电场足以导致分析物通过开口电泳转位。

43. 权利要求 42 的方法,其还包括鉴定分析物。

44. 权利要求 43 的方法,其中鉴定分析物包括比较电流图案与已知的分析物的已知的电流图案。

45. 制造包含分枝菌酸的人工未支持的膜的方法,其包括:

(a) 用一个或更多分枝菌酸-己烷混合物的包被预处理直径约 500nm ~ 约 500 μ m 的孔及移出己烷以提供干分枝菌酸;

(b) 将烃溶剂应用于干分枝菌酸,之后是加热以促进烃溶剂掺入以提供分枝菌酸-烃溶剂组合物;

(c) 在第 1 液体导电介质和第 2 液体导电介质之间放置孔;

(d) 将分枝菌酸-烃溶剂组合物应用于孔,同时监控离子电流通过孔直到孔电阻增加到 $1T\Omega$ 以上,之后是迫使液体导电介质之一从反式侧通过孔以根据需要消除离子电流阻断;以及

(e) 在孔上放置气泡,之后是气泡收缩,其中由孔电阻增加到 $1T\Omega$ 以上指示膜形成,且其中如果纳米孔可在膜内形成,则指示双层膜形成。

46. 权利要求 45 的方法,其中烃溶剂是十六碳烷。

47. 权利要求 45 的方法,其中烃溶剂是十六碳烯。

人工分枝菌酸膜

【0001】 【相关申请的交叉引用】

【0002】 本申请要求 2010 年 2 月 23 日提交的美国临时申请系列 No. 61/307, 441, 和 2010 年 8 月 20 日提交的美国临时申请系列 No. 61/375, 707 的权利, 各通过引用以其整体并入本文。

【0003】 【政府许可证权利的声明】

【0004】 本发明在由美国健康学会颁发的资金号 R01HG005115 和 5R21HG004145 的政府支持下完成。政府在本发明中具有特定权利。

【背景技术】

【0005】 分枝杆菌属(*Mycobacterium*), 包括结核分枝杆菌(*Mycobacterium tuberculosis*), 开发了抵抗当代多-药物治疗方案的株。随着由结核分枝杆菌(*M. tuberculosis*) 的感染导致的几乎 2 百万每年死亡及随着由麻风分枝杆菌(*M. leprae*) 感染衰弱的多于 200, 000 人, 有对明白分枝杆菌复原的机理的伴随的需求。这些疾病的存留和致死部分是由于不通透的分枝杆菌属(*Mycobacterium*) 细胞壁。相比大肠埃希氏菌(*Escherichia coli*) 的细胞壁, 分枝杆菌属(*Mycobacterium*) 的独特~ 8nm 厚外部细胞箱具有远远更低的对亲水剂的通透性, 且在分枝杆菌属(*Mycobacterium*) 的药物和环境抗性中是关键因素。

【0006】 尽管含有其他成分, 分枝杆菌外膜含有 30%~ 40% 分枝菌酸。分枝菌酸含有具有不等的长度的 2 个疏水尾的羧酸头基。见图 1 例示分枝菌酸。在体内, 分枝菌酸通过羧酸基团与肽聚糖或海藻糖共价连接。分枝杆菌膜的显著不通透性导致对亲水溶质通路的需求。此通路由蛋白孔介导。

【0007】 对与结核分枝杆菌(*M. tuberculosis*) 密切关联的耻垢分枝杆菌(*M. smegmatis*) 的分枝杆菌细胞-壁中的孔隙蛋白的体内研究导致外膜孔隙耻垢分枝杆菌(*M. smegmatis*) 孔蛋白 A (*MspA*) 的发现。在耻垢分枝杆菌(*M. smegmatis*) 中, *MspA* 是最丰富的蛋白和形成用于亲水营养的主要通路, 以横过外膜。已在分枝杆菌属(*Mycobacterium*) 物种中分离 *OmpATb*, 另一蛋白孔, 及离子转运子, 但在它们的天然的环境中它们的行为尚待探索。

【0008】 为了研究孔隙的各种性质, 该孔隙常常被嵌入膜中。有对开发用于涉及分枝菌酸膜的这些和其他实验的适合的膜的需求。

【0009】 【发明概述】

【0010】 本文提供包含分枝菌酸的人工膜, 包含该膜的系统, 及制造及使用该膜的方法。因此, 一些实施方式提供包含分枝菌酸的人工膜。一些实施方式提供由多种分枝菌酸组成的人工膜。一些实施方式提供由多种分枝菌酸和其他脂质的混合物组成的人工膜。其他实施方式提供基本上由多种分枝菌酸和纳米孔组成的人工膜。其他实施方式提供基本上由多种分枝菌酸, 其他脂质的混合物, 及纳米孔组成的人工膜。

【0011】 还提供的是包含人工膜的系统, 所述人工膜包含在第 1 液体导电介质和第 2 液体导电介质之间定位的分枝菌酸。也提供的是包括向系统施加电场的方法。

[0012] 也提供制备包含分枝菌酸的人工未支持的分支菌膜的方法。一实施方式提供制造包含分枝菌酸的人工未支持的膜的方法,其包括:(a)用一个或更多分枝菌酸-己烷混合物的包被预处理直径约 500nm ~ 约 500 μm 的孔及移出己烷以提供干分枝菌酸;(b)将烃溶剂应用于干分枝菌酸,之后是加热以促进烃溶剂掺入以提供分枝菌酸-烃溶剂组合物;(c)在第 1 液体导电介质和第 2 液体导电介质之间放置孔;(d)将分枝菌酸-烃溶剂组合物应用于孔,同时监控离子电流通过孔直到孔电阻增加到 $1\text{T}\Omega$ 以上,之后是迫使液体导电介质之一从反式侧通过孔以根据需要消除离子电流阻断;及(e)在孔上放置气泡,之后是气泡收缩,其中由孔电阻增加到 $1\text{T}\Omega$ 以上指示膜形成,且其中如果纳米孔可在膜内形成,则指示双层膜形成。

【附图说明】

[0013] 本发明的上述方面和许多伴随优势会变得更加容易被同意,如通过参考以下详述结合附图变得更佳明白。

[0014] 图 1 显示存在于分枝杆菌外膜的例示分枝菌酸的化学结构。见 *Annu Rev Biochem* 64:29 (1995)。

[0015] 图 2 显示在添加 MspA 之后随着离散电流步骤而跨分枝菌酸(MA)膜的导电性的变化。观察的电流步骤与在 DPhPC 膜中观察到的电流水平无法区别。

[0016] 图 3 显示 MA 膜(N=330)和 DPhPC 膜(N=205)的破裂电压的计量的直方图,及包括来自几个不同 ~ 20 μm 孔的数据。MA 膜经 MspA 蛋白的插入确认。

[0017] 图 4 提供 MA 和 DPhPC 膜中 MspA I-V 曲线的比较。在负电压,在两个膜中的 MspA 选通使开放状态电流不明了,及忽略。对于 DPhPC 膜中的 MspA, N=9 孔隙,对于 MA 中的 MspA, N=2。

[0018] 图 5 是在难以形容的膜中 MspA 的风格化的形式;双链 DNA 不能穿过孔隙的最小收缩。像不可计量。

[0019] 图 6 显示由嵌入 MA (红色)或 DPhPC (蓝色)膜的 MspA 孔隙中暂时把持的同聚物腺嘌呤发卡尾导致的离子电流阻断水平。电流表达为在给定电压下开放状态电流的级分。在 200mV 以上的电压, DPhPC 膜变得对于扩展实验太脆性,而 MA 膜允许在更高得多的电压测量。记录到 400mV 以上的事件的持续时间对于安心地提取特征性离子电流而言太短。

【发明详述】

[0021] 本文提供包含分枝菌酸的人工膜。这些膜寿命长且高度抗电穿孔,展示它们的普遍强度。分枝菌酸膜适合于,例如,分枝杆菌外膜的控制的研究,和可用于纳米孔分析物转位实验以及以下描述的其他应用。

[0022] 因此,本文提供包含分枝菌酸的人工膜。在本文中的任何实施方式中,膜可为未支持的或栓系的。分枝菌酸可还被定义为修饰的分枝菌酸。修饰的分枝菌酸可为交联的分枝菌酸。分枝菌酸可还被定义为不修饰的分枝菌酸。在一些实施方式中,膜具有平均厚度约 5 ~ 约 22nm。在一些实施方式中,平均厚度为约,至多约,或至少约 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 或 22nm 或更大,或可来源于它们的任何范围。测量膜厚度的方法在本领域熟知和一例在以下实施例中提供。

[0023] 在一些实施方式中,膜在用 10mM HEPES 缓冲到 pH8.0 \pm 0.05 的去离子水中的 1.0M

KCl 溶液存在下制备,其当施加穿过膜的电压以约 100mV/s 渐变时具有约 2.0V 的平均破裂电压。测定平均破裂电压的一例在本文提供。在一些实施方式中,膜在用 10mM HEPES 缓冲到 pH8.0±0.05 的用去离子水制备的 1.0M KCl 溶液存在下具有抵御大于 1V 的电压大于几小时的能力。在一些实施方式中,在用 10mMHEPES 缓冲到 pH8.0±0.05 的用去离子水制备的 1.0M KCl 溶液存在下,膜具有抵御大于 1V 的电压至少约 2,3,4 或 5 或更多小时,或可来源于它们的任何范围的能力。在一些实施方式中,膜当除去顺式或反式侧上的缓冲剂时具有对破裂的抗性。当暴露于提供到其顺式侧的 pH2 ~ pH9 缓冲剂时,膜可形成及改性。在一些实施方式中,膜可在超过 55°C 的温度形成及改性。

[0024] 包含分枝菌酸的人工膜可还包含各种物质。在一些实施方式中,膜还包含纳米孔。在一些实施方式中,膜还包含多个纳米孔。任选地,纳米孔是蛋白孔。在一些实施方式中,蛋白孔还被定义为 α -溶血素或其变体,耻垢分枝杆菌(*Mycobacterium smegmatis*)孔蛋白(Msp)孔蛋白,或 OmpA/Tb。在一些实施方式中,蛋白孔还被定义为 α -溶血素或其变体。在一些实施方式中,纳米孔是突变 MspA 孔蛋白。在一些实施方式中,突变 MspA 孔蛋白的第 90,91 和 93 位的氨基酸各用天冬酰胺取代。任何膜可还包含药物。在一些实施方式中,药物是结核治疗用抗菌性药物。在一些实施方式中,膜还包含酶,分子发动机,纳米粒子,光珠,磁珠,或光。

[0025] 在一些实施方式中,膜还被定义为双层膜。

[0026] 也提供的是由多种分枝菌酸组成的人工膜。也提供的是由多种分枝菌酸和其他脂质的混合物组成的人工膜。

[0027] 一些实施方式提供基本上由多种分枝菌酸和纳米孔组成的人工膜。不实质上影响该实施方式的基本和新特征的材料或步骤包括不影响分枝菌酸的脂质性质或预防膜的形成的那些。

[0028] 一些实施方式提供基本上由多种分枝菌酸,其他脂质的混合物,及纳米孔组成的人工膜。不实质上影响该实施方式的基本和新特征的材料或步骤包括不影响分枝菌酸或其他脂质的脂质性质或预防膜形成的那些。

[0029] 在一些实施方式中,系统提供为包含人工膜,所述人工膜包含在第 1 液体导电介质和第 2 液体导电介质之间定位的分枝菌酸。至少一个液体导电介质可包含分析物,诸如核酸或蛋白。在一些实施方式中,膜还包含纳米孔。纳米孔可还被定义为插入膜后形成隧道的蛋白。在一些实施方式中,蛋白是突变耻垢分枝杆菌(*M. smegmatis*)孔蛋白 A 孔蛋白(突变 MspA 孔蛋白)。在一些实施方式中,突变 MspA 孔蛋白的第 90,91 和 93 位的氨基酸各用天冬酰胺取代。在一些实施方式中,至少一个液体导电介质包含分析物,其中膜还包含纳米孔,且其中系统用于检测分析物的性质。分析物的性质可为电,化学或物理性质。在一些实施方式中,至少一个液体导电介质包含分析物,其中膜还包含具有隧道的蛋白孔,且其中系统用于通过隧道电泳转位分析物。系统可还包含膜片钳放大器,数据获取装置,一个或更多与第 1 液体导电介质或第 2 液体导电介质,或其任何组合交通的温度调控装置。

[0030] 方法可包括向本文所述的任何系统施加电场。一些方法还包括以包括测量随着分析物与纳米孔的开口相互作用的离子电流的方法检测系统中的分析物以提供电流图案,其中电流图案中中断的呈现指示分析物的存在。在一些实施方式中,施加电场足以导致分析物通过开口电泳转位。方法可还包括鉴定分析物。鉴定分析物可包括比较电流图案与已知

的分析物的已知的电流图案。

[0031] 也提供的是制造包含分枝菌酸的人工未支持的膜的方法,其包括:(a)用一个或更多分枝菌酸-己烷混合物的包被预处理直径约 500nm~约 500 μ m 的孔及移出己烷以提供干分枝菌酸;(b)将烃溶剂应用于干分枝菌酸,之后是加热以促进烃溶剂掺入以提供分枝菌酸-烃溶剂组合物;(c)在第 1 液体导电介质和第 2 液体导电介质之间放置孔;(d)将分枝菌酸-烃溶剂组合物应用于孔,同时监控离子电流通过孔直到孔电阻增加到 1TΩ 以上,之后是迫使液体导电介质之一从反式侧通过孔以根据需要消除离子电流阻断;及(e)在孔上放置气泡,之后是气泡收缩,其中由孔电阻增加到 1TΩ 以上指示膜形成,且其中如果纳米孔可在膜内形成,则指示双层膜形成。烃溶剂可为十六碳烷或十六碳烯或可掺入膜的任何其他烃溶剂。采用的烃溶剂类型依赖于想要制备膜的温度。

[0032] 分枝菌酸是作为全部分枝杆菌属(*Mycobacterium*)的细胞包膜组分的高分子量 α 分支的,β-羟基脂肪酸。分枝菌酸含有具有不等的长度的 2 个疏水尾的羧酸头基。分枝菌酸具有基本结构 $R_2CH(OH)CHR_1COOH$,其中 R_1 是 $C_{20} \sim C_{24}$ 线性烷和 R_2 是可含有各种数的碳-碳双键,环丙烷环,甲基枝或氧官能诸如羰基,羧酸和甲氧基的 30~60 个碳原子的更复合结构。分枝菌酸的结构根据家族和物种而改变。

[0033] 在分枝杆菌细胞包膜中,分枝菌酸作为游离的脂质,诸如二分枝菌酸海藻糖(TDM)或索状因子和单分枝菌酸海藻糖(TMM)存在。它们也可酯化为阿拉伯糖半乳聚糖,肽聚糖-连接的多糖的末端五-阿拉伯糖呋喃糖基单元。在本文中,分枝菌酸可还被定义为任何这些变体。在一些实施方式中,分枝菌酸还被定义为本领域知道的可为天然存在的或合成的海藻糖-修饰的分枝菌酸。见,例如,美国专利 No. 4,307,229,4,720,456,5,006,514 和 5,049,664,各通过引用以其整体并入本文。该长链脂肪酸的存在是大大地负责分枝杆菌细胞包膜的高疏水性和非常低通透性。分枝菌酸已报道于非分枝杆菌属(*Mycobacterium*)的细菌物种,例如,棒杆菌属(*Corynebacterium*)和诺卡菌属(*Nocardia*)。因而,3 种主要类别的分枝菌酸是区别的(The Merck Index,1989),即:

[0034] (i) 白喉菌酸($C_{28} \sim C_{40}$ 酰基链长度)

[0035] (ii) 诺卡分枝菌酸($C_{40} \sim C_{60}$ 酰基链长度)及

[0036] (iii) 分枝杆菌分枝菌酸($C_{60} \sim C_{90}$ 酰基链长度)。

[0037] MA 的结构,基序和变异的详述在 Prog Lipid Res 37:143 (1998) 中提供。MA 可购买,诸如自 Sigma Aldrich,或如本领域已知制备。见例如,美国专利 No. 6,171,830,通过引用以其整体并入本文。

[0038] 分枝菌酸的定义也包括修饰的分枝菌酸。因此,膜可包含一种或更多修饰的分枝菌酸。例如,分枝菌酸可由交联分枝菌酸修饰。分枝菌酸膜可通过端基聚合或通过分枝菌酸的内部基团的交联制成更凝胶-样和稳定。可采用类似于交联二棕榈酰磷脂酰胆碱(DPhPC)或其他脂质的方法的交联方法,如本领域知道,来制备修饰的分枝菌酸。见,例如,A. Singh and J.M. Schnur, Polymerizable Phospholipids in Phospholipids Handbook, C. Cevc, ed., Marcel Dekker Inc., NY, pp233-287 (1993)。

