



(12) 发明专利

(10) 授权公告号 CN 109072295 B

(45) 授权公告日 2022. 05. 31

(21) 申请号 201680081416.8

(22) 申请日 2016.12.08

(65) 同一申请的已公布的文献号  
申请公布号 CN 109072295 A

(43) 申请公布日 2018.12.21

(30) 优先权数据  
62/264,709 2015.12.08 US

(85) PCT国际申请进入国家阶段日  
2018.08.08

(86) PCT国际申请的申请数据  
PCT/IB2016/001841 2016.12.08

(87) PCT国际申请的公布数据  
W02017/098322 EN 2017.06.15

(73) 专利权人 鲁汶大学研究与发展中心  
地址 比利时鲁汶

(72) 发明人 乔瓦尼·马利亚  
洛伦佐·弗兰切斯基尼  
蒂内·邦斯 安德鲁·约翰·赫伦  
拉科马·尼尚萨·贾亚辛格  
伊丽莎白·杰恩·华莱士

(74) 专利代理机构 北京同立钧成知识产权代理有限公司 11205  
专利代理师 彭雪瑞 臧建明

(51) Int.Cl.  
C12Q 1/6869 (2018.01)

(56) 对比文件  
WO 2014153625 A1, 2014.10.02  
WO 2015166276 A1, 2015.11.05  
WO 2015051378 A1, 2015.04.09  
Misha Soskine等.Tuning the size and properties of ClyA nanopores assisted by directed evolution.《J Am Chem Soc》.2013, 第135卷(第36期),

Lorenzo Franceschini等.A nanopore machine promotes the vectorial transport of DNA across membranes.《Nat Commun》.2013, 第4卷(第1期),

Alicia Guasch等.Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage phi29 connector particle.《J Mol Biol》.2002, 第315卷(第4期),

审查员 谢培

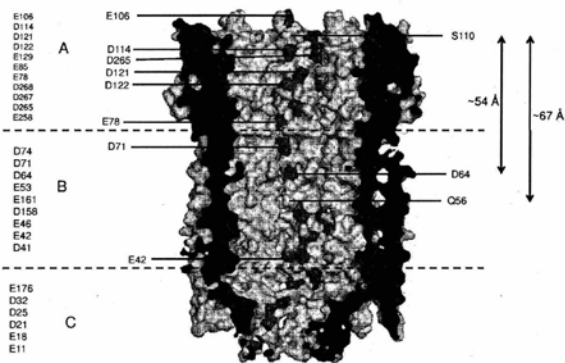
权利要求书2页 说明书47页  
序列表7页 附图20页

(54) 发明名称

修饰的纳米孔, 包含其的组合物及其应用

(57) 摘要

本文提供涉及修饰或突变形式的溶细胞素A (ClyA) 及包含其的组合物。特别地, 修饰或突变形式的ClyA允许在低或生理离子强度下通过修饰的或突变的ClyA纳米孔对带负电荷的靶分子或聚合物有效捕获和/或转移。因此, 还提供了使用修饰或突变形式的ClyA和组合物, 例如表征带负电荷的靶分析物, 例如靶多核苷酸的方法。



CN 109072295 B

1. 一种修饰的溶细胞素A纳米孔, 包含顺式开口, 中间部分和反式开口, 其中所述顺式开口的内表面包含第一带正电荷的氨基酸取代; 所述中间部分的内表面包含第二带正电荷的氨基酸取代; 并且所述反式开口包含带负电荷的收缩部;

其中所述第一带正电荷的氨基酸取代是用带正电荷的氨基酸取代SEQ ID NO:1或SEQ ID NO:2的位置S110处的氨基酸; 和

其中所述第二带正电荷的氨基酸取代是用带正电荷的氨基酸取代SEQ ID NO:1或SEQ ID NO:2的选自D74, D71, D64, E53, E161, D158, E46, E42, 和D41中的一个位置处的氨基酸。

2. 根据权利要求1所述的修饰的溶细胞素A纳米孔, 其中所述第一和第二带正电荷的氨基酸取代各自包含精氨酸。

3. 根据权利要求1所述的修饰的溶细胞素A纳米孔, 其中所述第一带正电荷的氨基酸取代是SEQ ID NO:2的氨基酸序列中的S110R突变, 和/或

其中所述第二带正电荷的氨基酸取代是SEQ ID NO:2的氨基酸序列中的D64R突变。

4. 根据权利要求1所述的修饰的溶细胞素A纳米孔, 其中所述修饰的溶细胞素A纳米孔是十二聚体孔或13聚体孔。

5. 根据权利要求1所述的修饰的溶细胞素A纳米孔, 其中所述修饰的溶细胞素A孔是十二聚体孔。

6. 一种使DNA移位穿过溶细胞素A纳米孔的方法, 包括:

a. 在150mM到300mM的低离子强度溶液中提供根据权利要求1-4中任一项所述的修饰的溶细胞素A纳米孔, 和人造膜, 其中所述修饰的溶细胞素A纳米孔存在于所述人造膜中, 使得所述修饰的溶细胞素A纳米孔的顺式开口存在于所述低离子强度溶液的顺式侧并且所述修饰的溶细胞素A纳米孔的反式开口存在于所述低离子强度溶液的反式侧;

b. 在所述低离子强度溶液的顺式侧提供DNA; 和

c. 跨所述修饰的溶细胞素A纳米孔施加电势, 从而使DNA从所述顺式侧移位穿过所述修饰的溶细胞素A纳米孔到所述反式侧。

7. 一种多核苷酸, 包含这样的核苷酸序列: 所述核苷酸序列编码根据权利要求1-5中任一项所述的修饰的溶细胞素A纳米孔的亚单位多肽。

8. 一种同源多聚体修饰的溶细胞素A纳米孔, 包含多个根据权利要求1所述的修饰的溶细胞素A纳米孔的亚单位多肽。

9. 一种异源多聚体修饰的溶细胞素A纳米孔, 包含至少一个根据权利要求1所述的修饰的溶细胞素A纳米孔的亚单位多肽。

10. 根据权利要求1所述的修饰的溶细胞素A纳米孔, 包含亚单位多肽, 所述亚单位多肽具有与SEQ ID NO:1或SEQ ID NO:2中所示的氨基酸序列至少95%同一性的氨基酸序列。

11. 根据权利要求1所述的修饰的溶细胞素A纳米孔, 包含具有多达10个取代的亚单位多肽, 与包含所述第一和第二带正电荷的氨基酸取代的SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列相比。

12. 一种组合物, 包含至少一个根据权利要求1-5或8-11中任一项所述的修饰的溶细胞素A纳米孔。

13. 一种表征靶多核苷酸的方法, 所述方法包括:

(a) 在150mM至300mM的低离子强度溶液中, 提供根据权利要求1-5或8-11中任一项所述

的修饰的溶细胞素A纳米孔,和膜,其中所述修饰的溶细胞素A纳米孔存在于所述膜中;

(b) 在步骤(a)的所述低离子强度溶液中加入靶多核苷酸;和

(c) 在跨所述纳米孔施加电势期间,测量经过所述修饰的溶细胞素A纳米孔的离子流,其中所述离子流测量值代表所述靶多核苷酸的一个或多个特征。

## 修饰的纳米孔,包含其的组合物及其应用

[0001] 相关申请的交叉引用

[0002] 本申请要求2015年12月8日根据35 U.S.C.§119(e)提交的美国临时申请号62/264,709的权益,其内容通过全文引用的方式并入本文。

### 技术领域

[0003] 本文提供了修饰的或突变形式的溶细胞素A(ClyA)以及包含其的组合物。还提供了使用所述修饰的或突变形式的ClyA和所述组合物例如表征目标分析物如目标多核苷酸的方法。

### 背景技术

[0004] 跨膜孔(如纳米孔)已被用于识别小分子或折叠蛋白,并监测单分子水平的化学或酶反应。DNA穿过重组在人工膜中的纳米孔的电泳移位在实际应用如DNA测序和生物标记识别中前景广阔。然而,双链或单链DNA穿过具有朝向内表面的带负电荷的氨基酸的纳米孔的移位是无效的。具体而言,在具有负内表面电荷和与溶液的德拜长度相当的半径的纳米孔中,由内部纳米孔壁上的双电荷层(EDL)产生的表面电势重叠,导致对于DNA进入纳米孔的大的静电能垒。因此,只能高离子强度溶液中或在不对称盐浓度下使用大纳米孔(例如10nm)或使用小纳米孔(例如~3.5nm)才能观察到DNA穿过这种纳米孔的移位。

### 发明内容

[0005] 本发明至少部分基于意外的发现,即虽然某些蛋白质纳米孔,例如溶细胞素A(ClyA)纳米孔具有带负电荷的狭窄收缩部(或抑制或降低移位效率的区域),但可以实现成功地捕获和移位带负电荷的分子或聚合物(例如,双链或单链DNA)穿过低离子强度溶液中的这种具有带负电荷的狭窄收缩部的蛋白质纳米孔,通过在所述蛋白质纳米孔(例如,ClyA纳米孔)的腔表面内引入正电荷例如带正电荷的氨基酸(例如,精氨酸)来在所述蛋白质纳米孔内捕获和定向带负电荷的分子或聚合物(例如,双链或单链DNA)来实现。例如,正电荷,如带正电荷的氨基酸(例如精氨酸)可被引入到具有带负电荷的狭窄收缩部的蛋白质纳米孔(例如,ClyA纳米孔)的靠近其开口(如用于使带负电荷的分子或聚合物进入的开口)和位于其中间部分的腔表面内。

[0006] 在某些实例中,使用ClyA-AS,由于其在平面脂质双分子层中的有利特性而选择的人造ClyA版本(engineered ClyA version),来产生如本文所述的修饰的ClyA纳米孔。ClyA-AS的内部电荷被重新排列以诱导纳米孔在生理离子强度下对的DNA的捕获。例如,修饰的ClyA纳米孔包含顺式开口,中间部分和反式开口,其中所述顺式开口的内表面包含第一带正电荷的氨基酸取代;所述中间部分的内表面包含第二带正电荷的氨基酸取代;并且所述反式开口包含带负电荷的收缩部。在一些情况下,第一带正电荷的氨基酸取代(例如,用精氨酸取代)可以位于顺式开口内,以允许捕获DNA进入所述修饰的ClyA纳米孔和/或第二带正电荷的氨基酸取代(例如用精氨酸取代)可以位于所述中间部分内以允许DNA移位穿

过所述修饰的Cl<sub>y</sub>A纳米孔。例如，第一带正电荷的氨基酸取代可以对应于Cl<sub>y</sub>A-AS的氨基酸序列中的S110R突变和/或第二带正电荷的氨基酸取代可以对应于Cl<sub>y</sub>A-AS的氨基酸序列中的D64R突变。

[0007] 因此，本发明的一个方面的特征在于修饰的Cl<sub>y</sub>A纳米孔，例如允许捕获带负电荷的聚合物进入所述修饰的Cl<sub>y</sub>A纳米孔和/或移位带负电荷的聚合物穿过所述修饰的Cl<sub>y</sub>A纳米孔。修饰的Cl<sub>y</sub>A纳米孔包含第一开口，中间部分，第二开口，和从第一开口穿过中间部分延伸至第二开口的腔，其中第一开口的腔表面包含第一带正电荷修饰（例如，第一带正电荷的氨基酸取代）并且中间部分的腔表面包含第二带正电荷修饰（例如，第二带正电荷的氨基酸取代）。第二开口的腔表面限定了一个带负电荷的收缩部。

[0008] 在本文所述的任何修饰的Cl<sub>y</sub>A纳米孔中，从第一带正电荷修饰（例如，第一带正电荷的氨基酸取代）到第二带正电荷修饰（例如，第二带正电荷的氨基酸取代）的腔内距离可以在约0.5nm至约10nm的范围内变化。在一些实施方案中，从第一带正电荷修饰（例如，第一带正电荷的氨基酸取代）到第一开口表面的腔内距离可以在约3nm至约7nm的范围内变化。

[0009] 可以使用任何形式的Cl<sub>y</sub>A来制备本文所述的修饰的Cl<sub>y</sub>A纳米孔。例如，野生型Cl<sub>y</sub>A（Cl<sub>y</sub>A-WT）和Cl<sub>y</sub>A-AS的氨基酸序列和编码它们的核苷酸序列是本领域已知的。因此，在一些实施方案中，修饰的Cl<sub>y</sub>A纳米孔可以包含亚单位多肽（subunit polypeptide），其具有与对应于野生型Cl<sub>y</sub>A的SEQ ID NO:1所示的氨基酸序列至少约80%（包括例如至少约85%，至少约90%，至少约95%或更高）同一性的氨基酸序列。或者，修饰的Cl<sub>y</sub>A纳米孔可以包含亚单位多肽，其具有与对应于Cl<sub>y</sub>A-AS的SEQ ID NO:2所示的氨基酸序列至少约80%（包括例如至少约85%，至少约90%，至少约95%或更高）同一性的氨基酸序列。在一些实施方案中，与包含第一和第二带正电荷的氨基酸取代的SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列相比，修饰的Cl<sub>y</sub>A纳米孔可以包含多达15个取代。

[0010] 在本文所述的任何修饰的Cl<sub>y</sub>A纳米孔中，第一带正电荷修饰（例如，第一带正电荷的氨基酸取代）可以位于第一开口内以允许捕获溶液中暴露于所述第一开口的带负电荷的聚合物（例如但不限于脱氧核糖核酸（DNA）如双链DNA或单链DNA）。例如，用正电荷（例如带正电荷的氨基酸）的取代可以发生在以下位置中的一个或多个：SEQ ID NO:1或SEQ ID NO:2的E106,D114,D121,D122,E129,E85,E78,D268,D267,D265,E258。

[0011] 在本文所述的任何修饰的Cl<sub>y</sub>A纳米孔中，第二带正电荷修饰（例如，第二带正电荷的氨基酸取代）可位于中间部分内以允许带负电荷的聚合物（例如，但不限于脱氧核糖核酸（DNA），如双链DNA或单链DNA）移位穿过所述孔的腔。例如，用正电荷（例如带正电荷的氨基酸）的取代可以发生在以下位置中的一个或多个：SEQ ID NO:1或SEQ ID NO:2的D74,D71,D64,E53,E161,D158,E46,E42,D41。

[0012] 第一和第二带正电荷修饰（例如，第一和第二带正电荷的取代）之间的距离优选为约0.5nm至约10nm。该距离可以在约3nm至约7nm之间。

[0013] 修饰的Cl<sub>y</sub>A纳米孔可以是同源多聚体（例如纳米孔内的所有亚基是相同的）或异源多聚体（例如纳米孔内的至少一个亚基与其他亚基不同）。修饰的Cl<sub>y</sub>A纳米孔可以包含任何数量的亚单位多肽，其足以形成足够大以允许靶聚合物（例如多核苷酸）通过的腔。在一些实施方案中，修饰的Cl<sub>y</sub>A纳米孔可以包含12个亚单位多肽或更多，包含例如13个亚单位多肽，和14个亚单位多肽，其中至少一个或多个亚单位多肽包含如本文所述的第一和第二

带正电荷的氨基酸取代。

[0014] 第一和第二带正电荷修饰(例如,第一和第二带正电荷的氨基酸取代)可以发生在纳米孔的所有亚基中。

[0015] 因此,本文还提供了修饰的ClyA纳米孔亚单位多肽和包含编码所述修饰的ClyA纳米孔亚单位多肽的核苷酸序列的多核苷酸。例如,修饰的ClyA纳米孔亚单位多肽包含的氨基酸序列与SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列具有至少约80%(包括例如至少约85%,至少约90%,至少约95%或更高)的同一性,并且其中所述氨基酸序列包含在SEQ ID NO:1或SEQ ID NO:2的106-78范围内位置处的第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)和在SEQ ID NO:1或SEQ ID NO:2的41-74范围内的位置处的第二带正电荷修饰(例如,第二带正电荷的氨基酸取代)。在一个实例中,第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)可位于SEQ ID NO:1或SEQ ID NO:2的第110位置处;和/或第二带正电荷修饰(例如,第二带正电荷的氨基酸取代)可位于SEQ ID NO:1或SEQ ID NO:2的第64位置处。第一和/或第二带正电荷的氨基酸取代的例子包括但不限于用精氨酸,组氨酸和赖氨酸中的一个的取代。

[0016] 同样在本发明的范围内的是组合物,比如用于表征靶聚合物,如带负电荷的靶聚合物,如靶多核苷酸的组合物。该组合物包含本文所述的任何修饰的ClyA纳米孔。该组合物可以进一步包含所述修饰的ClyA纳米孔位于其中的膜(例如人造膜)。所述组合物可以进一步包含低离子强度溶液,例如具有离子强度约100mM至约300mM或约150mM至约300mM的盐溶液。更一般地,所述盐溶液可具有约50mM至约1M的离子强度。在一些实施方案中,所述组合物可进一步包含多核苷酸结合蛋白,其可选择性地与修饰的ClyA纳米孔偶联。

[0017] 如本文所述的修饰的ClyA纳米孔和组合物可用于各种生物传感器或分析物检测应用,但不限于多核苷酸测序。所述分析物可以是蛋白质。在一个方面,本文描述了在低离子强度下使DNA移位的方法。该方法包括(a)在低离子强度溶液中,提供本文所述的任何一种修饰的ClyA纳米孔和膜(例如人造膜),其中修饰的ClyA纳米孔存在于所述膜中,使得修饰的ClyA纳米孔的顺式开口存在于低离子强度溶液的顺式侧中,并且修饰的ClyA纳米孔的反式开口存在于低离子强度溶液的反式侧中;(b)在低离子强度溶液的顺式侧提供DNA;和(c)跨修饰的ClyA纳米孔施加电势,使得DNA从修饰的ClyA纳米孔的顺式侧移位到反式侧。在一个实例中,低离子强度溶液可以是具有离子强度约150mM至约300mM的盐溶液(例如氯化钠溶液)。这种方法可用于表征多核苷酸(例如DNA或RNA)。