[0039] 本文所述的膜可包含一种或更多类型的分枝菌酸(即,分枝菌酸的混合物)。在一些实施方式中,膜包含约,至少约,或至多约 1%,2%,5%,10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%,98%,99% 或更多分枝菌酸,或可来源于

它们的任何范围。在一些实施方式中,膜包含 100% 分枝菌酸。在一些实施方式中,分枝菌酸源于结核分枝杆菌(*M. tuberculosis*)。

[0040] 说到“人工”,其是指不是天然存在的而是人造的膜。

[0041] 构建包含分枝菌酸的人工膜的能力可提供检查膜中分枝菌酸脂质的设置和构型的新工具。此外,膜可允许通过见于分枝杆菌外膜的外膜孔隙,诸如 MspA 和可能 Rv1689 的药物和化学品的控制的检查。该检查可辅助改善分枝杆菌感染的治疗。

[0042] 在分枝杆菌属(*Mycobacterium*) 中的研究之外,膜可提供用于依赖于脂质膜的稳定性的纳米技术(例如,生物-纳米技术)应用的构件。这些包括下一代核酸测序及纳米孔力光谱学。在这点上,本申请通过引用以其整体并入由 Jens H. Gundlach, Ian M. Derrington, 和 Marcus D. Collins 在 2011 年 2 月 23 日提交到美国受理局的标题为“Analyte Sequencing with Nanopores”的国际申请。本申请也通过引用各以其整体合并美国临时申请系列 No. 61/098,938 和其相关的 PCT 申请, WO 2010/034018, 标题为“Msp Nanopores and Related Methods”。可随本文公开的分枝菌酸膜采用这些申请中公开的方法。

[0043] 而且,作为在纳米孔测序及分析物检测中的读数机理探索几种类型的可观察的信号。原本建议的,最直接的和最探索的读数方法依赖于通过核苷酸或占据孔隙中最窄的收缩的其他分析物的鉴定来唯一地测定的离子“阻断电流”或“共经过电流”。此方法被称为“阻断电流纳米孔测序”,或 BCNS。已在蛋白孔 α -溶血素,突变 MspA 孔蛋白,及固态纳米孔中展示核酸的阻断电流检测和表征。阻断电流检测和表征已显示提供有关在各种情景中经过,或把持进纳米孔的 DNA 的结构的信息的宿主。类似实验可用嵌入本文所述的分枝菌酸膜的该纳米孔进行。

[0044] 一般而言,“阻断”通过与噪音波动明显可区分的及通常与在孔隙的中心开口分析物分子的存在关联的离子电流的变化来证明。阻断强度会依赖于存在的分析物类型。更特别,“阻断”指称离子电流降到未阻断的电流水平的约 5~100% 的阈值以下,保持至少 1.0 μ s,及自发返回到未阻断的水平的时间间隔。例如,离子电流可降到约,至少约,或至多约 5%,10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95% 或 100%,或可来源于它们的任何范围的阈值以下。如果直接在其前或后的未阻断的信号具有自典型未阻断的水平偏离未阻断的信号的 rms 噪音大于两倍的平均电流,则阻断被拒绝。“深度阻断”鉴定为离子电流降到未阻断的水平的时间间隔。电流保持未阻断的水平的时间间隔的 80% 和 50% 之间的间隔鉴定为“部分阻断”。

[0045] 在一些实施方式中,可将通过孔隙的离子电流幅度转化为荧光光学系统,如本领域熟知。见,例如, *J Amer Chem Soc* 13:1652 (2009)。

[0046] 人工分枝菌酸膜也可用于筛选改善了通透性的药物,以明白对药物的膜不通透性,及评价膜中的已知的及未知的药物流出泵(见 *Trends Microbiol* 18:109 (2010))。膜也可用于评价膜蛋白诸如 Msp 孔蛋白(例如, MspA 和突变 MspA) 和 OmpATb 作为药物靶。

[0047] 也涵盖的是方法,其中在药物已呈递到膜的反式或顺式侧之后观察到分枝菌酸膜的穿孔。可通过监控离子电流或通过采用荧光粒子,或通过观察在膜的另一侧的药物来观察膜的穿孔。也涵盖的是适应于执行该方法的系统。

[0048] 如本文所用,“未支持的膜”是跨在沿着膜表面的任意侧上无支持物的孔的开口

的膜。膜在任意侧或两侧具有液体,气体或真空,但不与任意侧上的固体(基材)接触。

[0049] 如本文所用,“栓系的膜”是分枝菌酸的头基附接,或栓系到基材(例如,塑料,玻璃,芯片,珠)的膜。通过头基的化学修饰将脂质附接到基材以形成栓系的膜的方法为本领域熟知,及该方法可用于类似地修饰及附接分枝菌酸的头基。

[0050] “纳米孔”指称具有在其最窄点具有约 0.3nm ~ 约 2nm 的直径的开口的孔隙。例如,纳米孔可为固态纳米孔,石墨烯纳米孔,弹性体纳米孔,或可为插入双层,薄膜,膜或固态孔后形成隧道的天然存在的或重组蛋白,也在本文称之为蛋白孔或蛋白纳米孔(例如,跨膜孔隙)。如果蛋白插入膜,则蛋白是隧道-形成蛋白。测定是否蛋白是隧道-形成蛋白的方法为本领域熟知。例如,可通过测定是否蛋白插入双层来测定是否 Msp 孔蛋白形成隧道,诸如描述于美国临时申请系列 No. 61/098,938 和其相关的 PCT 申请,WO 2010/034018,各通过引用以其整体并入本文,及 Proc Natl Acad Sci 105:20647 (2008)。一般而言,隧通形成通过观察导电性的离散变化来检测。见,例如,Mo1 Microbiol33:933(1999)。开口一般处于与纳米孔的顺式和反式侧液体或气体交通。纳米孔可包含固态材料,诸如氮化硅,修饰的氮化硅,硅,氧化硅,或石墨烯或其组合(例如,纳米孔可通过制造第 1SiN 孔来制备,将石墨烯片材放到其上,然后在石墨烯中制造纳米孔)。蛋白纳米孔(也称之为蛋白孔)的非限制性例包括 α -溶血素和其变体,耻垢分枝杆菌(*Mycobacterium smegmatis*)孔蛋白(Msp)孔蛋白,及 OmpATb。

[0051] “液体介质”包括水性,有机-水性和仅有机液体介质。有机介质包括,例如,甲醇,乙醇,二甲基亚砷和其混合物。在本文所述的方法中可采用的液体为本领域熟知。该介质,包括导电液体介质的描述和例提供于美国专利 No. 7,189,503,例如,通过引用以其整体并入本文。可将盐,去污剂或缓冲剂加入该介质。可采用该剂来改变液体介质的 pH 或离子强度。粘度-改变物质,诸如甘油或各种聚合物(例如,聚乙烯吡咯烷酮,聚乙二醇,聚乙烯基醇,纤维素聚合物)和其混合物,可包括在液体介质中。

[0052] 任何实施方式中采用的第 1 和第 2 液体介质可相同或不同,及任意者或两者可包含一种或更多盐,去污剂,或缓冲剂。任选地,至少一个液体介质是导电的。任选地,至少一个液体介质不是导电的。液体介质可包含本文所述的任何分析物。

[0053] 在本文中的任何实施方式中,分析物可为核苷酸,核酸,氨基酸,肽,蛋白,聚合物,药物,离子,污染物,纳米观对象,或生物学战剂。任选地,分析物是聚合物,诸如蛋白,肽,或核酸。任选地,聚合物是核酸。核酸可为 ssDNA, dsDNA, RNA, 或其组合。本文所述的任何分析物可包含光珠或磁珠。

[0054] 如本文所用,“药物”指称可改变受试者的生物学过程的任何物质。药物可设计或用于受试者的疾病,病症,综合征或其他健康痛苦的诊断,治疗或预防。药物可为在性质上休养的,即,简单用来改变生物学过程及不用于受试者的疾病,病症,综合征或其他健康痛苦的诊断,治疗或预防。生物制品,其指称通过涉及重组 DNA 技术的生物学机理产生的物质,也被术语“药物”包括。药物包括,例如,抗细菌剂,抗炎剂,抗凝剂,抗病毒剂,抗高血压药,抗抑郁药,抗微生物剂,镇痛药,麻醉剂, β -阻断剂,双膦酸盐,化学治疗剂,造影剂,生育药品,致幻剂,激素,麻醉药,阿片剂,镇静剂,抑制素,甾体和血管扩张剂。药物的非限制性例也可见于 Merck Index。结核治疗中使用的抗细菌性药物,例如,包括异烟肼,利福平,吡嗪酰胺和乙胺丁醇。

[0055] 如本文所用,“聚合物”指称包含 2 个或多个线性单元的分子(也称之为“聚体”),其中各单元可相同或不同。聚合物的非限制性例包括核酸,肽和蛋白,以及各种烃聚合物(例如,聚乙烯,聚苯乙烯)及官能化的烃聚合物,其中聚合物主链包含碳链(例如,聚氯乙烯,聚甲基丙烯酸酯)。聚合物包括共聚物,嵌段共聚物及分支的聚合物诸如星状聚体和树状聚体。

[0056] 如本文所用,“生物学战剂”指称任何能导致植物或动物(包括人)死亡或疾病或食品或水供给降解,或环境降解的该微生物的任何生物或任何天然存在的,生物加工的或合成的组分。非限制性例包括埃博拉病毒,Marburg 病毒,炭疽芽孢杆菌(*Bacillus anthracis*)和肉毒梭菌(*Clostridium botulinum*),重型天花(*Variola major*),轻型天花(*Variola minor*),炭疽和蓖麻毒素。

[0057] 如本文所用,“污染物”指称污染空气,水或土壤的物质。污染物的非限制性例包括肥料,杀虫剂,杀昆虫剂,去污剂,石油烃,烟和含有重金属的物质,诸如含有锌,铜或汞(例如,甲基汞)的那些。

[0058] 如本文所用,“氨基酸”指称见于蛋白的 20 种天然存在的氨基酸的任何,天然存在的氨基酸的 D- 立体异构体(例如, D- 苏氨酸),非天然的氨基酸,及化学修饰的氨基酸。氨基酸的各这些类型不是相互排他的。 α 氨基酸包含氨基,羧基,氢原子,及称之为“侧链”的特征性的基团所键合的碳原子。天然存在的氨基酸的侧链为本领域熟知及包括,例如,氢(例如,如在甘氨酸),烷基(例如,如在丙氨酸,缬氨酸,亮氨酸,异亮氨酸,脯氨酸),取代的烷基(例如,如在苏氨酸,丝氨酸,甲硫氨酸,半胱氨酸,天冬氨酸,天冬酰胺,谷氨酸,谷氨酰胺,精氨酸和赖氨酸),芳基烷基(例如,如在苯丙氨酸和色氨酸),取代的芳基烷基(例如,如在酪氨酸),及杂芳基烷基(例如,如在组氨酸)。

[0059] 以下缩写用于 20 种天然存在的氨基酸:丙氨酸(Ala;A),天冬酰胺(Asn;N),天冬氨酸(Asp;D),精氨酸(Arg;R),半胱氨酸(Cys;C),谷氨酸(Glu;E),谷氨酰胺(Gln;Q),甘氨酸(Gly;G),组氨酸(His;H),异亮氨酸(Ile;I),亮氨酸(Leu;L),赖氨酸(Lys;K),甲硫氨酸(Met;M),苯丙氨酸(Phe;F),脯氨酸(Pro;P),丝氨酸(Ser;S),苏氨酸(Thr;T),色氨酸(Trp;W),酪氨酸(Tyr;Y),及缬氨酸(Val;V)。

[0060] 非天然的氨基酸(即,不天然地见于蛋白的那些)也为本领域熟知,如见于,例如, *Mol Cell Biol* 9:2574 (1989); *J Amer Chem Soc* 112:4011-4030 (1990); *J Amer Chem Soc* 56:1280-1283 (1991); *J Amer Chem Soc* 113:9276-9286 (1991); 及该文引用的全部参考文献。 β - 和 γ 氨基酸为本领域所知及也涵盖在本文为非天然的氨基酸。以下表显示在本文涵盖的非天然的氨基酸的非限制性例。

[0061] 表 1. 例示非天然的氨基酸

[0062]

缩写	氨基酸	缩写	氨基酸
Aad	2- 氨基己二酸	EtAsn	N- 乙基天冬酰胺
Baad	3- 氨基己二酸	Hyl	羟基赖氨酸

Bala	β -丙氨酸, β -氨基-丙酸	AHyl	别-羟基赖氨酸
Abu	2-氨基丁酸	3Hyp	3-羟脯氨酸
4Abu	4-氨基丁酸, 哌啶酸	4Hyp	4-羟脯氨酸
Acp	6-氨基己酸	Ide	异锁链素
Ahe	2-氨基庚酸	AIle	别-异亮氨酸
Aib	2-氨基异丁酸	MeGly	N-甲基甘氨酸, 肌氨酸
Baib	3-氨基异丁酸	MeIle	N-甲基异亮氨酸
Apm	2-氨基庚二酸	MeLys	6N-甲基赖氨酸
Dbu	2, 4-二氨基丁酸	MeVal	N-甲基缬氨酸
Des	锁链素	Nva	正缬氨酸
Dpm	2, 2' -二氨基庚二酸	Nle	正亮氨酸
Dpr	2, 3-二氨基丙酸	Orn	鸟氨酸
EtGly	N-乙基甘氨酸		

[0063] 如本文所用, “化学修饰的氨基酸”指称其侧链已化学修饰的氨基酸。例如, 可将侧链修饰为包含信号传导部分, 诸如荧光团或放射性标记。可将侧链修饰为包含新官能团, 诸如氢硫基, 羧酸, 或氨基。翻译后修饰的氨基酸也包括在化学修饰的氨基酸的定义中。

[0064] 如本文所用, “肽”指称由酰胺键(即, “肽键”)接合在一起的 2 个或更多氨基酸。肽包含或包括 50 个氨基酸。肽可为线性或环状。肽可为 α , β , γ , δ 或更高, 或混合的。肽可包含本文定义的氨基酸的任何混合物, 诸如包含 D, L, α , β , γ , δ 或更高氨基酸的任何组合。

[0065] 如本文所用, “蛋白”指称具有 51 个或更多氨基酸的氨基酸序列。

[0066] 术语“核酸”指称单链或双链形式的脱氧核糖核苷酸或核糖核苷酸聚合物, 及除非另外限制的, 包括以类似于天然存在的核苷酸的方式与核酸杂交的天然的核苷酸的已知的类似物, 诸如肽核酸(PNA)和硫代磷酸酯 DNA。除非另外指示的, 特定核酸序列包括其互补序列。核苷酸包括, 但不限于, ATP, dATP, CTP, dCTP, GTP, dGTP, UTP, TTP, dUTP, 5-甲基-CTP, 5-甲基-dCTP, ITP, dITP, 2-氨基-腺苷-TP, 2-氨基-脱氧腺苷-TP, 2-三磷酸硫代胸苷, 吡咯并-三磷酸嘧啶, 及 2-硫代胞苷, 以及以上全部的 α 硫代三磷酸酯, 及全部以上碱基的 2'-O-甲基-核糖核苷酸三磷酸酯。修饰的碱基包括, 但不限于, 5-Br-UTP, 5-Br-dUTP, 5-F-UTP, 5-F-dUTP, 5-丙炔基 dCTP, 及 5-丙炔基-dUTP。

[0067] “分子发动机”为本领域熟知且指称与分析物, 诸如聚合物(例如, 多核苷酸)物理学相互作用及能相对固定的位置, 诸如纳米孔的开口(例如, Msp 孔蛋白的隧道)物理学运

动分析物的分子(例如,酶)。尽管不旨在被理论束缚,分子发动机利用化学能量以产生机械力。在一些实施方式中,分子发动机可以连续方式与聚合物的各单元(或“聚体”)相互作用。分子发动机的非限制性例包括 DNA 聚合酶, RNA 聚合酶,解螺旋酶,核糖体和外切核酸酶。也知道非-酶促发动机,诸如包装 DNA 的病毒发动机。见 Nature 413:748(2001)。各种分子发动机和该发动机的期望性质描述于美国专利 No. 7, 238, 485, 通过引用以其整体并入本文。分子发动机可布置在膜的顺式侧或反式侧和可任选地固定,诸如描述于' 485 专利。将分子发动机合并入纳米孔的方法可使用,例如,描述于' 485 专利的方法进行。关于包含本文也所述的纳米孔的膜可采用描述于' 485 专利的系统和设备。分子发动机也讨论于,例如, J Amer Chem Soc 130:818 (2008); Nature Nanotech 2:718 (2007); 及 ACS Nano 3:1457(2009)。也可在本文所述的纳米孔和膜的情景中采用如描述于 WO 2010/034018, 通过引用以其整体并入本文的分子发动机。