[0018] 因此,本文还提供了表征靶多核苷酸的方法。该方法包括(a)在低离子强度溶液(例如约150mM至约300mM)中提供本文所述的任何一种修饰的ClyA纳米孔和膜,其中修饰的ClyA纳米孔存在于所述膜中;(b)向步骤(a)的低离子强度溶液中加入靶多核苷酸;和(c)在跨所述纳米孔的施加电势期间,测量穿过所述修饰的ClyA纳米孔的离子流,其中离子流测量值指示靶多核苷酸的一个或多个特征。可以使用本文描述的方法测定的靶多核苷酸的特征的非限制性实例包括(i)靶多核苷酸的长度,(ii)靶多核苷酸的同一性,(iii)靶多核苷酸的序列,(iv)靶多核苷酸的二级结构,(v)靶多核苷酸是否是修饰的,并从而表征靶多核苷酸,及其任何组合。

[0019] 在本文所述的任何方面中,所述靶多核苷酸可以是单链DNA或双链DNA。

[0020] 在本文所述的任何方面中,该方法可进一步包括在低离子强度溶液中添加多核苷

酸结合蛋白,使得所述多核苷酸结合蛋白与靶多核苷酸结合并控制所述靶多核苷酸穿过修饰的ClyA纳米孔的移动。

[0021] 在下面的描述中阐述了本发明的一个或多个实施例的细节。根据以下附图和若干实施例的详细描述以及所附权利要求,本发明的其他特征或优点将变得显而易见。

## 附图说明

[0022] 以下附图形成本说明书的一部分并且被包括以进一步说明本发明的某些方面,通过参考这些附图中的一个或多个并结合本文给出的具体实施例的详细描述,可以更好地理解本发明的某些方面。

[0023] 图1用于DNA移位的人造ClyA纳米孔。版A) ClyA-AS和ClyA-RR纳米孔的横截面, ClyA-AS和ClyA-RR纳米孔嵌入通过由大肠杆菌ClyA结构 (PDB:2WCD, 90%序列同一性) 同源建模构造的脂质双分子层中。内部孔隙显示为表面表示,并且根据“真空”静电(负域为更暗阴影,正域为较浅阴影,PyMol) 带上阴影。测试的氨基酸取代在ClyA-AS(左)中显示。ClyA-RR孔包括在每个蛋白质组的位置110和64处的两个额外的精氨酸残基(右)。版B) ClyA-AS和ClyA-RR的电流与电压关系。版C) ssDNA (1a, 1 $\mu$ M) 和版D) dsDNA (1, 170nM) 在+70mV的生理离子强度下移位穿过ClyA-RR纳米孔。底部电流迹线显示DNA移位事件的放大图。使用2kHz低通贝塞尔滤波器在10kHz采集电流信号。缓冲液是150mM NaCl, 15mM Tris HCl, pH 7.5, 温度22 $^{\circ}$ C。

[0024] 图2在150mM NaCl溶液中形成DNA轮烷。版A) 在+50mV下通过将用中性抗生物素蛋白 (1.2 $\mu$ M, 单体) 复合的杂交dsDNA/ssDNA链1a/1c (1.0 $\mu$ M) 添加到顺式侧腔, 形成dsDNA轮烷。1a/1c在5'处含有31个碱基的单链的突出端, 其用于与1d (1.0 $\mu$ M) 杂交, 生物素化的ssDNA分子与1a/1c的ssDNA突出端互补。因此, 只有在1a/1c移位纳米孔时才形成纳米孔/DNA轮烷。当DNA占据ClyA的腔时, 开孔电流在正施加电势 (IRES+50=84 $\pm$ 7, 平均值 $\pm$ S.D., N=3) 下减小并且在负施加电势下增强 (IRES-50=1.11 $\pm$ 0.06, 平均值 $\pm$ S.D., N=3)。版B) 在+50mV下, 通过将用中性抗生物素蛋白 (1.2 $\mu$ M, 单体) 复合的5'生物素化的ssDNA链2a (1.0 $\mu$ M, 黑线) 添加到ClyA-RR纳米孔的顺式侧腔, 形成ssDNA/dsDNA杂交轮烷。将与2a的3'末端互补并与中性抗生物素蛋白 (1.2 $\mu$ M, 单体) 复合的第二个5'生物素化ssDNA分子2b (1.0 $\mu$ M) 添加到反式侧腔中。轮烷形成后, 逆转施加的电势至-50mV诱导电流增强 (IRES-50=1.16 $\pm$ 0.03, 平均值 $\pm$ S.D., N=3), 表明组装了杂交ssDNA/dsDNA。电流迹线的右侧显示了轮烷构型中游离ClyA-RR和ClyA-RR的电压关系 (IV曲线)。图2版A和B中的黑色和灰色线条, 表示两种轮烷的DNA构型。使用的缓冲液是15mM Tris HCl, pH7.5, 温度22 $^{\circ}$ C。DNA序列如表3所示。

[0025] 图3DNA移位和穿过 (threading) 的离子强度依赖性。版A) 对于dsDNA (圆圈) 与ssDNA (三角形) 的移位频率的德拜强度的依赖性。dsDNA移位事件的频率符合线性回归 ( $R^2=0.98$ ), 而ssDNA的频率更适合单一指数 ( $R^2=0.99$ ) 而非线性回归 ( $R^2=0.78$ )。版B) dsDNA (三角形) 与中性抗生物素蛋白:dsDNA复合物 (圆圈) 的剩余电流阻断 (blockades) 对溶液德拜长度的依赖性。这些线条表示线性回归。版C) 与版B相同, 但用于ssDNA。版D) DNA穿过的离子强度依赖性。在+70mV施加电势下, ssDNA (1a, 1 $\mu$ M) 初始添加至ClyA-RR的顺式侧诱导ClyA-RR开孔电流的快速电流阻断。随后添加的中性抗生物素蛋白 (1.2 $\mu$ M, 顺式) 在150和300mM NaCl溶液中诱导持久的电流阻断, 这很可能是由于ssDNA穿过引起的。这在1M NaCl溶

液(或更高)中没有观察到,其阻断仍然是短暂的。由于dsDNA穿过,进一步添加的互补ssDNA (1a, 1 $\mu$ M, 顺式)在所有离子强度下诱导永久性阻断。在每次永久性DNA捕获事件之后,通过手动反转电势至-70mV,开孔再生。高于和低于开孔电流水平的尖峰代表电势反转之后的电容性瞬变。在15mM Tris HCl, pH7.5, 22 $^{\circ}$ C下进行电记录。通过应用10kHz低通贝塞尔滤波器并使用20 $\mu$ s (50kHz) 采样率来记录数据,并列于表7中。在150mM NaCl, 额外的数字2kHz低通贝塞尔滤波器被应用于电流迹线。

[0026] 图4单向DNA移位穿过ClyA-RR纳米孔。版A) 在150mM NaCl溶液中,在ClyA-RR纳米孔的顺式和反式两侧添加3 $\mu$ M dsDNA 1仅在正施加电势下诱导瞬时电流阻断(灰色垂直线)。版B) 在1.0M NaCl溶液中,在两种施加电势下观察到DNA阻断。DNA诱导的阻断显示为灰色垂直线。在21秒内施加的电势自动从+70变为-70mV(版A) 或从+100变为-100mV(版B)。在15mM Tris HCl, pH7.5, 22 $^{\circ}$ C下进行电记录。通过应用2kHz(版A) 和10kHz(版B) 低通贝塞尔滤波器并使用100 $\mu$ s (10kHz, 版A) 和50kHz(版B) 采样率来记录数据。

[0027] 图5dsDNA和ssDNA移位穿过ClyA纳米孔的机理。版A) dsDNA移位。(1) dsDNA最初与纳米孔的顺式入口处的电荷相互作用。(2) dsDNA穿入纳米孔内部,在此其与第二人工电荷相互作用。两种电荷对于使DNA对齐以移位穿过带负电荷的反式收缩部非常重要。(3) 然后dsDNA可以移位并且然后(4) 离开所述孔。版B) (1) 顺式入口处的额外电荷介导纳米孔内的DNA的有效捕获。(2) ssDNA最有可能以卷曲结构进入顺式内腔。(3) 为了移位反式收缩部, ssDNA需要展开,然后在纳米孔外重新卷曲。(4) DNA离开纳米孔。DNA分子和纳米孔按比例绘制。R<sub>g</sub>表示ssDNA的回转半径。在实验条件下,dsDNA是刚性棒,并且ssDNA是具有~6nm的回转半径的卷曲结构。

[0028] 图6DNA在0.15M NaCl溶液中从ClyA纳米孔的顺式侧移位。对于每个指示的突变体(版A-G),报告了在正VG施加电势(表5)下,在将1 $\mu$ M生物素化ssDNA (1a, 表3)添加到顺式侧腔之前和之后的IV关系(在21s内和10mV电压阶跃下从+100到-100mV的电压斜升)和代表性电流迹线。随后向顺式溶液中添加1.2 $\mu$ M中性抗生物素蛋白(单体)和1 $\mu$ M互补ssDNA (1b表1)后,还显示了多种电流迹线。在0.15M NaCl, 15mM Tris-HCl, pH 7.5, 22 $^{\circ}$ C下进行电记录。通过应用2kHz低通贝塞尔滤波器并使用100 $\mu$ s (10kHz) 采样率来记录数据。

[0029] 图7ssDNA移位的离子强度依赖性。版A-F显示不同盐浓度或离子强度的数据。(左侧) 代表性电流迹线,其显示在不同的NaCl浓度,+70mV下将1 $\mu$ M生物素化ssDNA (1a, 表3)添加至孔的顺式侧之前和之后的ClyA-RR纳米孔的开孔电流。右侧的直方图表示单个ssDNA移位事件的停留时间( $t_{\text{OFF}}$ , 左侧直方图)和事件间时间( $t_{\text{ON}}$ , 右侧直方图)。使用0.05ms的数据采集阈值,通过使用Clampfit软件(分子装置)中的“单通道搜索”功能分别收集单个 $t_{\text{OFF}}$ 和事件间时间 $t_{\text{ON}}$ 事件。从累积直方图的单指数拟合计算平均DNA移位停留时间 $t_{\text{OFF}}$ 。事件间时间 $t_{\text{ON}}$ 是使用对数箱(基数10)由直方图的指数对数概率拟合来计算。在15mM Tris-HCl, pH 7.5, 22 $^{\circ}$ C下进行电记录。通过应用10kHz低通贝塞尔滤波器并使用20 $\mu$ s (50kHz) 采样率来记录数据。在0.15M NaCl溶液中使用额外的2kHz低通贝塞尔滤波器收集数据。

[0030] 图8dsDNA移位的离子强度依赖性。版A-E显示不同盐浓度或离子强度的数据。电流迹线显示在+70mV和指定的NaCl浓度下,在将170nM dsDNA (1, 表3)添加到孔的顺式侧之前(左)和之后(右)的ClyA-RR纳米孔的开孔电流。右侧的直方图代表单个ssDNA移位事件的停留时间( $t_{\text{OFF}}$ , 左侧直方图)和事件间时间( $t_{\text{ON}}$ , 右侧直方图)。使用0.05ms的数据采集阈值,通

过使用Clampfit软件(分子装置)中的“单通道搜索”功能分别收集各个 $t_{\text{OFF}}$ 和事件间时间 $t_{\text{ON}}$ 事件。从累积直方图的双指数拟合计算平均DNA移位停留时间 $t_{\text{off}}$ 。事件间时间 $t_{\text{on}}$ 是使用对数箱(基数10)的直方图的指数对数概率拟合计算。在15mM Tris-HCl, pH 7.5, 22°C下进行电记录。通过应用10kHz低通贝塞尔滤波器并使用20 $\mu$ s (50kHz) 采样率来记录数据。在0.15M NaCl溶液中使用额外的2kHz低通贝塞尔滤波器用于收集数据。

[0031] 图9在1M NaCl下从反式侧的DNA轮烷的形成。版A) 在-70mV施加电势下, 通过将中性抗生物素蛋白(1.2 $\mu$ M, 单体) 复合的杂交dsDNA/ssDNA链T1d (1a和1c, 1.0 $\mu$ M, 表3, 显示为在电流迹线上方的黑线) 添加到反式纳米孔腔形成dsDNA轮烷。将与T1d的突出端互补的3'生物素化的ssDNA分子1d (1.0 $\mu$ M, 表3, 对应于电流迹线上方的灰色线) 添加到顺式侧腔中。由于纳米孔/DNA轮烷仅在T1d移位穿过纳米孔与1杂交时形成, 因此该实验证明DNA从ClyA的顺式穿过至反式的移位。在-70mV时, 穿过的DNA的阻断孔电流为 $64 \pm 2.0$ , 平均值 $\pm$  S.D.,  $N=3$ 。在轮烷形成后, 施加的电势逆转至+70mV显示了阻断的孔电流 ( $I_{\text{RES}+70} = 73 \pm 0.5$ , 平均值 $\pm$  S.D.,  $N=3$ ), 表明dsDNA占据纳米孔。版B) 轮烷构象中ClyA-RR和ClyA-RR的IV关系。

[0032] 图10人造的用于观察0.15M NaCl溶液中从反式侧的DNA移位的孔。对于版A-I中指示的每个突变体, 报告了在正VG施加电势下, 在加入1 $\mu$ M生物素化ssDNA (1a, 表3) 到反式侧腔的之前和之后的IV关系(在21s内和10mV电压阶跃下从+100到-100mV的电压斜升) 和代表电流迹线。在随后将1.2 $\mu$ M中性抗生物素蛋白(单体) 和1 $\mu$ M互补ssDNA (1b表1) 添加到反式溶液之后, 还显示了多条电流迹线。虽然在将DNA添加到反式侧腔之后, 1yA-3R-E7S显示出电流阻断, 但是不能形成轮烷, 表明阻断不是由于DNA的移位引起的。在0.15M NaCl, 15mM Tris-HCl, pH 7.5, 22°C下进行电记录。通过应用2kHz低通贝塞尔滤波器并使用100 $\mu$ s (10kHz) 采样率来记录数据。

[0033] 图11人造的用于DNA移位的ClyA纳米孔。版A) 嵌入脂质双分子层中的ClyA-AS (左) 和ClyA-RR (右) 纳米孔的横截面, 所述脂质双分子层使用VMD和NAMD (PDB: 2WCD, 90%序列同一性), 通过来自大肠杆菌ClyA结构的同源建模构建。使用溶剂可及表面积显示内部孔腔(所述溶剂可及表面积是由PyMOL (版本1.8Schrödinger, LLC) 计算) 并根据150mM NaCl溶液中的静电势带上阴影(由自适应泊松-波尔兹曼解算器 (APBS) 计算)。阴影区域对应于负和正电势(范围为-2至+2kBT/e或-51.4至+51.4mV)。版B) 在150mM NaCl浓度下在ClyA-AS和ClyA-RR纳米孔中心处的静电势。

[0034] 图12在+50mV, 150mM NaCl溶液中形成DNA轮烷。版A) 通过将1a/1c (1.0 $\mu$ M, 黑线) 和1d (1.0 $\mu$ M, 灰线) 分别添加至顺式和反式侧腔形成dsDNA轮烷。在两种溶液中也添加中性抗生物素蛋白 (NA, 0.3 $\mu$ M, 四聚体)。版B) 通过将5'-生物素化的ssDNA链2a (1.0 $\mu$ M, 黑色线) 添加至顺式侧腔和将与2a的3'末端互补的5'-生物素化的ssDNA分子 (2b, 1.0 $\mu$ M, 灰线) 添加到反式侧腔形成ssDNA/dsDNA杂交轮烷。NA (0.3 $\mu$ M, 四聚体) 存在于两侧。电流迹线的右手边的曲线显示了轮烷构象中的ClyA-RR和ClyA-RR的电压关系 (I-V曲线)。实验在22°C下含有150mM NaCl和15mM Tris-HCl (pH 7.5) 的缓冲液中进行。DNA序列显示在表3中。

[0035] 图13+70mV下DNA移位和穿过的离子强度依赖性。版A-B) 每1 $\mu$ M DNA的dsDNA (版A) 和ssDNA (版B) 移位的频率的德拜长度依赖性。(版A) 中的虚线描述了扩散限制过程 (diffusion-limited process) 的移位频率的理论预测。(版B) 中的线是指示能垒限制过程

(barrier-limited process)的指数回归。

[0036] 图14dsDNA和ssDNA移位穿过ClyA-RR纳米孔的机理。版A) dsDNA移位是扩散受限的。(i) 在实验条件下dsDNA为刚性棒,通过电场线对齐并以确定的方向进入纳米孔。(ii) dsDNA穿入纳米孔内部,在此其与有人造电荷的第二层相互作用。(iii) dsDNA然后可以移动穿过收缩部和(iv) 离开孔。纳米孔的顺式口处的电荷有助于最初的捕获。版B) ssDNA移位是反应受限的。(i) ssDNA具有回转半径( $R_g \approx 6\text{nm}$ )的卷曲结构,其大约是纳米孔半径的两倍。(ii) ssDNA尚未在孔中,并且它搜索入口。(iii) ssDNA的一端发现顺式侧腔的入口并开始展开。由于进入纳米孔存在熵能垒,因此在成功的移位事件之前可进行多次尝试。(四) 为了移动穿过收缩部,ssDNA需要完全展开。(v) DNA离开纳米孔然后重新卷曲。顺式入口处的额外电荷很可能介导纳米孔内的DNA的有效捕获。DNA分子和纳米孔按比例绘制。

[0037] 图15DNA穿过的离子强度依赖性。首先将ssDNA (1a,  $1.0\mu\text{M}$ ) 添加至ClyA-RR的顺式侧,然后加入中性抗生物素蛋白 (NA,  $0.3\mu\text{M}$ , 顺式),最后加入互补ssDNA (1b,  $1\mu\text{M}$ , 顺式)。在150和500mM NaCl溶液中,ssDNA:NA复合物诱导持久电流阻断,这很可能是由于ssDNA的穿过导致的。在1.0M NaCl溶液(或更高)中,ssDNA:NA阻断是瞬时的,表明ssDNA不能完全穿过孔。所有离子强度下dsDNA:NA复合物诱导永久性阻断。在开孔电流水平之上和之下的尖峰代表在用于将纳米孔从DNA中释放出来的手动电势反转之后的电容性瞬变。在15mM Tris-HCl, pH 7.5,  $22^\circ\text{C}$ 下进行电记录。