[0068] 可采用的珠包括磁珠和光珠。例如,可使用链霉亲和素-包被的磁珠来施加拉 DNA 通过纳米孔的开口的静电力的相反的力。在此后来的技术中,磁珠附接于生物素化的 DNA, 及会使用强磁场梯度施加与静电驱动力可比较的力($\sim 10\text{pN}$)。见 Biophys J 82:3314 (2002)。以此方式,阻断-电流读数会未受影响,但对 DNA 的力可独立地控制。几十或几百的各 DNA 的完全,独立读数可然后关联及组装以重构精确的 DNA 序列。在一些实施方式中,珠可用于可视化膜的位置或指示膜破裂了。后者在不有用或可能测量离子电流的情况中 useful。

[0069] 如本文所用,“纳米粒子”指称具有一个或更多 100nm 或更小量级的尺度的粒子。“纳米观对象”是在 2 个其尺度小于 100nm 的对象。

[0070] 如本文所用,“膜的顺式侧”指称放置任何分析物的膜侧,其中分析物任选地转位。如果无分析物检查,经灌注可接近的膜侧被认为是膜的顺式侧。如果膜的两侧是经灌注等同可接近的,则顺式侧可被定义为操作子。

[0071] 如本文所用,“膜的反式侧”指称与膜的顺式侧相反的膜侧。

[0072] 分枝菌酸膜可包含非分枝菌酸的脂质。脂质是本领域中已知的一类分子且含有疏水尾和亲水头基。见,例如,美国专利 No. 7, 514, 267, 通过引用以其整体并入本文。在一些实施方式中,脂质是 3~28 个碳的链长度及具有 0~6 个不饱和的键的饱和的或不饱和的脂肪酸。脂质可具有 2 个烃链,一般是酰基链,及头基,极性或非极性的。有各种合成和天然存在的脂质,包括磷脂,诸如磷脂酰胆碱,磷脂酰乙醇胺,磷脂酸,磷脂酰肌醇和鞘磷脂,其中 2 个烃链一般长度在约 14~22 个碳原子之间,及具有变化程度的不饱和。

[0073] 可包含在分枝菌酸膜的磷脂包括天然或合成磷脂。非限制性例包括磷脂酰胆碱(PC),磷脂酰乙醇胺(PE),磷脂酰肌醇(PI),磷脂酰甘油(PG),磷脂酸(PA),磷脂酰丝氨酸(PS),及鞘磷脂(SM)。磷脂中的脂肪酰基链通常是至少约 7 个碳原子长,一般 12~20 个碳长,及可完全饱和或部分不饱和的。磷脂的其他例包括磷脂酰胆碱,诸如二棕榈酰磷脂酰胆碱(DPPC 或 DPhPC),二月桂基磷脂酰胆碱(DLPC) C12:0,二肉豆蔻酰磷脂酰胆碱(DMPC) C14:0,二硬脂酰磷脂酰胆碱(DSPC),二植烷酰磷脂酰胆碱,十九烷酰磷脂酰胆碱,花生酰磷脂酰胆碱,二油酰磷脂酰胆碱(DOPC) (C18:1),二棕榈油酰磷脂酰胆碱(C16:1),亚油酰磷脂酰胆碱(C18:2),二棕榈酰磷脂酰乙醇胺,二油酰磷脂酰乙醇胺(DOPE),二油酰磷脂酰甘油(DOPG),棕榈酰油酰磷脂酰甘油(POPG),二硬脂酰磷脂酰丝氨酸(DSPS) 大豆卵磷脂,卵

黄卵磷脂,鞘磷脂,磷脂酰丝氨酸,磷脂酰甘油,磷脂酰肌醇,二磷脂酰甘油,磷脂酰乙醇胺和磷脂酸。

[0074] 可使用的其他脂质包括 1, 2- 二酰基 -sn- 甘油基 -3-[磷酸 - 消旋 -(1- 甘油)], 1, 2- 二酰基 -sn- 甘油基 -3-[磷酸 -L- 丝氨酸], 1, 2- 二酰基 -sn- 甘油基 -3- 磷酸胆碱, 1, 2- 二酰基 -sn- 甘油基 -3- 磷酸和 1, 2- 二酰基 -sn- 甘油基 -3- 磷酸乙醇胺, 其中二酰基可为对称或不对称的, 且含有链长度 3 ~ 28 个碳及具有达 6 个不饱和的键的各种类型的饱和的或不饱和的脂肪酸。其他脂质包括卵磷脂酰胆碱(EPC), 1, 2- 二肉豆蔻酰 -sn- 甘油基 -3- 磷酸胆碱(DMPC), 及 1, 2- 二肉豆蔻酰 -sn- 甘油基 -3-[磷酸 - 消旋 -(1- 甘油)]。

[0075] 在一些实施方式中, 非 - 磷脂, 中性脂质, 糖脂, 胆甾醇, 甾醇, 甾等包括在膜中。在一些实施方式中, 使用阴离子脂质。阴离子脂质的例包括磷脂酸(PA), 磷脂酰丝氨酸(PS) 和磷脂酰甘油(PG), 磷脂酰胆碱(PC), 1, 2- 二肉豆蔻酰 -sn- 甘油基 -3-[磷酸 - 消旋 -(1- 甘油)] (DMPG)。阳离子脂质也可使用, 在一些实施方式中。该阳离子脂质一般具有亲脂性部分, 诸如甾醇, 酰基或二酰基链, 且其中脂质具有总体净正电荷。脂质的头基可携带正电荷。例示阳离子脂质包括 1, 2- 二油基氧基 -3-(三甲基氨基) 丙烷(DOTAP); N-[1-(2, 3, - 二 - 十四烷基氧基) 丙基]-N, N- 二甲基 -N- 羟乙基铵溴化物(DMRIE); N-[1-(2, 3, - 二油基氧基) 丙基]-N, N- 二甲基 -N- 羟基乙基铵溴化物(DORIE); N-[1-(2, 3- 二油基氧基) 丙基]-N, N, N- 三甲基铵氯化物(DOTMA); 3[N-(N', N' - 二甲基氨基乙烷) 氨基甲酰基] 胆甾醇(DC-Chol); 及二甲基二 - 十八烷基铵(DDAB)。脂质也可作为中性脂质, 诸如二油酰磷脂酰乙醇胺(DOPE)或两亲性脂质, 诸如磷脂, 用阳离子脂质衍生的, 诸如聚赖氨酸或其他聚胺脂质。

[0076] 在一些实施方式中, 选择脂质以达到膜的特定的程度的流动性或刚性, 以控制膜的稳定性, 或以控制膜内截留的剂(例如, 分析物)的释放速度。例如, 饱和的脂质可在脂质双层中贡献更大膜刚性。其他脂质组分, 诸如胆甾醇, 也已知在脂质双层结构贡献于膜刚性。

[0077] 如本文所用, “转位”和语法变体是指进入纳米孔的开口的一侧及移动到及移动出开口的另一侧。特别涵盖包含转位的本文的任何实施方式可指称电泳转位或非 - 电泳转位, 除非特别提及。电场可移动分析物使得其与开口相互作用。说到“相互作用”, 其是指分析物移动入及, 任选地, 通过开口, 其中“通过开口”(或“转位”)是指进入开口的一侧及移动到及移动出开口的另一侧。任选地, 涵盖不采用电泳转位的方法, 诸如物理压力或磁压力, 当采用磁珠时。

[0078] “耻垢分枝杆菌(*Mycobacterium smegmatis*)孔蛋白(Msp)”或“Msp孔蛋白”指称包含 2 个或更多 Msp 单体的多聚体复合物。Msp 单体由耻垢分枝杆菌(*M. smegmatis*)中的基因编码。耻垢分枝杆菌(*M. smegmatis*)具有 4 种鉴定的 Msp 基因, 表示为 MspA, MspB, MspC 和 MspD。Msp 孔蛋白可, 例如, 包含野生型 MspA 单体, 突变 MspA 单体, 野生型 MspA 旁系同原物或同源物单体, 或突变 MspA 旁系同原物或同源物单体。任选地, Msp 孔蛋白是单链 Msp 孔蛋白或为几种单链 Msp 孔蛋白的多聚体。单链 Msp 孔蛋白可, 例如包含由 2 个或更多由一种或更多氨基酸接头肽连接的 Msp 单体(例如, 8 单体)形成的多聚体。部分单链 Msp 孔蛋白指称必需二聚化, 三聚化等而形成孔蛋白的单链多聚体复合物。全单链 Msp 孔蛋白指称无需二聚化, 三聚化等而形成孔蛋白的形成孔蛋白的单链多聚体复合物。Msp 孔蛋白

为本领域所知,是制造突变 Msp 孔蛋白的方法。国际申请 W02010/034018,通过引用以其整体并入本文,描述许多这些孔蛋白和制造这些孔蛋白的方法。

[0079] “前庭”指称 Msp 孔蛋白内部的锥-形部分,其直径通常从一端到另一端沿着中心轴减小,其中前庭的最窄的部分连接于收缩区域。前庭也可被称为“杯状体”。见 W0 2010/034018 的图 1 的野生型 MspA 孔蛋白的前庭的一例。前庭和收缩区域一起定义 Msp 孔蛋白的隧道。

[0080] 当说到 Msp 孔蛋白的前庭直径时,需知,因为前庭是锥-样形状,直径沿着中心轴的路径变化,其中直径是一端相比对置的端更大。直径可在约 2nm ~ 约 6nm 范围。任选地,直径为约,至少约,或至多约 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 或 6.0nm, 或可来源于它们的任何范围。中心轴的长度可在约 2nm ~ 约 6nm 范围。任选地,长度为约,至少约,或至多约 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 或 6.0nm, 或可来源于它们的任何范围。当在本文指称“直径”时,可通过测量中心-到-中心距离或原子表面-到-表面距离来测定直径。

[0081] “收缩区域”指称 Msp 孔蛋白的隧道的最窄的部分,从连接于前庭的直径而言。野生型 MspA 孔蛋白的收缩区域显示于 W0 2010/034018 的图 1(标记的“内部收缩”)。收缩区域的长度可在约 0.3nm ~ 约 2nm 范围。任选地,长度为约,至多约,或至少约 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2 或 3nm, 或可来源于它们的任何范围。收缩区域的直径可在约 0.3nm ~ 约 2nm 范围。任选地,直径为约,至多约,或至少约 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2 或 3nm, 或可来源于它们的任何范围。

[0082] “中性收缩区域”指称当浸渍到水溶液时包含累积呈现无净电荷的氨基酸侧链的收缩区域。与收缩区域接触的液体介质(例如,缓冲的水溶液)的 pH 可影响是否收缩区域表征为中性或非中性。

[0083] “隧道”指称气体,液体,离子或分析物可经过的被定义为前庭和收缩区域的 Msp 孔蛋白的中心,空部分。隧道是纳米孔的开口的一例。

[0084] “突变 MspA 孔蛋白”是与其对应野生型 MspA 孔蛋白具有至少或至多 70, 75, 80, 85, 90, 95, 98 或 99% 或更高同一性,或可来源于它们的任何范围,但小于 100%, 及保留隧道-形成能力的多聚体复合物。突变 MspA 孔蛋白可为重组蛋白。任选地,突变 MspA 孔蛋白是在野生型 MspA 孔蛋白的收缩区域或前庭内具有突变的蛋白。任选地,突变可在野生型 MspA 孔蛋白的边缘或周质环之外发生。突变 MspA 孔蛋白可在本文所述的任何实施方式中采用。

[0085] 尤其是说到 MspA 孔蛋白,任选地,MspA 孔蛋白是由 8 个 184-氨基酸 MspA 单体组成的八聚体。一个或更多突变可在野生型 MspA 孔蛋白的一个或更多氨基酸 MspA 单体中发生,以产生突变 MspA 孔蛋白。此外, MspA 孔蛋白可具有少于或多于 8 个单体,其中任何一个或更多可包含突变。

[0086] 野生型 MspA 孔蛋白包含由 13 个氨基酸组成及直接相邻于收缩区域的周质环。见

J Biol Chem 284:10223(2009)。野生型 MspB, C 和 D 孔蛋白也含有周质环。一个或更多突变可在野生型 Msp 孔蛋白的周质环中发生以产生突变 Msp 孔蛋白。例如, 达全部 13 个氨基酸的缺失可在野生型 MspA 孔蛋白的周质环中发生。一般而言, 周质环中的缺失不影响 Msp 孔蛋白的隧道 - 形成能力。

[0087] Msp 孔蛋白或 Msp 单体也可为化学或生物学修饰的。例如, 可用化学品修饰 Msp 孔蛋白或 Msp 单体以产生二硫桥, 如本领域技术人员知道。

[0088] Msp 孔蛋白可包含核苷酸结合位点。如本文所用, “核苷酸结合位点” 指称 Msp 孔蛋白中的位点其中核苷酸与氨基酸保持接触, 或处于与氨基酸接触长于可归因于扩散移动的时间, 诸如大于 1ps 或 1ns。可采用分子动力学计算来评定这些临时静止时间。

[0089] Msp 孔蛋白中的一个或更多突变可在蛋白的前庭或收缩区域发生。任选地, 突变 Msp 孔蛋白相比野生型 Msp 孔蛋白, 在其周质环, 前庭或收缩区域氨基酸序列中具有至少一个差异(例如, 缺失, 取代, 添加)。其他任选的突变在本文所述。

[0090] 本文的任何实施方式的 Msp 孔蛋白可为本文所述的任何 Msp 孔蛋白, 诸如野生型 MspA 孔蛋白, 突变 MspA 孔蛋白, 野生型 MspA 旁系同源物或同源物孔蛋白, 或突变 MspA 旁系同源物或同源物孔蛋白。Msp 孔蛋白可由编码单链 Msp 孔蛋白的核酸序列编码。这里的任何 Msp 孔蛋白可包含本文所述的任何 Msp 单体, 诸如突变 Msp 单体。

[0091] 营养经过分枝杆菌属 (*Mycobacterium*) 的野生型孔蛋白。野生型 MspA 孔蛋白, 野生型 MspB 孔蛋白, 野生型 MspC 孔蛋白, 及野生型 MspD 孔蛋白是野生型隧道 - 形成孔蛋白的例。Msp 孔蛋白可还被定义为本文所述的任何 Msp 孔蛋白, 包括旁系同源物, 同源物, 突变体和单链孔蛋白。

[0092] 例示野生型 MspA 旁系同源物和同源物提供于表 2。提供的是野生型 MspA 旁系同源物, 其包括野生型 MspB, 野生型 MspC 和野生型 MspD。本文定义的“旁系同源物”是自具有类似结构和功能的相同的细菌物种的基因。本文定义的“同源物”是自具有类似结构和进化来源的另一细菌物种的基因。作为一例, 提供的是野生型 MspA 同源物, 其包括 MppA, PorM1, PorM2, PorM1 和 Mmcs4296。

[0093] 表 2. 例示野生型 MspA 和野生型 MspA 旁系同源物和同源物单体

[0094]

蛋白#	生物	与 MspA 的同一性/相似性(%)	长度(aa)	参考文献
MspA/Msmeg0965	耻垢分枝杆菌 (<i>M. smegmatis</i>)	100/100	211	gb ABK74363.1 , (Stahl 等人, 2001)*
MspB/Msmeg0520	耻垢分枝杆菌 (<i>M. smegmatis</i>)	94/95	215	gb ABK73437.1 , (Stahl 等人, 2001)*
MspC/Msmeg5483	耻垢分枝杆菌 (<i>M. smegmatis</i>)	93/95	215	gb ABK74976.1 , (Stahl 等人, 2001)*
MspD/Msmeg6057	耻垢分枝杆菌 (<i>M. smegmatis</i>)	82/89	207	gb ABK72453.1 , (Stahl 等人, 2001)*
MppA	草分枝杆菌 (<i>M. phlei</i>)	100/100	211	AJ812030, (Dorner 等人, 2004)**
PorM1	偶发分枝杆菌 (<i>M. fortuitum</i>)	95/96	211	emb CAI54228.1
PorM2	偶发分枝杆菌 (<i>M. fortuitum</i>)	91/93	215	emb CAL29811.1
PorM1	外来分枝杆菌 (<i>M. peregrinum</i>)	94/96	211	emb CAI54230.1
Mmcs4296	分枝杆菌属 (<i>Mycobacterium</i> sp.) MCS	85/91	216	gb ABG10401.1
Mmcs4297	分枝杆菌属 (<i>Mycobacterium</i> sp.) MCS	85/91	216	gb ABG10402.1
Mmcs3857	分枝杆菌属 (<i>Mycobacterium</i> sp.) MCS	30/44	235	gb ABG09962.1
Mmcs4382	分枝杆菌属 (<i>Mycobacterium</i> sp.) MCS	85/91	216	gb ABL93573.1
Mmcs4383	分枝杆菌属 (<i>Mycobacterium</i> sp.) MCS	85/91	216	gb ABL93574.1
Mjls3843	分枝杆菌属 (<i>Mycobacterium</i> sp.) JLS	26/40	235	gb ABN99619.1
Mjls3857	分枝杆菌属 (<i>Mycobacterium</i> sp.) JLS	26/40	235	gb ABG09962.1
Mjls3931	分枝杆菌属 (<i>Mycobacterium</i> sp.) JLS	26/40	235	gb ABL93123.1
Mjls4674	分枝杆菌属 (<i>Mycobacterium</i> sp.) JLS	85/89	216	gb ABO00440.1
Mjls4675	分枝杆菌属 (<i>Mycobacterium</i> sp.) JLS	83/89	216	gb ABO00441.1
Mjls4677	分枝杆菌属 (<i>Mycobacterium</i> sp.) JLS	84/89	216	gb ABO00443.1
Map3123c	副结核鸟分枝杆菌 (<i>M. avium</i> paratuberculosis)	24/39	220	gb AAS05671.1
Mav3943	鸟分枝杆菌 (<i>M. avium</i>)	24/39	227	gb ABK66660.1
Mvan1836	<i>M. vanbaalenii</i> PYR-1	82/88	209	gb ABM12657.1
Mvan4117	<i>M. vanbaalenii</i> PYR-1	32/43	239	gb ABM14894.1