[0038] 图16ssDNA移位穿过ClyA-RR纳米孔的离子强度依赖性。版A-F显示不同盐浓度或离子强度的数据。电流迹线显示在不同的NaCl浓度下,在+70mV下将 $1.0\mu\text{M}$ 生物素化ssDNA (1a, 表3)添加至孔的顺式侧之前和之后的ClyA-RR的开孔电流。迹线右侧的直方图表示dsDNA移位事件的停留时间(左直方图,常规分箱(binning)单指数拟合)和事件间时间(右直方图,对数基数(logarithmic base) 10,指数对数概率拟合)。散射图代表电流相对于停留时间。在15mM Tris-HCl, pH 7.5,  $22^\circ\text{C}$ 下进行电记录。通过应用10kHz低通贝塞尔滤波器并使用 $20\mu\text{s}$  (50kHz) 采样率来记录数据。在0.15M NaCl溶液中使用额外的2kHz低通贝塞尔滤波器用于数据收集。

[0039] 图17dsDNA移位穿过ClyA-RR纳米孔的离子强度依赖性。版A-E显示不同盐浓度或离子强度的数据。电流迹线显示在不同的NaCl浓度下,在+70mV下将140-170nM的生物素化的dsDNA (1, 表3)添加至孔的顺式侧之前和之后的ClyA-RR的开孔电流。迹线右侧的直方图表示dsDNA移位事件的停留时间(左直方图,常规分箱单指数拟合)和事件间时间(右直方图,对数基数10,指数对数概率拟合)。散射图代表电流相对于停留时间。在15mM Tris-HCl, pH 7.5,  $22^\circ\text{C}$ 下进行电记录。通过使用10kHz低通贝塞尔滤波器并使用50kHz采样率来记录数据。在0.15M NaCl溶液中使用额外的2kHz低通贝塞尔滤波器用于数据收集。

[0040] 图18以1kHz过滤频率的DNA移位的离子强度依赖性。由电流迹线确定(版A) dsDNA和(版B) ssDNA的事件频率的盐依赖性,使用1kHz数字高斯滤波器(Clampfit, Molecular Devices)过滤。线条显示线性回归拟合(版A)和指数回归拟合(版B)。

[0041] 图19作用于纳米孔附近的ssDNA的熵的力和电泳力。ssDNA具有卷曲形状,并被预期通过能垒穿越(barrier crossing) (反应限制过程) 被孔捕获。能垒源于孔附近的熵起源的排斥力,其作用于有吸引力的电泳力的顶部。这两种贡献的自由能用细线表示,而粗线是两者的和(方程(15))。该图的上半部分分别显示了以反应坐标 $r_a$ 和 $r_b$ 为特征的ssDNA的两种

特征构象。构象 (b) 具有较低的熵并且对应于靠近能垒顶部的状态。

[0042] 图20显示ClyA的结构和表示为A的顺式部分,表示为B的中间部分,表示为C的反式部分。带负电荷的氨基酸D和E显示在图的左侧(连同极性不带电的氨基酸S和Q)。可以在A中进行一个或多个极性不带电的氨基酸或带负电荷的氨基酸的取代,可以在B中进行一个或多个带负电荷的氨基酸的取代。包含许多带负电荷的氨基酸的区域C可以保持原样,不被中性或带正电荷的氨基酸取代。

### 具体实施方式

[0043] 虽然跨膜孔(例如,蛋白质纳米孔或固态纳米孔)可用作传感器来检测或表征生物聚合物,但生物聚合物(例如多核苷酸)移位穿过低离子强度下(例如,约150mM至约300mM)的某些纳米孔可能具有挑战。具体而言,具有负内表面电荷部分且半径与带负电荷的生物聚合物的尺寸(例如,对于dsDNA的B型为~2.2nm,对于ssDNA为~1nm)相当的纳米孔,对于低离子强度下的带负电荷的生物聚合物进入所述纳米孔,可以产生大的静电能垒。因此,需要改造跨膜纳米孔,其允许更有效地捕获和/或移位带负电荷的生物聚合物,例如多核苷酸穿过纳米孔,其可用于例如多核苷酸图谱或测序的实际应用。

[0044] 本发明至少部分基于意外的发现,即可以将正电荷引入到跨膜纳米孔例如溶细胞素A (ClyA) 的腔表面中的某些位置处,以克服对于DNA移位穿过带负电荷的窄收缩部(例如,尺寸约为3.3nm)的熵和静电能垒。例如,发现在ClyA纳米孔的较宽入口(顺式侧)和中间部分引入正电荷(例如,带正电荷的氨基酸,例如精氨酸)足以“抓住”并定向DNA(例如双链或单链的)用于有效的电泳驱动滑动穿过窄且带负电荷的反式收缩部,即使在没有对带负电荷的反式收缩部本身进行任何修饰的情况下。进一步地,发现这种修饰允许DNA在低离子强度下例如低至50mM的移位。原则上,所述修饰允许在甚至低于50mM的离子强度下实施本文所述的任何方面的方法。然而较低的离子强度可能导致相应较低的离子电流,并因此在某些情况下可能不合意。没有这种修饰,单链或双链DNA穿过纳米孔的移位仅在2.0M离子强度以上才观察到。

[0045] 因此,在一些方面,本发明提供修饰的ClyA纳米孔亚单位多肽(例如,用于形成修饰的ClyA纳米孔)和包含其的纳米孔。如本文所述的修饰的ClyA纳米孔可以用于多种实际应用,例如表征多核苷酸。因此,本文描述的还有用于表征多核苷酸例如双链或单链多核苷酸的方法和组合物。本文描述的方法和组合物提供在生理离子强度(例如50mM-300mM)或低离子强度(例如,小于2M或小于1M)下的双链或单链多核苷酸的有效移位。

[0046] 本文所述的修饰的ClyA纳米孔和方法允许多核苷酸的单向移位,即多核苷酸不能以反式至顺式的方向进入和穿过纳米孔。这能够例如实现多核苷酸(例如,DNA)以顺式到反式方向的过滤。

[0047] 还预期其他纳米结构,其具有与ClyA纳米孔类似的纳米孔结构(例如,圆柱形腔,在顺式开口处有较大直径(例如5-7nm)和在反式开口处有负电荷的较窄收缩部(例如3-4nm直径),可以采用类似的修饰策略,以允许在低离子强度溶液下的DNA移位。

[0048] 修饰的ClyA纳米亚单位多肽

[0049] 本发明的一个方面提供了修饰的ClyA纳米孔亚单位多肽。修饰的ClyA纳米孔亚单位多肽是其序列从参考ClyA氨基酸序列变化而来的多肽。修饰的ClyA纳米孔亚单位多肽的

氨基酸序列包含 (i) 形成顺式开口的氨基酸序列, (ii) 形成中间部分的氨基酸序列, 和 (iii) 形成反式开口的氨基酸序列。当修饰的ClyA纳米孔亚单位多肽与其他亚单位多肽相互作用以在膜中形成纳米孔时, 形成顺式开口的氨基酸序列是所述氨基酸序列的形成部分的纳米孔的顺式开口的部分。当修饰的ClyA纳米孔亚单位多肽与其他亚单位多肽相互作用以在膜中形成纳米孔时, 形成中间部分的氨基酸序列是所述氨基酸序列的形成部分的纳米孔的中间部分的部分。当修饰的ClyA纳米孔亚单位多肽与其他亚单位多肽相互作用以在膜中形成纳米孔时, 形成反式开口的氨基酸序列是所述氨基酸序列的形成部分的纳米孔的反式开口的部分。鉴定ClyA氨基酸序列中对应于ClyA纳米孔的顺式部分, 中间部分和反式部分的部分的方法在本领域中是已知的, 并且也在实施例中描述。例如, 纳米孔, 其嵌入到膜中的部分, 可以通过使用VMD从已知ClyA结构进行同源性建模而构造, VMD例如Humphrey等人描述的, "VMD: Visual Molecular Dynamics" J. Mol. Graphics (1996) 14:33-38和NAMD例如Phillips等人描述的 "Scalable Molecular Dynamics with NAMD" J. Comput. Chem. (2005) 26:1781-1802。参阅例如图1A。

[0050] 如本文所用, 术语“参考ClyA氨基酸序列”是指ClyA纳米孔亚单位的已知氨基酸序列。本领域已知各种形式的ClyA纳米孔亚单位, 包括例如但不限于ClyA野生型 (ClyA-WT), ClyA-SS, ClyA-CS和ClyA-AS。参见例如Soskine等人 "Tuning the size and properties of ClyA nanopores assisted by directed evolution" J Am Chem Soc. (2013) 135: 13456-13463, 其描述了相对于ClyA-WT的ClyA-SS, ClyA-CS和ClyA-AS中的不同突变, 以及制备它们的方法。W02016/166232和W02014/153625中描述的任何ClyA氨基酸序列也可以用作参考ClyA氨基酸序列。在一个实施方案中, 参考ClyA氨基酸序列是如SEQ ID NO:1所示的ClyA-WT的氨基酸序列。在一个实施方案中, 参考氨基酸是如SEQ ID NO:2所示的ClyA-AS的氨基酸序列, 与SEQ ID NO:1所示的ClyA-WT的氨基酸序列相比, 其包含以下突变: C87A, L99Q, E103G, F166Y, I203V, C285S, K294R。在一些实施方案中, 与ClyA-WT的氨基酸序列相比, ClyA-AS的氨基酸序列可以进一步包括H307Y。

[0051] 在一些实施方案中, 修饰的ClyA纳米孔亚单位多肽包含与参照ClyA氨基酸序列至少约80% (包括例如至少约85%, 至少约90%, 至少约95%, 至少约96%, 至少约97%, 至少约98%, 至少约99%或更高) 同一性的氨基酸序列。本领域的标准方法可用于确定同一性。例如, UWGCG软件包提供可用于计算同一性的BESTFIT程序, 例如用于其默认设置 (Devereux et al (1984) Nucleic Acids Research 12, p387-395)。PILEUP和BLAST算法可用于计算同一性或比对顺序 (例如鉴定等价残基或相应序列 (通常在其默认设置下)), 例如如Altschul S.F. (1993) J Mol Evol 36:290-300; Altschul, S.F et al (1990) J Mol Biol 215:403-10中描述的。执行BLAST分析的软件可通过国家生物技术信息中心 (<http://www.ncbi.nlm.nih.gov/>) 公开获得。

[0052] 修饰的ClyA纳米孔亚单位多肽的氨基酸序列包含 (i) 在形成顺式开口的氨基酸序列内的位置处的第一带正电荷修饰 (例如, 第一带正电荷的氨基酸取代) 和 (ii) 在形成中间部分的氨基酸序列内的位置处的第二带正电荷修饰 (例如第二带正电荷的氨基酸取代)。选择第一和第二带正电荷修饰 (例如第一和第二带正电荷的取代) 以提供更高效率的带负电荷的聚合物 (例如多核苷酸, 例如双链或单链DNA) 的捕获和/或移位穿过纳米孔, 与参考ClyA氨基酸序列相比。

[0053] 在一个实施方案中,第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)可以位于SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列的位置110处。在一些实施方案中,用正电荷(例如,带正电荷的氨基酸)的取代可以发生在以下位置的一个或多个处:SEQ ID NO:1或SEQ ID NO:2的E106,D114,D121,D122,E129,E85,E78,D268,D267,D265,E258。在一些实施方案中,C1yA氨基酸序列(例如,如SEQ ID NO 1或2所示)可以被修饰或设计为包括在其N-末端的额外氨基酸“MI”。

[0054] 在一个实施方案中,第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)可以位于SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列的位置64处。在一些实施方案中,用正电荷(例如带正电荷的氨基酸)的取代可以发生在SEQ ID NO:1或SEQ ID NO:2的以下位置的一个或多个处:SEQ ID NO:1或SEQ ID NO:2的D74,D71,D64,E53,E161,D158,E46,E42,D41。

[0055] 如本文所用的术语“带正电荷的氨基酸取代”是指对参考氨基酸的修饰,例如,如在pH7.0-8.0(例如,在pH8.0)和室温例如在20-25℃下检测的,增加参考氨基酸的净正电荷或减少净负电荷。例如,带正电荷的氨基酸取代可以包括但不限于(i)用带较少负电荷的氨基酸,中性氨基酸或带正电荷的氨基酸取代带负电的氨基酸,(ii)用带正电荷的氨基酸取代中性氨基酸,或(iii)用带较多正电荷的氨基酸取代带正电荷的氨基酸。在一些实施方案中,带正电荷的氨基酸取代包括带负电荷的氨基酸的缺失或带正电荷的氨基酸的添加。在一些实施方案中,带正电荷的氨基酸取代可以包括对一个或多个带负电荷的氨基酸的一个或多个化学修饰,其中和它们的负电荷。例如,所述一个或多个带负电荷的氨基酸可以与碳二亚胺反应。

[0056] 带正电荷的氨基酸是具有比溶液pH高的等电点(pI)的氨基酸,从而溶液中的氨基酸携带净正电荷。例如,在pH7.0-8.0(例如,在pH8.0)和室温(例如在20-25℃)检测到的带正电荷的氨基酸的实例包括但不限于精氨酸(R),组氨酸(H)和赖氨酸(K)。带负电荷的氨基酸是具有比溶液pH低的pI的氨基酸,从而溶液中的氨基酸携带净负电荷。如在pH7.0-8.0(例如,在pH8.0)和在室温(例如在20-25℃)检测到的带负电荷的氨基酸的实例包括但不限于天冬氨酸(D),谷氨酸(E),丝氨酸(S),谷氨酰胺(Q)。中性氨基酸是具有与溶液的pH相同的等电点(pI)的氨基酸,从而溶液中的氨基酸不带净电荷。氨基酸的pI值是本领域已知的。通过比较感兴趣的氨基酸的pI值与溶液的pH值,本领域普通技术人员将容易地确定是否溶液中存在的氨基酸是带正电的氨基酸,中性氨基酸或带负电荷的氨基酸。如本文所用,术语“氨基酸”可以是天然存在的或合成的氨基酸。

[0057] 在一些实施方案中,第一和/或第二带正电荷的氨基酸取代,例如,如在pH7.0-8.0(例如,在pH8.0)和室温(例如在20-25℃)检测到的,包括但不限于用精氨酸,组氨酸和赖氨酸中的一个取代参考氨基酸。

[0058] 在一些实施方案中,第一带正电荷的氨基酸取代是S110R,其中位置110对应于SEQ ID NO:1或SEQ ID NO:2的氨基酸110。

[0059] 在一些实施方案中,第二带正电荷的氨基酸取代是D64R,其中位置64对应于SEQ ID NO:1或SEQ ID NO:2的氨基酸64。

[0060] 除了本文所述的第一和第二带正电荷的氨基酸取代之外,可以对参考C1yA氨基酸序列进行氨基酸取代,例如高达1,2,3,4,5,6,7,8,9,10,15,20或30个氨基酸的取代。保守取代用具有相似化学结构,相似化学性质或相似侧链体积的其它氨基酸取代氨基酸。引入

的氨基酸可以具有与它们要取代的氨基酸具有相似的极性,亲水性,疏水性,碱性,酸性,中性或带电性。或者,所述保守取代可在预先存在芳香族或脂肪族氨基酸的位置引入其他芳香族或脂肪族氨基酸。保守氨基酸的改变在本领域公知的,并且可以根据如下表A中定义的20种主要氨基酸的特性进行选择。如果氨基酸具有相似的极性,这也可以通过参考表A中氨基酸侧链的亲水性大小来确定。

[0061] 表A-氨基酸的化学性质

[0062]

Ala	脂肪族,疏水,中性	Met	疏水,中性
Cys	极性,疏水,中性	Asn	极性,亲水,中性
Asp	极性,亲水,带电荷(-)	Pro	疏水,中性
Glu	极性,亲水,带电荷(-)	Gln	极性,亲水,中性
Phe	芳香族,疏水,中性	Arg	极性,亲水,带电荷(+)
Gly	脂肪族,中性	Ser	极性,亲水性,中性
His	芳香族,极性,亲水性,带电荷(+)	Thr	极性,亲水性,中性
Ile	脂肪族,疏水,中性	Val	脂肪族,疏水,中性
Lys	极性,亲水,带电荷(+)	Trp	芳香族,疏水,中性
Leu	脂肪族,疏水,中性	Tyr	芳香族,极性,疏水

[0063] 表B-亲水性大小

侧链 亲水性值

[0064]

Ile	4.5
Val	4.2
Leu	3.8
Phe	2.8
Cys	2.5
Met	1.9
Ala	1.8
Gly	-0.4

	Thr	-0.7
	Ser	-0.8
	Trp	-0.9
	Tyr	-1.3
	Pro	-1.6
[0065]	His	-3.2
	Glu	-3.5
	Gln	-3.5
	Asp	-3.5
	Asn	-3.5
	Lys	-3.9
	Arg	-4.5

[0066] SEQ ID NO:1或SEQ ID NO:2的氨基酸序列的一个或多个氨基酸残基可另外从上述多肽中缺失。可缺失高达1,2,3,4,5,10,20或30个或更多个氨基酸残基。

[0067] 一个或多个氨基酸可以替代地或另外地添加到上述多肽。可以在SEQ ID NO:1或2或多肽变体或其片段的氨基酸序列的氨基末端或羧基末端提供延伸。所述延伸可以很短,例如1至10个氨基酸长度。或者,延伸可以更长,例如长达50或100个氨基酸。载体蛋白可以与氨基酸序列例如修饰的ClyA纳米孔亚单位多肽的氨基酸序列融合。其他融合蛋白在下文中更详细地讨论。

[0068] 修饰氨基酸(例如通过取代,添加或缺失)的方法是本领域公知的。例如,通过在修饰的ClyA纳米孔亚单位多肽的多核苷酸中的相关位置处用编码靶氨基酸的密码子取代参考氨基酸的密码子,参考氨基酸可被靶氨基酸取代。然后可以按如下所述表达多核苷酸。如果氨基酸是非天然存在的氨基酸,其可以通过在IVTT系统中包含合成的氨酰-tRNA被引入,所述IVTT系统用于表达修饰的ClyA纳米孔亚单位多肽。或者,其可以通过在大肠杆菌中表达修饰的ClyA纳米孔亚单位多肽被引入,在特定氨基酸的合成的(即,非天然存在的)类似物存在的情况下,所述大肠杆菌对于特定氨基酸是营养缺陷的。如果使用部分肽合成制备修饰的ClyA纳米孔亚单位多肽,则它们也可以通过裸露连接(naked ligation)制备。