[0095]

Mvan4839	<i>M. vanbaalenii</i> PYR-1	83/88	209	gb ABM15612.1
Mvan4840	<i>M. vanbaalenii</i> PYR-1	83/89	209	gb ABM15613.1
Mvan5016	<i>M. vanbaalenii</i> PYR-1	30/41	238	gb ABM15788.1
Mvan5017	<i>M. vanbaalenii</i> PYR-1	25/35	227	gb ABM15789.1
Mvan5768	<i>M. vanbaalenii</i> PYR-1	21/32	216	gb ABM16533.1
MUL_2391	溃疡分枝杆菌 (<i>M. ulcerans</i>) Agy99	21/34	233	gb ABL04749.1
Mflv1734	淡黄分枝杆菌 (<i>M. gilvum</i>) PYR-GCK	21/32	225	gb ABP44214.1
Mflv1735	淡黄分枝杆菌 (<i>M. gilvum</i>) PYR-GCK	32/41	226	gb ABP44215.1
Mflv2295	淡黄分枝杆菌 (<i>M. gilvum</i>) PYR-GCK	25/40	250	gb ABP44773.1
Mflv1891	淡黄分枝杆菌 (<i>M. gilvum</i>) PYR-GCK	84/90	217	gb ABP44371.1
MCH4691c	龟亚科分枝杆菌 (<i>M. chelonae</i>)	70/80	223	gb ACV04474.1
MCH4689c	龟亚科分枝杆菌 (<i>M. chelonae</i>)	66/78	223	gb ACV04472.1
MCH4690c	龟亚科分枝杆菌 (<i>M. chelonae</i>)	72/81	217	gb ACV04473.1
MAB1080	脓肿分枝杆菌 (<i>M. abscessus</i>)	69/79	223	emb CAM61170.1
MAB1081	脓肿分枝杆菌 (<i>M. abscessus</i>)	68/78	222	emb CAM61171.1
MAB2800	脓肿分枝杆菌 (<i>M. abscessus</i>)	27/44	246	emb CAM62879.1
RHA1 ro08561	<i>Rhodococcus jostii</i> RHA1	34/51	233	gb ABG99605.1
n.d.	浑浊红球菌 (<i>Rhodococcus opacus</i>) B4	34/51	233	gbj BAH52196.1
RHA1 ro04074	红球菌属 (<i>Rhodococcus sp.</i>) RHA1	34/50	233	gb ABG95871.1
RHA1 ro03127	红球菌属 (<i>Rhodococcus sp.</i>) RHA1	34/50	233	gb ABG94930.1
n.d.	红串红球菌 (<i>Rhodococcus erythropolis</i>) PR4	35/50	229	gbj BAH30938.1

[0096] 仅包括在蛋白全长上具有显著氨基酸相似性的蛋白。数据由 PSI-Blast 算法 (BLOSUM62 矩阵) 使用 www.ncbi.nlm.nih.gov/blast/Blast.cgi 上的 NIH GenBank 数据库获得。

[0097] n. d. : " 未测定的 "

[0098] *Mol Microbiol 40:451 (2001)

[0099] **Biochim Biophys Acta 1667:47-55 (2004)

[0100] " 突变 MspA 旁系同源物或同源物孔蛋白 " 是与其对应野生型 MspA 旁系同源物或同源物孔蛋白具有至少或至多 70, 75, 80, 85, 90, 95, 98 或 99% 或更高同一性, 或可来源于它们的任何范围, 但小于 100%, 及保留隧道 - 形成能力的多聚体复合物。突变 MspA 旁系同源物或同源物孔蛋白可为重组蛋白。任选地, 突变 MspA 旁系同源物或同源物孔蛋白是在野生型 MspA 旁系同源物或同源物孔蛋白的收缩区域或前庭中具有突变的蛋白。任选地, 突变可在野生型 MspA 旁系同源物或同源物孔蛋白的边缘或周质环之外发生。任何突变 MspA 旁系同源物或同源物孔蛋白可在本文所述的任何实施方式中采用, 及可包含本文所述的任何突

变。

[0101] Msp 孔蛋白可包含 2 个或更多 Msp 单体。“Msp 单体”是野生型 MspA 单体,突变 MspA 单体,野生型 MspA 旁系同原物或同源物单体,或突变 MspA 旁系同原物或同源物单体,及当与一个或更多其他 Msp 单体结合时保留隧道-形成能力的蛋白单体。本文所述的任何 Msp 孔蛋白可包含一个或更多本文所述的任何 Msp 单体。任何 Msp 孔蛋白可包含,例如,2~15 个 Msp 单体,其中各单体可相同或不同。

[0102] “突变 MspA 单体”指称与野生型 MspA 单体具有至少或至多 70, 75, 80, 85, 90, 95, 98 或 99% 或更高同一性,或可来源于它们的任何范围,但小于 100%,及当与一个或更多其他 Msp 单体结合时保留隧道-形成能力的 Msp 单体。任选地,突变 MspA 单体还被定义为在贡献于完全-形成的,隧道-形成孔蛋白的前庭或收缩区域的形成的序列部分中包含突变。突变 Msp 单体可为重组蛋白,例如重组 MspA 单体可包含本文所述的任何突变。

[0103] 在本文中的任何实施方式中, Msp 单体可为野生型 MspA 旁系同原物或同源物,诸如 MspA/Msmeg0965, MspB/Msmeg0520, MspC/Msmeg5483, MspD/Msmeg6057, MppA, PorM1, PorM2, PorM1, Mmcs4296, Mmcs4297, Mmcs3857, Mmcs4382, Mmcs4383, Mjls3843, Mjls3857, Mjls3931Mjls4674, Mjls4675, Mjls4677, Map3123c, Mav3943, Mvan1836, Mvan4117, Mvan4839, Mvan4840, Mvan5016, Mvan5017, Mvan5768, MUL_2391, Mflv1734, Mflv1735, Mflv2295, Mflv1891, MCH4691c, MCH4689c, MCH4690c, MAB1080, MAB1081, MAB2800, RHA1ro08561, RHA1ro04074 和 RHA1ro03127。

[0104] “突变 MspA 旁系同原物或同源物单体”指称与野生型 MspA 旁系同原物或同源物单体具有至少或至多 70, 75, 80, 85, 90, 95, 98 或 99% 或更高同一性,或可来源于它们的任何范围,但小于 100%,及保留隧道-形成能力的 MspA 旁系同原物或同源物单体。任选地,突变 MspA 旁系同原物或同源物单体还被定义为在贡献于完全-形成的,隧道-形成孔蛋白的前庭和/或收缩区域的形成的序列部分内包含突变。突变 MspA 旁系同原物或同源物单体可为重组蛋白,例如任何突变 MspA 旁系同原物或同源物单体可任选地在本文中的任何实施方式中采用。

[0105] Msp 孔蛋白可作为 2 个或更多野生型 MspA 单体,突变 MspA 单体,野生型 MspA 旁系同原物或同源物单体,或突变 MspA 旁系同原物或同源物单体的组合表达。由此, Msp 孔蛋白可为或包含二聚体,三聚体,四聚体,五聚体,六聚体,七聚体,八聚体,九聚体,等。例如, Msp 孔蛋白可包含野生型 MspA 单体和野生型 MspB 单体的组合。Msp 孔蛋白可包含 1~15 个单体,其中各单体相同或不同。的确,本文所述的任何 Msp 孔蛋白可包含至少或至多 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 或 15 个单体,或可来源于它们的任何范围,其中各单体相同或不同。例如, Msp 孔蛋白可包含相同或不同的一个或更多突变 MspA 单体。作为另一例, Msp 孔蛋白可包含至少一个突变 MspA 单体和至少一个 MspA 旁系同原物或同源物单体。

[0106] 如上定义,单链 Msp 孔蛋白包含由一种或更多氨基酸接头肽连接的 2 个或更多 Msp 单体。包含 2 个 Msp 单体的单链 Msp 孔蛋白,其中 Msp 单体由氨基酸接头序列连接,可被称为单链 Msp 孔蛋白二聚体。包含 8 个 Msp 单体的单链 Msp 孔蛋白,其中 Msp 单体由氨基酸接头序列连接,可被称为单链 Msp 孔蛋白八聚体。单链 Msp 孔蛋白可包含 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 或更多 Msp 单体,或可来源于它们的任何范围,由氨基酸接头序列连接。任选地,单链 Msp 孔蛋白可,例如,包含 2 个或更多单链 Msp 孔蛋白二聚体,2 个或更多

单链 Msp 孔蛋白三聚体, 2 个或更多单链 Msp 孔蛋白四聚体, 2 个或更多单链 Msp 孔蛋白五聚体, 1 个或更多单链 Msp 孔蛋白六聚体, 1 个或更多单链 Msp 孔蛋白七聚体, 1 个或更多单链 Msp 孔蛋白八聚体, 或其组合。例如, 单链 Msp 孔蛋白可包含单链 Msp 孔蛋白二聚体和 2 个单链 Msp 孔蛋白三聚体。作为另一例, 单链 Msp 孔蛋白可包含单链 Msp 孔蛋白四聚体和 2 个单链 Msp 孔蛋白二聚体。

[0107] 野生型单链 Msp 孔蛋白包含野生型 Msp 单体。任选地, 单链 Msp 孔蛋白中的一个或更多突变存在于单链 Msp 孔蛋白的前庭或收缩区域。突变单链 Msp 孔蛋白, 例如, 相比野生型单链 Msp, 在氨基酸序列中具有对于周质环, 前庭或收缩区域(例如, 缺失, 取代或添加)的至少一个突变。单链的多聚体也可形成孔蛋白, 其中各单链包括 2, 3, 4, 5, 6, 7 或更多 Msp 单体。

[0108] 突变 MspA 序列的非限制性例提供于表 3。任选地, 突变 MspA 包含在氨基酸 138 的 A 到 P 取代, 在氨基酸 139 的 E 到 A 取代, 或其组合。任选地, 突变 MspA 包含在氨基酸 90 的 D 到 K 或 R 取代, 在氨基酸 91 的 D 到 N 取代, 在氨基酸 93 的 D 到 N 取代, 或其任何组合。任选地, 突变 MspA 包含在氨基酸 90 的 D 到 Q 取代, 在氨基酸 91 的 D 到 Q 取代, 在氨基酸 93 的 D 到 N 取代, 或其任何组合。任选地, 突变 MspA 包含在氨基酸 88 的 L 到 W 取代, 在氨基酸 105 的 I 到 W 取代, 在氨基酸 91 的 D 到 Q 取代, 在氨基酸 93 的 D 到 N 取代, 或其任何组合。任选地, 突变 MspA 包含在氨基酸 105 的 I 到 W 取代, 在氨基酸 108 的 N 到 W 取代, 或其组合。任选地, 突变 MspA 包含在氨基酸 118 的 D 到 R 取代, 在氨基酸 139 的 E 到 K 取代, 在氨基酸 134 的 D 到 R 取代, 或其任何组合。对于以下所列的突变 MspB 单体序列, 参照 MspB 序列是本领域知道的成熟野生型 MspB 单体序列。任选地, 突变 MspB 包含在氨基酸 90 的 D 到 K 或 R 取代, 在氨基酸 91 的 D 到 N 取代, 在氨基酸 93 的 D 到 N 取代, 或其任何组合。

[0109] 表 3 :MspA 突变

[0110]

行 1	行 2
MspA D90A	MspA T84C
MspA D91A	MspA I87C
MspA D90A/D91A	MspA D91C
MspA D90E	MspA D93C
MspA D91E	MspA A96C
MspA D90E/D91E	MspA P97C
MspA D90F	MspA G100C
MspA D91F	MspA N102C
MspA D90F/D91F	MspA P107C
MspA D90G	MspA G112C
MspA D91G	MspA V113C
MspA D90G/D91G	MspA S114C
MspA D90H	MspA D118C
MspA D91H	MspA N121C
MspA D90H/D91H	MspA E127C
MspA D90K	MspA F131C
MspA D91K	MspA D134C
MspA D90K/D91K	MspA S136C
MspA D90L	MspA A138C
MspA D91L	MspA E139C
MspA D90L/D91L	MspA G141C
MspA D90R	MspA V144C
MspA D91R	MspA H148C
MspA D90R/D91R	MspA T150C
MspA D90S	MspA A155C
MspA D91S	MspA R161C
MspA D90S/D91S	MspA R165C
MspA D90W	MspA S173C
MspA D91W	MspA T175C
MspA D90W/D91W	MspA E179C
MspA D90Y	MspA V184C
MspA D91Y	MspA N79C/D90K/D91N/P97C
MspA D90Y/D91Y	MspA K47S/D90K/D91N/P97C/D134C
MspA Q126C	MspA ΔA96-P98
MspA D90N	MspA ΔT95-F99
MspA D91N	MspA ΔI94-G100
MspA D93N	MspA ΔD93-L101
MspA D90N/D91N	MspA ΔG92-N102
MspA D90N/D91N/D93N	MspA N79R/D90N/D91N/D93N
MspA D90Q/D91N/D93N	MspA N79W/D90N/D91N/D93N
MspA D90Q/D91Q/D93N	MspA D90N/D91N/D93N/Q126R
MspA D90T/D91N/D93N	MspA D90N/D91N/D93N/T130R
MspA D90T/D91T/D93N	MspA D90N/D91N/D93N/D134R
MspA D91E	MspA D90N/D91N/D93N/Q126W
MspA D90E	MspA D90N/D91N/D93N/T130W
MspA D90E/D91E	MspA D90N/D91N/D93N/D134W
MspA D90N/D91N/D93Q	MspA D90N/D91N/D93N/D118W/D134R/E139K
MspA D90N/D91N/G92Q/D93N	MspA D90N/D91N/D93N/D118F/D134R/E139K

[0111]

MspA G1C	MspA D90N/D91N/D93N/D118H/D134R/E139K
MspA D3C	MspA D90N/D91N/D93N/D118Y/D134R/E139K
MspA E5C	MspA N79W/D90N/D91N/D93N/D118R/E139K
MspA D10C	MspA N79F/D90N/D91N/D93N/D118R/E139K
MspA D13C	MspA N79H/D90N/D91N/D93N/D118R/E139K
MspA R14C	MspA N79Y/D90N/D91N/D93N/D118R/E139K
MspA T17C	MspA D90N/D91K/D93N
MspA W21C	MspA D90N/D91R/D93N
MspA D22C	MspA D90N/D91W/D93N
MspA G27C	MspA D90N/D91W/D93N
MspA R33C	MspA D90N/D91T/D93N
MspA R38C	MspA D90N/D91L/D93N
MspA G44C	MspA D90N/D91H/D93N
MspA K47C	MspA D90N/D91S/D93N
MspA I49C	MspA D90N/D91N/D93N/D118R
MspA E57C	MspA D90N/D91N/D93N/D118R/E139R
MspA G60C	MspA D90N/D91N/D93N/D118R/E139K
MspA E63C	MspA D90N/D91N/D93N/D118R/D134R/E139K
MspA G69C	MspA D90Q/D91N/D93N/D118R/D134R/E139K
MspA S73C	MspA D90Q/D91Q/D93N/D118R/D134R/E139K
MspA L74C	MspA D90T/D91N/D93N/D118R/D134R/E139K
MspA V76C	MspA D90T/D91T/D93N/D118R/D134R/E139K

[0112] MspA 单体可在任何以下氨基酸位置包含一个或更多突变 :88, 105, 108, 118, 134 或 139。MspA 单体可包含一种或更多以下突变 :L88W, D90K/N/Q/R, D91N/Q, D93N, I105W, N108W, D118R, D134R 或 E139K。MspA 单体可包含以下突变 :D90N/D91N/D93N。MspA 单体可包含以下突变 :D90N/D91N/D93N/D118R/D134R/E139K。MspA 单体可包含以下突变 :D90Q/D91Q/D93N。MspA 单体可包含以下突变 :D90Q/D91Q/D93N/D118R/D134R/E139K。MspA 单体可包含以下突变 :D90 (K, R) /D91N/D93N。MspA 单体可包含以下突变 : (L88, I105)W/D91Q/D93N。MspA 单体可包含以下突变 :I105W/N108W。而且, MspA 单体可包含本文所述的任何其他突变。

[0113] 在本文中的任何实施方式中, 突变 Msp 孔蛋白, 诸如突变 MspA 孔蛋白或突变 MspA 旁系同原物或同源物, 可包含, 分别相比野生型 Msp 孔蛋白的前庭或收缩区域, 至少一个另外的带正电的氨基酸; 分别相比野生型 MspA 孔蛋白的前庭或收缩区域, 至少一个另外的带负电的氨基酸; 分别相比野生型 MspA 孔蛋白的前庭或收缩区域, 至少一个欠带正电的氨基酸; 或分别相比野生型 MspA 孔蛋白的前庭或收缩区域, 至少一个欠带负电的氨基酸。

[0114] 任选地, 野生型 Msp 孔蛋白的前庭和收缩区域中各带正电的氨基酸用带负电的氨基酸取代, 及各带负电的氨基酸相同或不同; 或野生型 Msp 孔蛋白的前庭和收缩区域中各带负电的氨基酸用带正电的氨基酸取代, 及各带正电的氨基酸相同或不同。

[0115] 任选地, 突变 Msp 孔蛋白的前庭或收缩区域, 分别相比野生型 Msp 孔蛋白的前庭或收缩区域, 包含更大数的带正电的残基; 或前庭或收缩区域, 分别相比野生型 Msp 孔蛋白的前庭或收缩区域, 包含更大数的带负电的残基; 或野生型 Msp 孔蛋白的前庭或收缩区域中的至少一个带正电的氨基酸, 诸如野生型 MspA 孔蛋白或野生型 MspA 旁系同原物或同源物孔蛋白被删除或被带负电的氨基酸取代; 或野生型 Msp 孔蛋白的前庭或收缩区域中的至少一个带负电的氨基酸被删除或被带正电的氨基酸取代。

[0116] 野生型 Msp 孔蛋白的前庭或收缩区域中的至少一个氨基酸, 诸如野生型 MspA 孔蛋白或野生型 MspA 旁系同原物或同源物孔蛋白, 可被具有立体地更大侧链的氨基酸; 具有立体地更小侧链的氨基酸; 具有更极性侧链的氨基酸; 具有欠极性侧链的氨基酸; 或具有更疏水侧链的氨基酸; 具有欠疏水侧链的氨基酸取代。

[0117] 在本文中的任何实施方式中, 突变 Msp 孔蛋白的前庭或收缩区域中的至少一个氨基酸可包含非天然的氨基酸或化学修饰的氨基酸。

[0118] 突变 Msp 孔蛋白, 诸如突变 MspA 孔蛋白或突变 MspA 旁系同原物或同源物孔蛋白, 可包含中性收缩区域。突变 Msp 孔蛋白, 诸如突变 MspA 孔蛋白或突变 MspA 旁系同原物或同源物孔蛋白, 可包含相比其对应野生型 Msp 孔蛋白的通过隧道的传导率更高, 诸如 2 倍更高的通过隧道的传导率。突变 Msp 孔蛋白, 诸如突变 MspA 孔蛋白或突变 MspA 旁系同原物或同源物孔蛋白, 可包含小于其对应野生型 Msp 孔蛋白的通过隧道的传导率的通过隧道的传导率。

[0119] 本文讨论的任何 Msp 孔蛋白可包含具有长度约 2 ~ 约 6nm 和直径约 2 ~ 约 6nm 的前庭, 及具有长度约 0.3 ~ 约 3nm 和直径约 0.3 ~ 约 3nm 的收缩区域, 其中前庭和收缩区域一起定义隧道。本文提供的也是突变 MspA 孔蛋白, 其包含具有长度约 2 ~ 约 6nm 和直径约 2 ~ 约 6nm 的前庭, 及具有长度约 0.3 ~ 约 3nm 和直径约 0.3 ~ 约 3nm 的收缩区域, 其中前庭和收缩区域一起定义隧道, 及还包含至少第 1 突变 MspA 旁系同原物或同源物单体。

[0120] 突变 Msp 孔蛋白, 诸如突变 MspA 孔蛋白或突变 MspA 旁系同原物或同源物的收缩区域的直径, 可小于其对应野生型 Msp 孔蛋白, 诸如野生型 MspA 孔蛋白或野生型 MspA 旁系同原物或同源物的收缩区域的直径。突变 Msp 孔蛋白, 诸如突变 MspA 孔蛋白或突变 MspA 旁系同原物或同源物, 可在前庭或收缩区域中包含突变, 其允许分析物具有小于分析物与其对应野生型 Msp 孔蛋白, (例如, 野生型 MspA 孔蛋白, 野生型 MspA 旁系同原物或同源物) 的隧道相互作用的速度或平均速度的其与隧道相互作用的速度或平均速度。

[0121] 本文讨论的野生型 Msp 单体的序列公开于 GenBank, 位于 www.pubmed.gov, 及这些序列和其他通过引用以它们整体并入本文, 如该文含有的个体亚序列或片段。例如, 野生型 MspA 单体的核苷酸和氨基酸序列可分别见于 GenBank 登录 No. AJ001442 和 CAB56052。野生型 MspB 单体的核苷酸和氨基酸序列可例如分别见于 GenBank 登录 No. NC_008596.1 (核苷酸 600086 ~ 600730) 和 YP_884932.1。野生型 MspC 单体的核苷酸和氨基酸序列可例如分别见于 GenBank 登录 No. AJ299735 和 CAC82509。野生型 MspD 单体的核苷酸和氨基酸序列可例如分别见于 GenBank 登录 No. AJ300774 和 CAC83628。由此提供的是 MspA, MspB, MspC 和 MspD 单体的核苷酸序列, 其包含与上述的核苷酸 GenBank 登录号的核苷酸序列具有至少约 70, 75, 80, 85, 90, 95, 98, 99% 或更高, 或可来源于它们的任何范围的同一性的核苷酸序列。MspA, MspB, MspC 和 MspD 单体的氨基酸序列可见于 W0 2010/034018 的图 18, 其包含与上述的氨基酸 GenBank 登录号的序列具有至少约 70, 75, 80, 85, 90, 95, 98, 99% 或更高, 或可来源于它们的任何范围的同一性的氨基酸序列。

[0122] 也提供的是 MspA 旁系同原物和同源物单体的氨基酸序列, 其包含与野生型 MspA 旁系同原物或同源物单体具有至少约 70, 75, 80, 85, 90, 95, 98, 99% 或更高, 或可来源于它们的任何范围的同一性的氨基酸序列。野生型 MspA 旁系同原物和同源物单体为本领域熟知。见表 2。

[0123] α -溶血素孔隙由 7 个相同亚基形成(七聚体)。编码 α -溶血素的一个亚基的多核苷酸序列显示于美国专利申请系列 No. 2010/0196203 的 SEQ ID NO:1, 通过引用以其整体并入本文。 α -溶血素的一个亚基的全长氨基酸序列显示于美国专利申请系列 No. 2010/0196203 的 SEQ ID NO:2。SEQ ID NO:2 的头 26 个氨基酸对应于信号肽。无信号肽的 α -溶血素的一个成熟亚基的氨基酸序列显示于美国专利申请系列 No. 2010/0196203 的 SEQ ID NO:3。SEQ ID NO:3 代替显示于 SEQ ID NO:2 的 26 个氨基酸信号肽而在第 1 位具有甲硫氨酸残基。

[0124] 变体是七聚体孔隙, 其中 7 个亚基中的 1 个或多个具有自 SEQ ID NO:2 或 3 变化而来的氨基酸序列及保留孔隙活性。突变 α -溶血素中亚基的 1, 2, 3, 4, 5, 6 或 7 个可具有自 SEQ ID NO:2 或 3 变化而来的氨基酸序列。突变孔隙之内的 7 个亚基一般相同但可不同。

[0125] 突变体可为由生物, 例如由葡萄球菌属(*Staphylococcus*) 细菌表达的天然存在的变体。变体也包括由重组技术产生的非-天然存在的变体。跨 SEQ ID NO:2 或 3 的氨基酸序列的全长, 变体可基于氨基酸同一性至少 50% 同源于该序列。亚基多肽可跨整个序列基于氨基酸同一性至少 80%, 至少 90%, 至少 95%, 至少 98%, 至少 99% 同源于 SEQ ID NO:2 或 3 的氨基酸序列。

[0126] 可对 SEQ ID NO:2 或 3 的氨基酸序列进行氨基酸取代, 例如可进行单氨基酸取代或可进行 2 个或多个取代。在一些实施方式中, 将在 SEQ ID NO:2 中的第 34 位和 SEQ ID NO:3 中的第 9 位的赖氨酸用半胱氨酸取代(即 K34C 或 K9C)。可进行的非-保守性取代的另一例是将在 SEQ ID NO:2 的第 43 位或 SEQ ID NO:3 的第 18 位的天冬酰胺残基用半胱氨酸取代(即 N43C 或 N17C)。在 SEQ ID NO:2 或 3 中包括这些半胱氨酸残基在关联位置提供氢硫基附接点。可在全部其他位置, 及在相同的亚基上的多个位置进行类似变化。

[0127] 在一些实施方式中, SEQ ID NO:2 或 3 的氨基酸序列的一个或多个氨基酸残基可替代性地或另外地删除。达 50% 的残基可删除, 由于贯穿氨基酸链长度分布的连续的区或多更小区。

[0128] 变体可包括由 SEQ ID NO:2 或 3 的片段组成的亚基。该片段保留它们的插入双层的能力。片段可为至少 100, 诸如 150, 200 或 250, 氨基酸长。该片段可用于产生嵌合孔隙。片段可包含 SEQ ID NO:2 或 3 的 β -桶结构域。

[0129] 变体包括包含 SEQ ID NO:2 或 3 的片段或部分的嵌合蛋白。嵌合蛋白由各包含 SEQ ID NO:2 或 3 的片段或部分的亚基形成。嵌合蛋白的 β -桶部分一般由 SEQ ID NO:2 或 3 的片段或部分形成。

[0130] 一个或多个氨基酸残基可替代性地或另外地插入, 或在氨基酸序列 SEQ ID NO:2 或 3 的一端或另一端或两端。1 个, 2 个或多个另外的氨基酸插入肽序列的 C 末端是欠可能扰乱蛋白的结构和 / 或功能, 及这些添加可为实质性的, 但达 10, 20, 50, 100 或 500 个氨基酸或更多的肽序列可使用。在单体的 N 末端的添加也可为实质性的, 具有添加的 1 个, 2 个或多个另外的残基, 但也添加 10, 20, 50, 500 或更多残基。也可将附加序列加入跨膜区内的蛋白, SEQ ID NO:3 的氨基酸残基 119 和 139 之间。更精确地, 附加序列可添加在 SEQ ID NO:3 的残基 127 和 130 之间, 在移出残基 128 和 129 后。添加可在 SEQ ID NO:2 中的相当的位置进行。载体蛋白可融合到本发明的氨基酸序列。

[0131] 其他任选的突变在本文所述。

[0132] OmpATb 是外膜蛋白存在于例如, 结核分枝杆菌(*M. tuberculosis*) 和牛分枝杆菌(*M. bovis*) | [isc1]。OmpATb 的孔隙 - 形成活性是 pH- 依赖性的, 致使分枝杆菌属 (*Mycobacterium*) 在酸性环境条件下生存。结核分枝杆菌 (*M. tuberculosis*) 和牛分枝杆菌 (*M. bovis*) 中的 OmpATb 的序列相同。见美国专利申请系列 No. 2010/0150966 的图 1, 其通过引用以其整体并入本文。

[0133] 术语“ OmpATb ”包含分枝杆菌属 (*Mycobacterium*) 中天然形式的蛋白, 以及在转化任何种宿主的任何类型的表达载体中产生的重组蛋白, 或也化学合成的蛋白或肽。其也包括类似物蛋白, 即具有不影响 OmpATb 的抗原性的少数变异的蛋白, 例如, 与 OmpATb 的序列具有至少 70% 序列同一性的蛋白。其他任选的突变在本文所述。

[0134] 可对 Msp 孔蛋白, Msp 单体, α - 溶血素和其变体, 及本文提供的 OmpATb 和其他蛋白进行的另外的任选的取代的描述在以下描述。

[0135] 本文所述的蛋白修饰包括氨基酸序列修饰。氨基酸序列的修饰可作为等位基因变异天然地出现 (例如, 由于遗传多态性), 可由于环境影响 (例如, 由于暴露于紫外线辐射) 而出现, 或可由人干预 (例如, 由克隆的 DNA 序列的诱变) 产生, 诸如诱导的点, 缺失, 插入和取代突变体。这些修饰可导致氨基酸序列的变化, 提供沉默突变, 修饰限制性位点, 或提供其他特定突变。氨基酸序列修饰一般落入 3 类中的一类或更多: 取代, 插入或缺失修饰。插入包括氨基和 / 或末端融合以及单或多氨基酸残基的内序列插入。插入通常会相比氨基或羧基末端融合的那些是更小插入, 例如, 以 1 ~ 4 个残基的量级。缺失特征在于一个或更多氨基酸残基自蛋白序列移出。一般而言, 在蛋白分子之内的任何一个位点删除不多于约 2 ~ 6 个残基。氨基酸取代一般是单残基, 但可在许多不同位置立即发生; 插入通常会在约 1 ~ 10 个氨基酸残基的量级; 及缺失会在约 1 ~ 30 个残基的范围。缺失或插入可在相邻配对内进行, 即, 2 个残基的缺失或 2 个残基的插入。可组合取代, 缺失, 插入或其任何组合以到达最终构建体。突变可或不将序列放置在阅读框外和可或不创建可产生二级 mRNA 结构的互补区。取代修饰是移出至少一个残基及在其位置插入不同残基的那些。

[0136] 通过已知的方法进行修饰, 包括特定氨基酸取代。例如, 修饰通过编码蛋白的 DNA 中核苷酸的定点诱变来进行, 由此产生编码修饰的 DNA, 及然后在重组细胞培养中表达 DNA。熟知在具有已知的序列的 DNA 中预定的位点制造取代突变的技术, 例如 M13 引物诱变和 PCR 诱变。

[0137] 本文所述的肽, 多肽, 单体, 多聚体, 蛋白, 等可还修饰及改变, 只要期望的功能维持或增强。需知, 限定本文公开的基因和蛋白的任何已知的修饰和衍生物或可出现的那些的一种方式是通过就特定已知的序列的同一性限定修饰和衍生物。特别公开的是与本文提供的野生型 MspA 和野生型 MspA 旁系同原物或同源物 (例如, 野生型 MspB, 野生型 MspC, 野生型 MspD, MppA, PorM1, Mmc4296) 和突变体以及 α - 溶血素和其变体及 OmpATb 具有至少 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% 同一性的多肽。

[0138] 本领域技术人员容易明白如何测定 2 个多肽的同一性。例如, 同一性可在 2 个序列对齐之后计算, 从而同一性在其最高水平。例如, 为测定 2 个氨基酸序列或 2 个核酸的“百分率同一性”, 为最佳比较目的将序列对齐 (例如, 为与第 2 氨基或核酸序列最佳比对, 可将间隔导入第 1 氨基酸序列或核酸序列)。然后比较在对应氨基酸位置或核苷酸位置的氨

氨基酸残基或核苷酸。当第 1 序列中的位置被与第 2 序列中的对应的位置相同的氨基酸残基或核苷酸占据时,则分子在该位置同一。2 个序列之间的百分率同一性是由序列共享的相同位置数的函数(即,百分率同一性 = 相同位置数 / 位置总数(例如,重叠位置) $\times 100$)。在一实施方式中,2 个序列是相同的长度。