[0069] 在一些实施方案中,修饰的ClyA纳米孔亚单位多肽的形成反式开口的氨基酸序列可带有净负电荷,例如,如在pH7.0-8.0(例如,在pH8.0)和室温(例如在20-25℃)检测到的,所述净负电荷与参考ClyA氨基酸序列的形成反式开口部分的净负电荷相当(例如,10%以内,5%以内,4%以内,3%以内,2%以内,1%以内或更低)。例如,在一些实施方案中,修饰的ClyA纳米孔亚单位多肽的形成反式开口的氨基酸序列可以与参考ClyA氨基酸序列如SEQ ID NO:1或SEQ ID NO:2所示的形成反式开口相应部分具有至少约95%或更高(包括例如至少约96%,至少约97%,至少约98%,至少约99%或至多100%)的同一性。在一个实施方案中,修饰的ClyA纳米孔亚单位多肽的形成反式开口的氨基酸序列与SEQ ID NO:2中示出的氨基酸序列的形成反式开口相应部分具有100%的同一性。

[0070] 本文所述的修饰的ClyA纳米孔亚单位多肽可用于形成如本文所述的同源多聚体纳米孔或异源多聚体纳米孔。因此,在一些实施方案中,修饰的ClyA纳米孔亚单位多肽保留与其他亚单位多肽形成纳米孔的能力。评估修饰的单体形成纳米孔的能力的方法是本领域

公知的。例如,修饰的ClyA纳米孔亚单位多肽可以与其他适当的亚单位一起插入到两亲性层中,并且其寡聚化形成孔的能力可以被测定。对于将亚单位插入膜例如两亲性层中的方法,是本领域已知的。例如,亚基(即亚单位)可以以纯化形式悬浮在含有三嵌段共聚物膜的溶液中,从而其扩散到膜并通过与膜结合并组装成功能状态而插入。或者,可以使用M.A.Holden,H.Bayley.J.Am.Chem.Soc.2005,127,6502-6503和国际申请号PCT/GB2006/001057(公开号为W0 2006/100484)所述的“拾取和放置”方法将亚基直接插入到膜中。

[0071] 修饰的ClyA纳米孔亚单位多肽可以含有非特异性修饰,只要它们不干扰纳米孔的形成。许多非特异性侧链修饰在本领域中是已知的并且可以对氨基酸的侧链进行。这种修饰包括例如通过与醛反应然后用 $\text{NaBH}_4$ 还原进行氨基酸的还原性烷基化,用甲基乙酰亚胺酯(methylacetimidate)进行脒基化,或用乙酸酐进行酰化。

[0072] 修饰的ClyA纳米孔亚单位多肽可以使用本领域已知的标准方法制备。修饰的ClyA纳米孔亚单位多肽可以通过合成或通过重组手段制备。例如,修饰的ClyA纳米孔亚单位多肽可以通过体外翻译和转录(IVTT)来合成。国际申请号PCT/GB09/001690(公开为W02010/004273),PCT/GB09/001679(公开为W02010/004265)或PCT/GB10/000133(公开为W0 2010/086603)描述了合适的用于制备孔和修饰的ClyA纳米孔亚单位多肽的方法。

[0073] 如本文所述的修饰的ClyA纳米孔亚单位多肽可以使用D-氨基酸来制备。例如,如本文所述的修饰的ClyA纳米孔亚单位多肽可以包含L-氨基酸和D-氨基酸的混合物。这在制备这种蛋白质或肽领域是常规的。

[0074] 在一些实施方案中,修饰的ClyA纳米孔亚单位多肽可以被化学修饰。修饰的ClyA纳米孔亚单位多肽可以以任何方式和在任何位点进行化学修饰。例如,修饰的ClyA纳米孔亚单位多肽可以通过连接染料或荧光团而被化学修饰。在一些实施方案中,修饰的ClyA纳米孔亚单位多肽可以通过以下方式被化学修饰:将分子连接至一个或多个半胱氨酸(半胱氨酸键)、将分子连接至一个或多个赖氨酸、将分子连接至一个或多个非天然的氨基酸、表位的酶修饰或末端修饰。进行这种修饰的合适方法是本领域公知的。

[0075] 在一些实施方案中,修饰的ClyA纳米孔亚单位多肽可以用分子适配器进行化学修饰,所述分子适配器促进包含修饰的ClyA纳米孔亚单位多肽的纳米孔与靶核苷酸或靶多核苷酸序列之间的相互作用。适配器的存在改善了纳米孔和核苷酸或多核苷酸序列的主客体化学性质,从而改善了由修饰的ClyA纳米孔亚单位多肽形成的孔的测序能力。主客体化学原理是本领域公知的。适配器对纳米孔的物理或化学性质具有影响,改善了其与核苷酸或多核苷酸序列的相互作用。适配器可以改变孔的桶(barrel)或通道的电荷或特异性地与核苷酸或多核苷酸序列相互作用或结合,从而促进其与孔的相互作用。

[0076] 在一些实施方案中,分子适配器可以是环状分子,环糊精,能够杂交的物种,DNA结合剂或交联剂,肽或肽类似物,合成聚合物,芳族平面分子,带小正电荷分子或能够氢键结合的小分子。

[0077] 在一些实施方案中,分子适配器可共价连接至修饰的ClyA纳米孔亚单位多肽。可以使用本领域已知的任何方法将适配器共价连接到所述纳米孔。适配器通常通过化学键来连接。如果分子适配器通过半胱氨酸键来连接,则可以通过取代将一个或多个半胱氨酸引入到修饰的ClyA纳米孔亚单位多肽中。

[0078] 在其他实施方案中,修饰的ClyA纳米孔亚单位多肽可以连接或偶联到多核苷酸结

合蛋白,例如解旋酶,外切核酸酶和聚合酶。在一些实施方案中,修饰的ClyA纳米孔亚单位多肽可以连接或偶联至解旋酶,例如DNA解旋酶。适用于纳米孔测序的解旋酶,外切核酸酶和聚合酶的实例是本领域已知的。在一些实施方案中,修饰的ClyA纳米孔亚单位多肽可以连接或偶联至解旋酶,例如DNA解旋酶,Hel308解旋酶(例如,如W02013/057495中所述),RecD解旋酶(例如,如W02013/098562中所述),XPD解旋酶(例如,如W0201/098561中所述)或Dda解旋酶(例如,如W02015/055981中所述)。这形成了可用于表征靶多核苷酸的方法中的模块化测序系统。以下讨论多核苷酸结合蛋白。移位速度控制可以通过添加到系统中的多核苷酸结合蛋白的类型和/或燃料(ATP)的量来确定。例如,双链DNA分析物的移位速率可以通过双链DNA移位酶例如FtsK来控制。根据添加到系统中的燃料(ATP),靶多核苷酸的移位速度可以在约30B/s和1000B/s之间。

[0079] 在一些实施方案中,多核苷酸结合蛋白可以共价连接至修饰的ClyA纳米孔亚单位多肽。可以使用本领域已知的任何方法将多核苷酸结合蛋白共价连接至修饰的ClyA纳米孔亚单位多肽。修饰的ClyA纳米孔亚单位多肽和多核苷酸结合蛋白可以化学融合或基因融合。修饰的ClyA纳米孔亚单位多肽和多核苷酸结合蛋白质是遗传融合,如果整个构建体从单个多核苷酸序列表达。国际申请号PCT/GB09/001679(公开为W02010/004265)讨论了修饰的ClyA纳米孔亚单位多肽与多核苷酸结合蛋白的遗传融合。

[0080] 修饰的ClyA纳米孔亚单位多肽可以用分子适配器和多核苷酸结合蛋白进行化学修饰。

[0081] 本文所述的任何蛋白质,例如本文描述的修饰的ClyA纳米孔亚单位多肽和纳米孔,可以被修饰以辅助它们的鉴定或纯化,所述修饰例如通过添加组氨酸残基(his标签),天冬氨酸残基(asp标签),链霉素亲和素标签,Flag标签,SUMO标签,GST标签或MBP标签,或通过添加信号序列以促进它们从细胞中分泌,该细胞中的多肽不天然地含有该信号序列。引入遗传标签的另一方式是将标签化学反应到蛋白质上的天然或人工位点。其例子是将凝胶迁移试剂与设计在蛋白质外部的半胱氨酸反应。这已经被证明为分离溶血素异源寡聚体的方法(Chem Biol.1997Jul;4(7):497-505)。

[0082] 本文所述的任何蛋白质,例如本文所述的修饰的ClyA纳米孔亚单位多肽和纳米孔,可以用可检测标记物标记。可检测标记物可以是使得蛋白质被检测到的任何合适的标记物。合适的标记物包括但不限于荧光分子,放射性同位素,例如<sup>125</sup>I,<sup>35</sup>S,酶,抗体,抗原,多核苷酸和配体如生物素。