[0139] 存在测定百分率同一性的几种方法。可用以下方式测定百分率同一性。将靶核酸或氨基酸序列使用来自含有 BLASTN 2.0.14 版和 BLASTP 2.0.14 版的独立版本的 BLASTZ 的 BLAST 2Sequences (B12seq) 程序与鉴定的核酸或氨基酸序列比较。此独立版本的 BLASTZ 可从美国政府的 National Center for Biotechnology Information 网站(www.ncbi.nlm.nih.gov) 得到。解释如何使用 B12seq 程序的使用说明可见于伴随 BLASTZ 的 readme 文件。

[0140] B12seq 进行使用 BLASTN 或 BLASTP 算法在 2 个序列之间进行比较。BLASTN 用于比较核酸序列,而 BLASTP 用于比较氨基酸序列。为了比较 2 核酸序列,选项可设置如下:-i 设置为含有待比较的第 1 核酸序列的文件(例如,C:\seq1.txt);-j 设置为含有待比较的第 2 核酸序列的文件(例如,C:\seq2.txt);-p 设置为 blastn;-o 设置为任何期望的文件名(例如,C:\output.txt);-q 设置为 -1;-r 设置为 2;及全部其他选项保持它们的默认设置。以下命令会产生含有 2 序列之间的比较的输出文件:C:\B12seq-i c:\seq1.txt-j c:\seq2.txt-p blastn-oc:\output.txt-q-1-r 2。如果靶序列与鉴定的序列的任何部分共享同源性,则指定的输出文件会作为对齐的序列呈递那些同源性区。如果靶序列不与鉴定的序列的任何部分共享同源性,则指定的输出文件会不呈递对齐的序列。

[0141] 一旦对齐的,长度通过计数来自靶序列供与自始于任何匹配的位置及结束于任何其他匹配的位置的鉴定的序列的序列比对的连续核苷酸数来测定。匹配的位置是同一核苷酸存在于靶及鉴定的序列的任何位置。由于间隔不是核苷酸,不计数存在于靶序列中的间隔。同样,由于计数靶序列核苷酸,而非自鉴定的序列的核苷酸,不计数存在于鉴定的序列中的间隔。

[0142] 跨特定长度的百分率同一性可通过计数跨该长度的匹配的位置数及用该数除以长度,之后是将得到的值乘以 100 来测定。例如,如果(1)将 50 个核苷酸靶序列与编码野生型 MspA 的序列比较(2)B12seq 程序呈现 45 个核苷酸,其来自与该 45 个核苷酸区的第 1 和最后核苷酸匹配的编码野生型 MspA 的序列区对齐的靶序列,及(3)在那些 45 个对齐的核苷酸上匹配数是 40,则 50 个核苷酸靶序列含有 45 的长度,而跨该长度上的百分率同一性是 89(即, $40/45 \times 100=89$)。

[0143] 另一计算同一性的方式可由出版的算法实施。用于比较的序列的最佳比对可由下列方法进行:由 Smith 和 Waterman, *Adv Appl Math* 2:482 (1981) 的局部同一性算法,由 Needleman 和 Wunsch, *J Mol Biol* 48:443(1970) 的同一性比对算法,由 Pearson 和 Lipman, *Proc Natl Acad Sci USA* 85:2444 (1988) 的搜索相似性方法,由这些算法的计算机化的实现(Wisconsin 遗传学软件包中的 GAP, BESTFIT, FASTA 和 TFASTA, Genetics Computer Group, 575 Science Dr., Madison, WI) 或由检查。

[0144] 可对于核酸获得相同的类型的同一性,其算法,例如,公开于 *Science* 244:48-52 (1989); *Proc Natl Acad Sci USA* 86:7706-10 (1989); 及 *Methods Enzymol* 183:281-306 (1989), 对于至少与核酸比对相关的材料通过引用并入本文。需知,一般可使用任何方法及

在特定实例中,这些各种方法的结果可不同,但本领域技术人员明白,如果用至少这些方法之一发现同一性,序列会被认为具有陈述的同一性及在本文公开。

[0145] 也公开编码本文公开的蛋白序列,以及变体和其片段的核酸。这些序列包括与特定蛋白序列相关的全部简并序列,即,具有编码一种特定蛋白序列的序列的全部核酸以及全部核酸,包括简并核酸,其编码蛋白序列的公开的变体和衍生物。由此,即便各特定核酸序列可不在于本文写出,需知,各和每序列通过公开的蛋白序列实际上在本文公开及描述。

[0146] 蛋白的片段和部分序列可在本文所述的方法中 useful。如同全部肽和蛋白,包括其片段,需知,可发生不改变肽和蛋白的性质或功能的本文公开的蛋白的另外的氨基酸序列的修饰。需知,对于这些的唯一限制是实践,它们必需包含在关联实施方式中使用的必需功能元件(例如,隧道-形成能力)。该修饰包括保守性氨基酸取代及在下文进一步详细讨论。

[0147] 以下表提供可在测定如何选择用于本文所述的蛋白(例如,蛋白孔)的修饰的氨基酸中辅助本领域技术人员的氨基酸性质的非限制性例。

[0148] 表 4. 氨基酸性质

[0149]

氨基酸	百分率埋藏残基 a (%)	平均体积 b (Å ³)	范德华体积 c (Å ³)	可接近的表面积 d (Å ²)	氨基酸极性的排序 e
丙氨酸	38 (12)	92	67	67	9 (7)
精氨酸	0	225	148	196	15 (19)
天冬酰胺	10 (2)	135	96	113	16 (16)
天冬氨酸	14.5 (3)	125	91	106	19 (18)
半胱氨酸	47 (3)	106	86	104	7 (8)
谷氨酰胺	6.3 (2.2)	161	114	144	17 (14)
谷氨酸	20 (2)	155	109	138	18 (17)
甘氨酸	37 (10)	66	48		11 (9)
组氨酸	19 (1.2)	167	118	151	10 (13)
异亮氨酸	65 (12)	169	124	140	1 (2)
亮氨酸	41 (10)	168	124	137	3 (1)
赖氨酸	4.2 (0.1)	171	135	167	20 (15)
甲硫氨酸	50 (2)	171	124	160	5 (5)
苯丙氨酸	48 (5)	203	135	175	2 (4)
脯氨酸	24 (3)	129	90	105	13 (-)
丝氨酸	24 (8)	99	73	80	14 (12)
苏氨酸	25 (5.5)	122	93	102	12 (11)
色氨酸	23 (1.5)	240	163	217	6 (6)
酪氨酸	13 (2.2)	203	141	187	8 (10)
缬氨酸	56 (15)	142	105	117	4 (3)

[0150] a 此栏代表待埋在蛋白内部埋藏的氨基酸趋势(定义为 <5% 的溶剂中可利用的残基)及基于 9 个蛋白的结构(研究总 ~2000 个体残基, 这些中 587 个(29%)被埋藏)。值指示各氨基酸被发现埋藏的频度,相对于见于蛋白的此氨基酸的残基总数。括弧中的值指示相对于蛋白中全部埋藏残基发现的此氨基酸的埋藏残基数。自 BioTechnology8:308 (1990)的数据;对于具有类似结果的其他计算方法,见 Nature277:491 (1979);及 Science 229:834

(1985)。

[0151] b 埋藏残基的平均体积(V_r),由侧链的表面积计算。Annu Rev Biophys Bioeng 6:151 (1977);Protein Eng 2:329 (1989)。

[0152] c 自 Darby N.J.and Creighton T.E.Protein structure. In In focus(ed. D. Rickwood), p. 4. IRL Press, Oxford, United Kingdom(1993) 的数据。

[0153] d 具有延伸的构象中的主链的 Gly-X-Gly 三肽中残基 X 的氨基酸侧链的总可接近的表面积(ASA)。J Mol Biol 196:641 (1987)。

[0154] e 显示的值代表根据它们在用于 38 个公开的疏水性尺度的各序列排序上发生的频度的氨基酸平均值排序。Protein Eng 11:153 (1998)。尽管大多数这些疏水性尺度源于分离的氨基酸的化学行为或物理化学性质的实验测量(例如,水中的溶解度,水和有机溶剂之间的分隔部,层析迁移或对表面张力的效应),包括几个“操作”疏水性尺度基于蛋白中氨基酸的已知的环境特征,诸如它们的溶剂可接近性或它们占据蛋白核心的倾向性(基于在由 x-射线晶体学或 NMR 观察的三级结构中残基的位置)。更低排序代表最疏水氨基酸,及更高值代表最亲水氨基酸。为比较性目的, Radzicka 和 Wolfenden, Biochem 27:1664 (1988) 的疏水性尺度显示于括弧。该尺度源于氨基酸的测量的水合势,其基于它们的自蒸汽相转移到环己烷,1-辛醇和中性水溶液的自由能。

[0155] 或者,可考虑氨基酸的亲水性指数。各氨基酸已基于它们的疏水性和 / 或荷电特征分配亲水性指数,这些是:异亮氨酸(+4.5);缬氨酸(+4.2);亮氨酸(+3.8);苯丙氨酸(+2.8);半胱氨酸 / 胱氨酸(+2.5);甲硫氨酸(+1.9);丙氨酸(+1.8);甘氨酸(-0.4);苏氨酸(-0.7);丝氨酸(-0.8);色氨酸(-0.9);酪氨酸(-1.3);脯氨酸(-1.6);组氨酸(-3.2);谷氨酸(-3.5);谷氨酰胺(-3.5);天冬氨酸(-3.5);天冬酰胺(-3.5);赖氨酸(-3.9);和 / 或精氨酸(-4.5)。对蛋白赋予相互作用生物学功能的亲水性氨基酸指标的重要性为本领域通常所知。已知,特定氨基酸可被具有类似亲水性指数和 / 或评分和 / 或仍保留类似生物学活性的其他氨基酸取代。在基于亲水性指数进行变化中,氨基酸的取代,其亲水性指数可在 ± 2 之内;在 ± 1 之内,或在 ± 0.5 之内。

[0156] 本领域也明白,可基于亲水性有效进行相似氨基酸的取代。如详述于美国专利 4,554,101,并入本文通过引用,以下亲水性值已分配到氨基酸残基:精氨酸(+3.0);赖氨酸(+3.0);天冬氨酸(+3.0 ± 1);谷氨酸(+3.0 ± 1);丝氨酸(+0.3);天冬酰胺(+0.2);谷氨酰胺(+0.2);甘氨酸(0);苏氨酸(-0.4);脯氨酸(-0.5 ± 1);丙氨酸(-0.5);组氨酸(-0.5);半胱氨酸(-1.0);甲硫氨酸(-1.3);缬氨酸(-1.5);亮氨酸(-1.8);异亮氨酸(-1.8);酪氨酸(-2.3);苯丙氨酸(-2.5);色氨酸(-3.4)。在基于类似亲水性值进行变化中,其涵盖氨基酸的取代,其亲水性值可在 ± 2 之内,在 ± 1 之内或在 ± 0.5 之内。

[0157] 任何突变蛋白可相比野生型 Msp 孔蛋白或单体,包含保守性氨基酸取代。任何取代突变是以最低限度破坏蛋白的生物化学性质的保守性。导入以代替保守性氨基酸残基的突变的非限制性例包括:带正电的残基(例如, H, K 和 R) 取代为带正电的残基;带负电的残基(例如, D 和 E) 取代为带负电的残基;中性极性残基(例如, C, G, N, Q, S, T 和 Y) 取代为中性极性残基;及中性非 - 极性残基(例如, A, F, I, L, M, P, V 和 W) 取代为中性非 - 极性残基。保守性取代可根据以下表 5 进行。非保守的取代也可进行(例如, 脯氨酸为甘氨酸)。

[0158] 表 5:例示氨基酸取代

[0159]

氨基酸	取代
Ala	Ser, Gly, Cys
Arg	Lys, Gln, Met, Ile
Asn	Gln, His, Glu, Asp
Asp	Glu, Asn, Gln
Cys	Ser, Met, Thr
Gln	Asn, Lys, Glu, Asp
Glu	Asp, Asn, Gln
Gly	Pro, Ala
His	Asn, Gln
Ile	Leu, Val, Met
Leu	Ile, Val, Met
Lys	Arg, Gln, Met, Ile
Met	Leu, Ile, Val
Phe	Met, Leu, Tyr, Trp, His
Ser	Thr, Met, Cys
Thr	Ser, Met, Val
Trp	Tyr, Phe
Tyr	Trp, Phe, His
Val	Ile, Leu, Met

[0160] 纳米孔一般能插入脂质双层或其他薄膜,及这些技术为本领域熟知,如本文解释。此外,美国专利 No. 6, 746, 594, 通过引用并入本文,描述各种脂质双层和薄膜,包括可关于本文讨论的纳米孔采用的无机材料。方法,设备和技术描述于美国专利 No. 6, 267, 872, 通过引用以其整体并入本文,也可关于本文讨论的纳米孔采用。

[0161] 在一些实施方式中,多个纳米孔包含在包含分枝菌酸的人工膜中。例如,2,3,4,5,10,20,200,2000 或更多可包含在膜中。

[0162] 任选地, 2, 3, 4, 5, 10, 20, 200, 2000 或更多纳米孔包含在膜, 双层或薄膜中。的确, 任意 2 ~ 1010 纳米孔可在本文所述的实施方式中采用。该多个纳米孔可为纳米孔簇的形式。簇可随机组装或可设置成图案。如本文所用, “簇” 指称分组在一起及作为单元移动, 但不彼此共价结合的分子。

[0163] 作为“包含”的替代或附加, 本文的任何实施方式可叙述“由... 组成”。过渡短语“由... 组成”排除权利要求中不特定的任何要素, 步骤或成分。

[0164] 本文的任何实施方式可任选地排除本文中的任何其他实施方式。

[0165] 权利要求中术语“或”的使用用于表示“和 / 或”除非明确地指示指称仅替代或替代是相互排斥的, 尽管公开支持指称仅替代和“和 / 或”的定义。

[0166] 贯穿本申请, 术语“约”用于指示包括待采用以测定值的装置或方法的误差的标准偏差的值。在结合术语“约”使用的数值的情景中讨论的任何实施方式中, 特别涵盖术语约可忽略。

[0167] 根据长期专利法, 词汇“a”和“an”, 当在权利要求或说明书中结合词汇“包含”使用时, 表示 1 个或更多, 除非特别提及。

[0168] 公开的是可用于, 可联用于, 可用于制备的物质, 组合物和组分, 或是公开的方法和组合物的产物。这些和其他物质在本文公开, 及应知, 当公开这些物质的组合, 子集, 相互作用, 组, 等, 而可不明确地公开这些化合物的各种个体及集体组合和置换的特定参照时, 各在本文特别涵盖及描述。例如, 如果公开及讨论方法及讨论可对许多包括方法的分子进行许多修饰, 特别涵盖方法的各和每组合和置换, 及可能的修饰, 除非特别相反指示。同样, 也特别涵盖及公开这些的任何子集或组合。此概念应用于本公开全部方面, 包括但不限于使用公开的组合物中的方法中的步骤。由此, 如果有各种可进行的另外的步骤, 需知, 各这些另外的步骤可用任何特定方法步骤或公开的方法的方法步骤的组合进行, 及特别涵盖组合的各该组合或子集及应被认为是公开的。因此涵盖, 此说明书中讨论的任何实施方式可关于本文描述的任何方法, 系统或组合物, 等实施, 及反之亦然。例如, 本文所述的任何纳米孔可在本文所述的任何方法中采用。

【实施例】

[0169] 【实施例 1 : 实施例 1 ~ 5 中使用的材料和方法】

[0170] 下列实施例中使用的实验设置已之前描述 (Biophys J 77, 3227(1999) 和 Biophys J 90:190 (2006))。简言之, 20 ~ 90 μm 特氟隆® 孔由围绕精致地锐化的针或线的熔解特氟隆® 热缩管道 (Small Parts, Inc) 形成。移出针或线之后, 切割管道以形成期望的宽度的孔。然后将管道弯曲, 以连接在特氟隆® 支架中的 2 ~ 200 μL 孔。Ag-AgCl 电极用于接地得到的顺式孔及将反式孔连接到以电压钳模式操作的 Axopatch™ 200B, 1B 或 1C 膜片钳放大器。为了发现超过膜片钳放大器的 1.2V 最大输出电压的 MA 膜的膜破裂电压, 与 Keithley® 485 沙安计使用系列可变电电源。将 DI 水中的 1.0M KCl 溶液, 用 10mM HEPES 缓冲到 pH8.0 +/- 0.05, 电连接 2 个孔。

[0171] 以 250kHz 或 500kHz 采样离子通道电流, 和用 4- 极 Bessel 滤波器以 1/5 采样率过滤低通。由 LabWindows® 和 LabVIEW® 写的自定义软件控制数据获取。对于寿命测量, 以 10Hz 采样数据。

[0172] 【实施例 2:人工分枝菌酸膜形成】

[0173] 购买从结核分枝杆菌(*M. tuberculosis*)提取的分枝菌酸($\geq 98\%$, Sigma-Aldrich, St. Louis, MO)及溶于氯仿至 50g/L 及存储于 -20°C 直到使用。分枝菌酸膜使用涂技术形成,随二棕榈酰磷脂酰胆碱(DPhPC)广泛用在类似实验中(*Biophys J* 77:3227 (1999))。过程始于 2 个预备步骤。在第 1 步骤中,将特氟隆®孔用脂质-己烷混合物的包被预处理。将 $1\ \mu\text{L}$ 的 MA/氯仿溶液在玻璃试管中风干,然后重悬浮于 0.01g 己烷。将 $1\ \mu\text{L}$ 的重悬浮的 MA 预处理溶液应用于特氟隆®孔的顺式侧,然后将轻空气压力用注射器自反式侧施加,以随己烷蒸发清除孔。一旦自孔清除溶液之后,应用另一 $1\ \mu\text{L}$ 的预处理溶液及清除。允许系统风干 15 分钟之后,通过将 KCl 缓冲剂放入孔(aperture),管和孔(well)来建立 2 电极之间的电连接。

[0174] 在第 2 预备步骤中,将脂质涂到孔上。将 $10\ \mu\text{L}$ 的 MA/氯仿溶液在氯仿-清洁的玻璃载玻片上风干。然后, $\sim 0.1\ \mu\text{L}$ 十六碳烷应用于 MA 和将溶液加热到 35°C 大致 5min,以促进十六碳烷掺入脂质。当 MA-十六碳烷混合物达到凝胶-样一致性时,将 $\sim 1\text{mm}$ 直径团迹的混合物应用于单硬毛刷。在监控离子电流通过孔的同时,将脂质-溶剂混合物轻轻应用到特氟隆®孔上,直到电流下降到 0。手动迫使从反式侧通过孔的缓冲剂消除物理阻断。对于大于 $40\ \mu\text{m}$ 直径的孔,将脂质应用于孔的外部边缘,而非孔顶部上。

[0175] 通过使用微量移液器将 $3\sim 6\ \mu\text{L}$ 气泡放置在制备的孔上来形成膜,然后轻轻收缩气泡。由通过孔的电流急剧下降到 0 指示膜形成。如果纳米孔可在脂质系统之内形成,推定存在双层膜,而非多-薄层状脂质结构或物理脂质梗阻。这些孔隙的观察在以下描述。

[0176] 【实施例 3:测定人工分枝菌酸膜的性质】

[0177] 形成。以与 DPhPC 膜类似可靠性形成 MA 膜。而且, MspA 类似地以 ~ 0.2 孔隙/s、以 $\sim 10\text{nM}$ MspA 的浓度掺入 MA (图 2) 和 DPhPC 膜。MA 膜中的 MspA 孔隙延续几小时,之后自发离开膜,类似于 DPhPC 膜中 MspA 孔隙的行为。

[0178] MspA 具有大致 9nm 的高度和仅 $\sim 5\text{nm}$ 的疏水长度,其限制纳米孔可穿透的疏水屏障尺寸。在体内,分枝杆菌属(*Mycobacterium*)的外膜是 $\sim 8\text{nm}$ 。如果膜由 2 个或更多层的 MA (用油层分离为二)组成, MspA 不会跨膜。因此 MA 系统之内的跨膜 MspA 孔隙的插入强烈提示双层膜的形成,而非其他脂质构型。

[0179] 膜电阻,破裂,电容和寿命。使用 Axopatch®放大器检查 MA 膜,以测定它们的电阻的更低范围。在 $20\sim 40\ \mu\text{m}$ 孔上,当跨膜施加 $\pm 1.2\text{V}$ 时,测量的离子电流是 $<1\text{pA}$,对应于 $>1\text{T}\Omega$ 电阻。在相同的孔上形成的 DPhPC 膜也呈现电阻值 $>1\text{T}\Omega$ 。

[0180] 为了测定 MA 和 DPhPC 膜的破裂电压,施加的电压以约 100mV/s 渐变,直到跨膜电流在破裂电压, V_{rup} 突然增加。将膜然后通过应用另一气泡来改性,及重复过程。破裂电压的直方图显示于图 3。对于 MA 膜,测定得 $V_{\text{rup,MA}}=2.0\text{V}$ 的平均破裂电压,具有 0.7V 的标准偏差($N=330$)。为了比较,在相同的孔上及用相同的操作条件形成 DPhPC 脂质膜及发现 $V_{\text{rup,DPhPC}}=0.50\text{V}$,具有 0.09V 的标准偏差($N=209$)。因为 MA 膜抵御的相对高施加的电压, B&K PRECISION® 875b 电容计可用于量度它们的电容。对于具有 $56\ \mu\text{m}\sim 85\ \mu\text{m}$ 之间的直径的特氟隆®孔,发现该膜上最大的仍可形成,电容值跨 $0.9\sim 2.8\times 10^{-3}\text{F/m}^2$,指示 7 和 22nm 之间的平均厚度(介电常数 $\epsilon_r=2.3$)。这些厚度与体内发现的 $\sim 8\text{nm}$ 的膜厚度一致。不被理论束缚,发明人将大范围的厚度归因于几个因素,包括膜的实际面积的不确定度

和溶剂掺入膜的未知的程度。膜形成之后立即,电容以 ~ 5 分钟的时间常数升向渐近值。该电容的增加与用 DPhPC 脂质观察的双层形成一致(Biophys J10:1127 (1970))。

[0181] 通过监控用 MspA 纳米孔形成的膜的传导率来检查 MA 膜的寿命。膜保持 200mV 施加,直到在多于 3 天后实验终结(N=4)。此在 3 天赋予 MA 膜寿命下限,显著大于用 DPhPC 膜观察的 1 天的最长寿命。

[0182] pH 对膜稳定性的影响。通过用缓冲到 pH2 ~ pH12 之间的不同 1MKCl 溶液取代在顺式侧上 1M KCl pH8.0- 缓冲剂来测定 pH2 ~ pH12 的碱性和酸性条件的存在下膜的稳定性。将 pH8 的溶液用 10mMHEPES 缓冲,而其他溶液用 40mM CH₃COOH, 硼酸和 K₂HPO₄ 的适当的混合物缓冲。各种缓冲剂的 pH 使用用 **Beckman®** 电极在适当的 pH 范围内校准的 Orion **perpHect®** logR 计测量。

[0183] 可形成 MA 膜及用提供于它们的顺式侧的 pH2 ~ pH9 的缓冲剂改性。膜在各 pH 生存至少 20min。在低至 2 的 pH,膜可容易改性及孔隙成功地插入。在 pH9 以上,膜形成缓和,但使缓冲剂回到 <7 的 pH,恢复膜形成和稳定性。膜稳定性通过测量离子电流监控。用 200mV 施加的可测量的电流的存在指示漏的形成及显著减小双层稳定性。

[0184] 【实施例 4:单 MspA 通道实验】

[0185] M1-NNN-MspA 纳米孔的制备描述于美国临时申请系列 No. 61/098, 938 和其相关的 PCT 申请, WO 2010/034018, 各通过引用以其整体并入本文。也见 Proc Natl Acad Sci 105:20647 (2008)。

[0186] 在单通道实验中 M1-NNN-MspA 的浓度是在 0.01%wt/v **Genapol®** 和 DI- 水中的 0.04 μ g/mL;用于多-通道实验的浓度是在 0.1%wt/v 中的 0.4 μ g/mL。自此溶液,将 $\sim 1 \mu$ L 加入以上 100 ~ 200 μ L 的顺式体积 MA 膜然后彻底混合。将 MA 膜用以上描述的气泡技术改性,之后观察逐步传导率变化。如果测量传导率对于单通道适当,将蛋白溶液快速灌注到含工作缓冲剂的顺式孔中,以避免其他通道的插入。

[0187] 为了展示掺入 MA 的 MspA 是适当的跨膜通道,测量它们的 I-V 曲线及进行 DNA 转位实验。MspA 的 I-V 曲线与 DPhPC 膜中 MspA 的那些呈现相同的定量特征(图 4),指示 MspA 能跨膜及 MspA 的内部通道不是被其膜环境略微影响。为了还确保通道完整性和 MA-MspA 组合用于纳米孔测序的有用性,进行 DNA 转位实验。如已之前用 DPhPC 膜中的单 MspA 通道进行重复实验(Proc Natl Acad Sci107:16060 (2010))。由于单链 DNA 转位(>1nt/ μ s)对于观察良好-表征的电流标记太快,使用不可完全转位的 DNA 发卡,直到双链段解离。此短暂中止期间,发卡 DNA 的单链段把持在 MspA 的收缩内,产生良好-溶解的离子电流水平。这些电流水平是在收缩处的核苷酸的特征(图 5)。实验用几个 DNA 发卡尾序列进行,和对其测定,离子电流水平无法与用 DPhPC 膜中的 MspA 类似实验区别(图 6)。

[0188] 【实施例 5:实验 1 ~ 4 的结果的总结和注释】

[0189] MA 膜相比可比较的 DPhPC 膜呈现更稳定性。上述实验展示由 MA 制造的体外建立的第 1 未支持的膜。相比 DPhPC 膜,MA 膜在破裂之前抵御显著更大电压。以实验方式,已知,许多因素影响膜的稳定性。MA 和 DPhPC 膜具有显著不同的溶解温度;结核分枝杆菌(M. tuberculosis)细胞壁提取物已发现改变高至 63°C 的相,而 DPhPC 无知道的相变化。得出 MA 膜的显著的稳健性归因于 MA 的成分脂质链的结构和相互作用。

[0190] MA 和 DPhPC 分子之间的显著差异是脂质头基和尾的化学结构。例如,已关于脂质

链长度的重要性进行研究。虽然带负电的 MA 头基是基本上不同于两性离子 DPhPC 头基,均不在 pH8 形成氢键,提示头基是不可能解释破裂电压差异。由此,不被理论束缚,发明人假定 MA 尾的尺寸和构型及关联的得到的膜的厚度似乎是 MA 膜的稳定性的决定因素。尤其是,其已推荐,膜之内 MA 脂质尾的组装可在膜的体内功能中起到显著作用(Trends Microbio 18:109 (2010))。人工 MA 膜对于分枝杆菌外膜的更近检查有用。

[0191] MA 膜中的跨膜孔隙。膜对孔隙形成,构象和功能的影响是在药物开发,及理解孔蛋白的折叠力学中尚未回答的问题。对于 MA 及 DPhPC 膜中的通道 MspA,观察到,开放 - 通道电流呈现同一传导率和纠正。此外在 2 种膜中观察到同一通过 MspA 的 DNA 转位性质。这些传导率和 DNA 转位性质是对于 MspA 结构高度敏感的(Proc Natl Acad Sci 107:16060 (2010))。因此,这些观察提示,由 MA 和 DPhPC 组成的膜之间的实质性差异不略微改变 MspA 通道的跨膜构象和功能。

[0192] 虽然已例证及描述例证性实施方式,需知,可对其进行各种变化而不离开本发明的精神和范围。

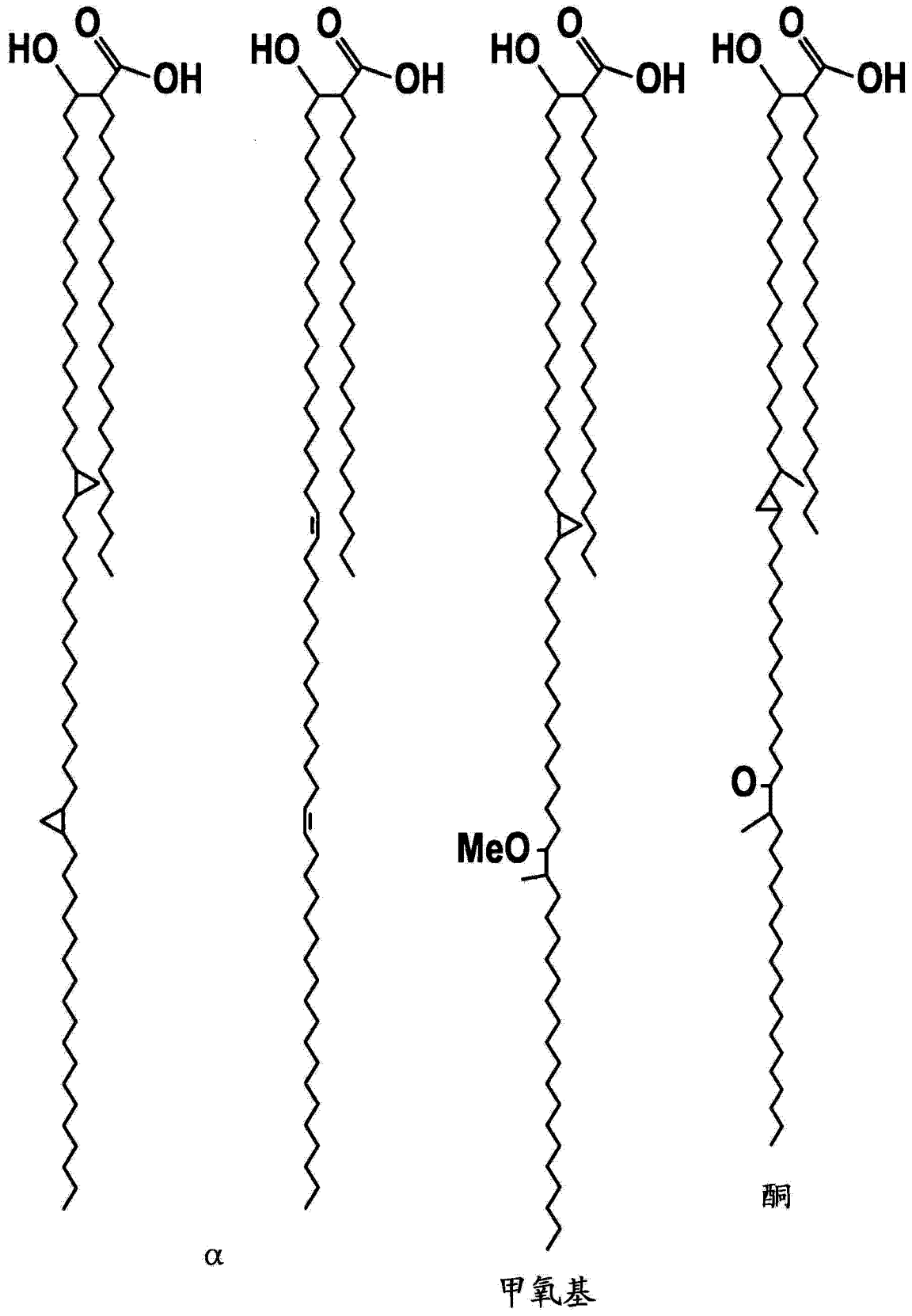


图 1

MspA逐步插入MA膜

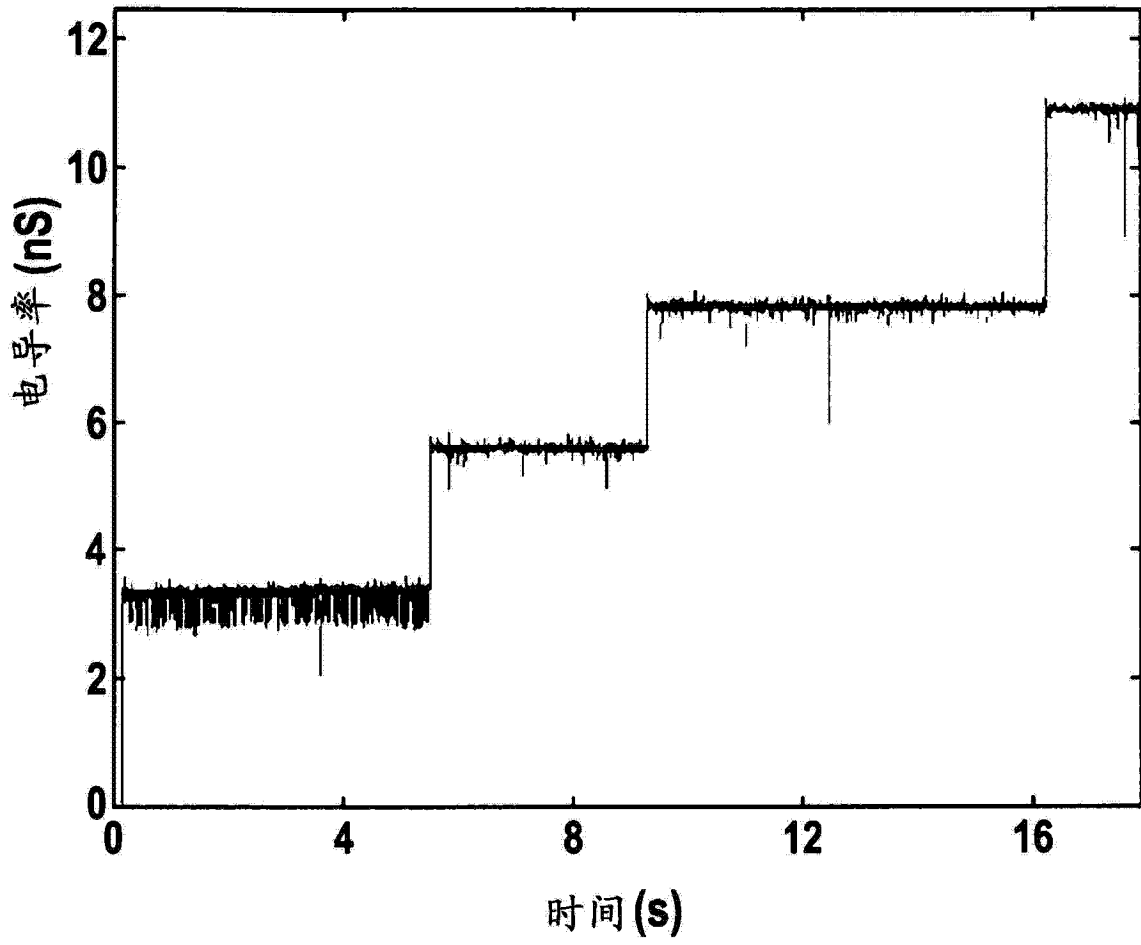


图 2

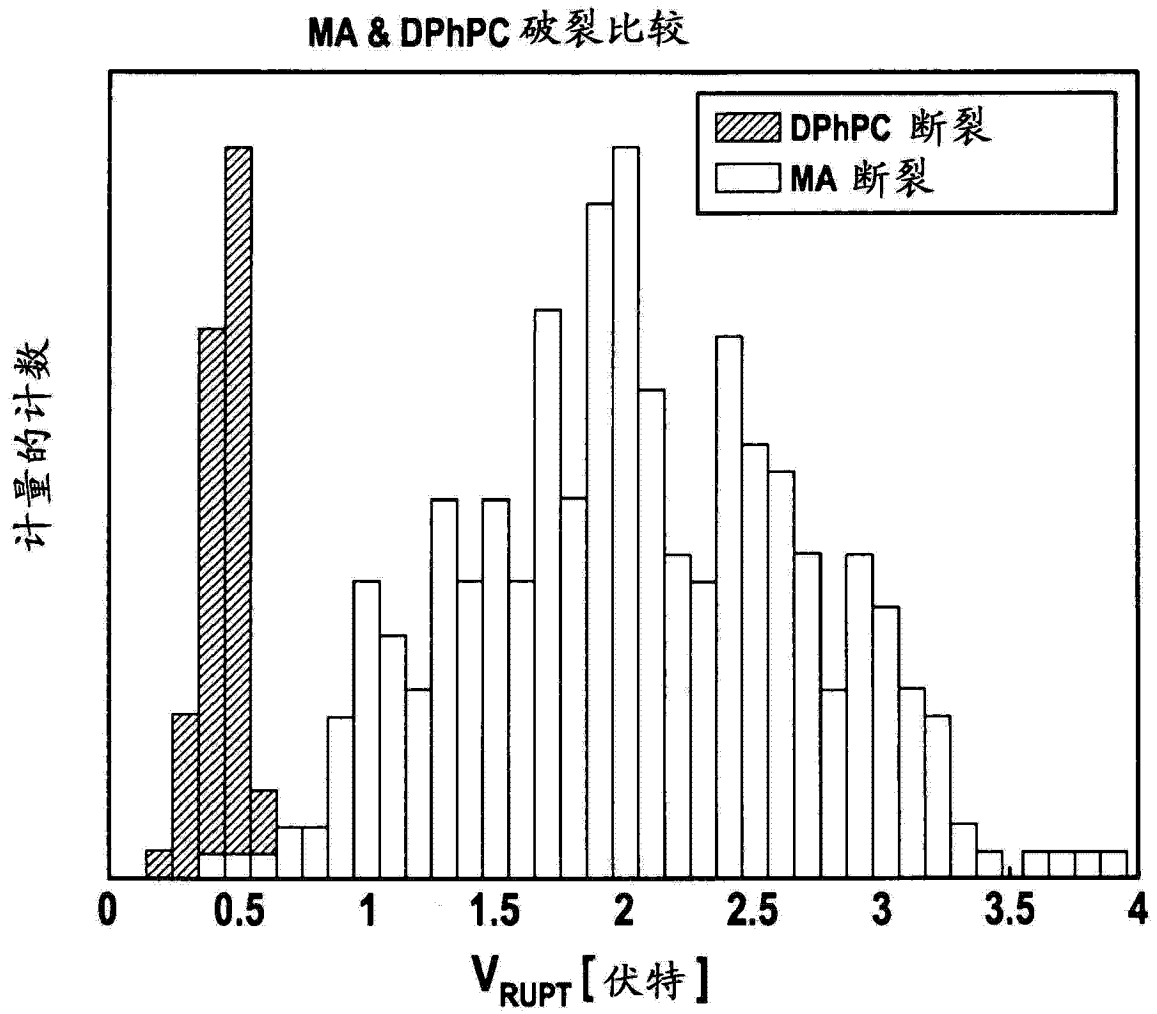


图 3

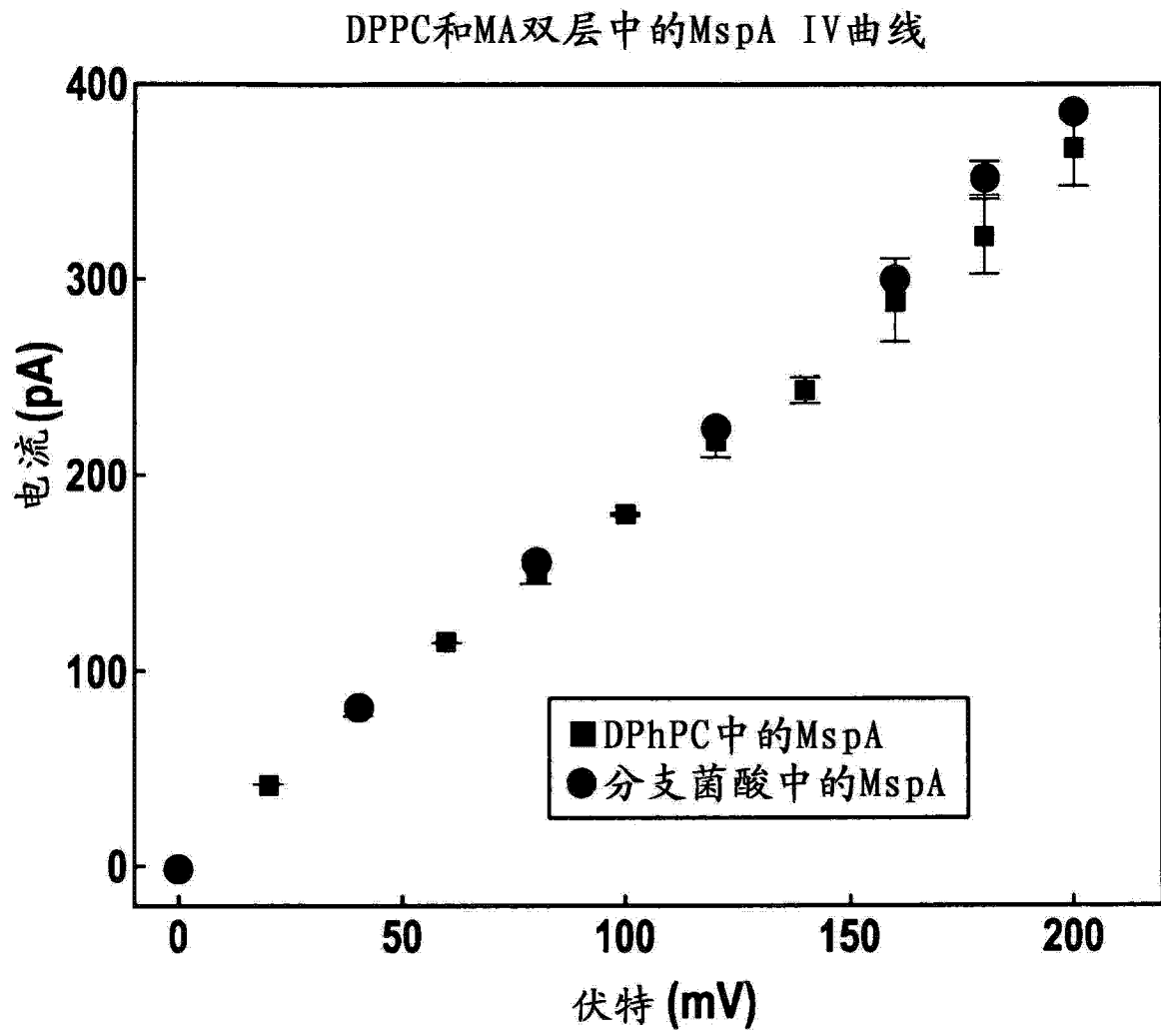


图 4

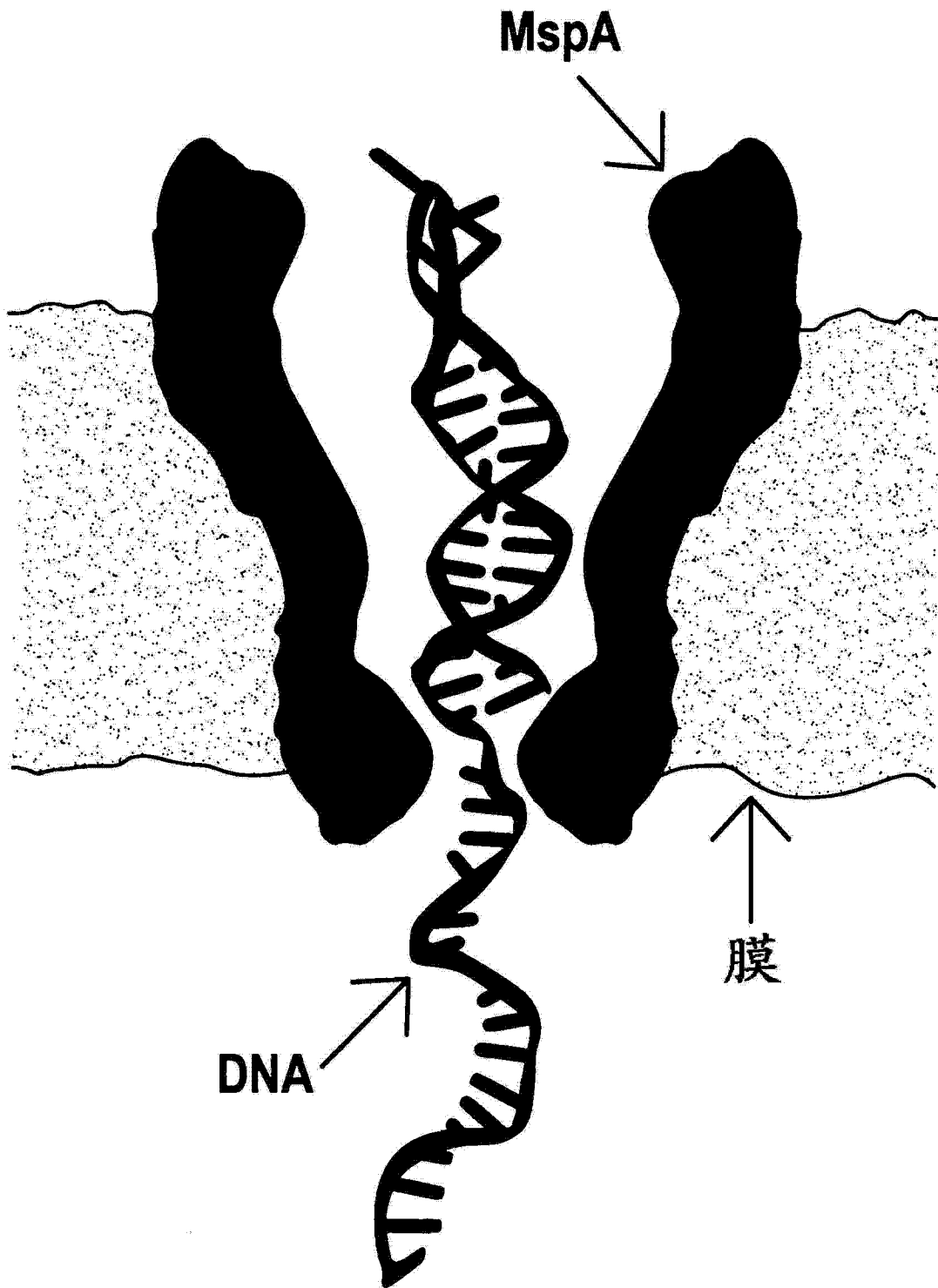


图 5

多聚A发卡阻断深度

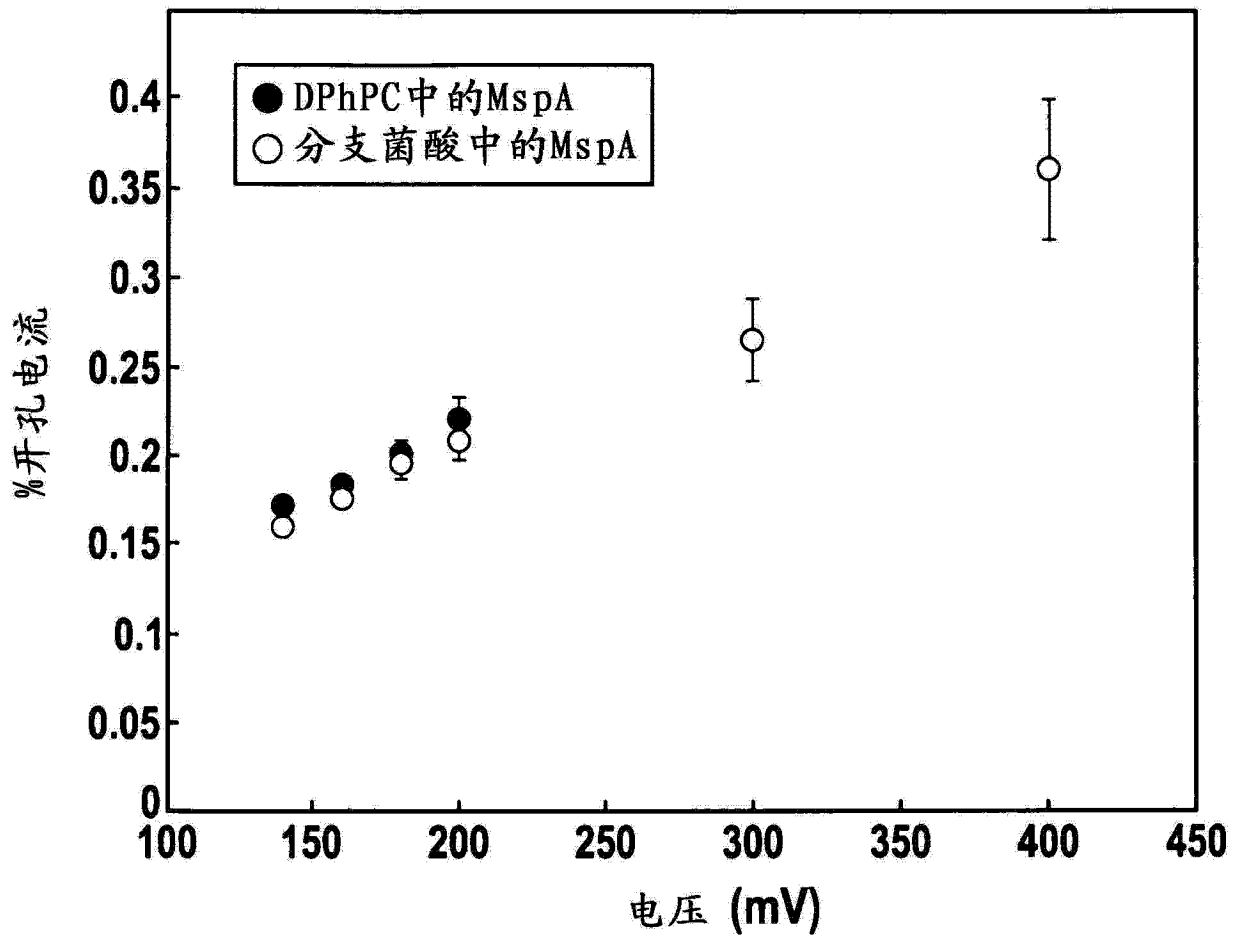


图 6