[0083] 本文所述的任何蛋白质,包括本文所述的修饰的ClyA纳米孔亚单位多肽,可以使用本领域已知的标准方法制备。编码蛋白质的多核苷酸序列可以使用本领域的标准方法衍生和复制。编码蛋白质的多核苷酸序列可以使用本领域中的标准技术在细菌宿主细胞中表达。蛋白质可以通过来自重组表达载体的多肽的原位表达而在细胞中制备。表达载体可选地携带诱导型启动子以控制多肽的表达。这些方法在Sambrook,J.and Russell,D.(2001).Molecular Cloning:A Laboratory Manual,3rd Edition.Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY中描述。

[0084] 由蛋白质制备生物体通过任何蛋白质液相色谱系统进行纯化后,或在重组表达后,蛋白质可以大规模制备。通常的蛋白质液相色谱系统包括FPLC、AKTA系统、Bio-Cad系统、Bio-Rad BioLogic系统和Gilson HPLC系统。

[0085] 编码修饰的ClyA纳米孔亚单位多肽的多核苷酸

[0086] 本文还提供编码如本文所述的修饰的ClyA纳米孔亚单位多肽中的任一种的多核苷酸序列。

[0087] 多核苷酸序列可以使用本领域的标准方法衍生和复制。编码野生型ClyA的染色体的DNA可以从孔制备生物体如伤寒沙门菌 (*Salmonella typhi*) 提取。编码孔亚基的基因可以使用包含特异性引物的PCR扩增。扩增的序列然后可以进行定点诱变。定点诱变的合适方法是本领域已知的并且包括例如组合链式反应。编码任何一种修饰的ClyA纳米孔亚单位多肽的多核苷酸可以使用公知技术制备, 例如Sambrook, J. and Russell, D. (2001) .Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY中描述的那些技术。

[0088] 然后可将得到的多核苷酸序列引入到重组可复制载体如克隆载体中。载体可用于在相容性宿主细胞中复制所述多核苷酸。因此, 可以通过将多核苷酸引入到可复制载体中, 将所述载体引入相容性宿主细胞中, 并在导致载体复制的条件下培养宿主细胞来制备多所述核苷酸序列。载体可以从宿主细胞中回收。用于克隆多核苷酸的合适宿主细胞是本领域已知的。

[0089] 本发明的另一方面包括制备本文所述的修饰的ClyA纳米孔亚单位多肽或构建体的方法。该方法包括在合适的宿主细胞中表达编码修饰的ClyA纳米孔亚单位多肽的任何实施方案的多核苷酸。所述多核苷酸优选是载体的一部分并且优选可操作地连接至启动子。

[0090] 修饰的ClyA纳米孔

[0091] 本发明的一个方面的特征在于修饰的ClyA纳米孔, 例如其允许捕获带负电荷的聚合物 (例如, 多核苷酸如DNA或RNA) 进入修饰的ClyA纳米孔和/或移位所述带负电荷的聚合物穿过修饰的ClyA纳米孔。修饰的ClyA纳米孔包含第一开口, 中间部分, 第二开口和从第一开口穿过中间部分延伸至第二开口的腔, 其中第一开口的腔表面包含第一带正电荷的氨基酸取代, 并且中间部分的腔表面包含第二带正电荷的氨基酸取代。第二开口的腔表面限定带负电荷的收缩部。第一带正电荷的氨基酸取代和第二带电荷的氨基酸取代在上文“修饰的ClyA纳米孔亚单位多肽”部分中详细描述。

[0092] 仅出于说明的目的, 图1 (版A) 显示根据本文描述的一个实施方案的修饰的ClyA纳米孔。修饰的ClyA纳米孔包括第一开口102, 中间部分104和第二开口106。腔108从第一开口102穿过中间部分104延伸到达第二开口106, 并且具有大约13nm至约15nm的总长度。第一开口102和中间部分104具有约5nm至约7nm的直径。第二开口106的腔表面限定带负电荷收缩部112, 其中最窄的横截面具有约3nm至约4nm的直径。修饰的ClyA纳米孔的第二开口106 (具有约3nm至约5nm的长度) 插入膜 (例如, 双分子层) 110中, 使得其中存在有修饰的ClyA纳米孔的溶液被分成两侧, 并且第一开口102存在于所述溶液的一侧中, 而带负电荷收缩部112存在于所述溶液的另一侧中。当在与第一开口102相同的一侧添加目标聚合物 (例如靶多核苷酸) 时, 第一开口102是顺式开口, 第二开口106是反式开口。

[0093] 如本文所用, 术语“腔表面”是指暴露于溶液的腔的内表面。

[0094] 如本文可互换使用的, 术语“带负电荷收缩部”或“带负电荷的收缩部”是指具有净负表面电荷的收缩部。例如, 限定带负电荷收缩部的第二开口的腔表面具有如图1 (版A) 所示的净负表面电荷。

[0095] 在本文所述的任何修饰的Cl<sub>y</sub>A纳米孔中,从第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)到第二带正电荷修饰(例如,第二带正电荷的氨基酸取代)的腔内距离可以在约0.5nm至约10nm,或约3nm至约7nm的范围内变化。在一些实施方案中,从第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)到第二带正电荷修饰(例如,第二带正电荷的氨基酸取代)的腔内距离可以是约1nm,约2nm,约3nm,约4nm,约5nm,约6nm,约7nm,约8nm或约9nm。

[0096] 可以使用任何形式的Cl<sub>y</sub>A来制备本文所述的修饰的Cl<sub>y</sub>A纳米孔。例如,如上所述,各种形式的Cl<sub>y</sub>A的氨基酸序列,包括例如但不限于野生型Cl<sub>y</sub>A(Cl<sub>y</sub>A-WT)和Cl<sub>y</sub>A-AS并且编码它们的核苷酸序列在本领域中是已知的。因此,在一些实施方案中,修饰的Cl<sub>y</sub>A纳米孔可以包含亚单位多肽,其具有与本文所述的参考Cl<sub>y</sub>A氨基酸序列至少约80%(包括例如至少约85%,至少约90%,至少约95%或更高)同一性的氨基酸序列。在一些实施方案中,修饰的Cl<sub>y</sub>A纳米孔可以包含亚单位多肽,其具有与对应于野生型Cl<sub>y</sub>A的SEQ ID NO:1所示的氨基酸序列至少约80%(包括例如至少约85%,至少约90%,至少约95%或更高)同一性的氨基酸序列。或者,修饰的Cl<sub>y</sub>A纳米孔可以包含亚单位多肽,其具有与对应于Cl<sub>y</sub>A-AS的氨基酸序列的SEQ ID NO:2所示的氨基酸序列至少约80%(包括例如至少约85%,至少约90%,至少约95%或更高)同一性的氨基酸序列。在一些实施方案中,与包含第一和第二带正电荷修饰(例如,第一和第二带正电荷的氨基酸取代)的SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列相比,所述修饰的Cl<sub>y</sub>A纳米孔可以包含多达15个取代(例如2,3,4,5,6,7,8,9,10,11,12,13,14或15个取代)。

[0097] 在本文所述的任何修饰的Cl<sub>y</sub>A纳米孔中,第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)可以位于第一开口内以允许捕获溶液中暴露于第一开口的带负电荷的聚合物(例如但不限于脱氧核糖核酸(DNA)如双链DNA或单链DNA)。例如,第一带正电荷修饰(例如,第一带正电荷的氨基酸取代)可位于SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列中的位置E106,S110,D114,D121,D122,E129,E85,E78,D268,D267,D265,E258或其组合。

[0098] 在本文所述的任何修饰的Cl<sub>y</sub>A纳米孔中,第二带正电荷修饰(例如,第二带正电荷的氨基酸取代)可位于中间部分内以便允许带负电荷的聚合物(例如,但不能限于脱氧核糖核酸(DNA),如双链DNA或单链DNA)移位穿过孔的内腔。例如,第二带正电荷修饰(例如,第二带正电荷的氨基酸取代)可位于SEQ ID NO:1或SEQ ID NO:2所示的氨基酸序列中的位置D74,D71,D64,E53,E161,D158,E46,E42,D41或其组合。

[0099] 修饰的Cl<sub>y</sub>A纳米孔可以是同源多聚体(例如纳米孔内的所有亚基是相同的)或异源多聚体(例如至少一个亚基与纳米孔内的其他亚基不同)。修饰的Cl<sub>y</sub>A纳米孔可以包含任何数量的亚单位多肽,其足以形成足够大的腔以允许靶聚合物(例如多核苷酸)通过。在一些实施方案中,修饰的Cl<sub>y</sub>A纳米孔可以包含12个亚单位多肽或更多,包括例如13个亚单位多肽,和14个亚单位多肽,其中至少一个或多个亚单位多肽包含如本文所述的第一个和第二带正电荷的氨基酸取代。

[0100] 例如基于纳米孔中的停留时间和流过孔的电流,修饰的Cl<sub>y</sub>A纳米孔可用于区分双链多核苷酸与单链多核苷酸。另外,修饰的Cl<sub>y</sub>A纳米孔可以用于表征诸如测序多核苷酸序列。修饰的Cl<sub>y</sub>A纳米孔也可用于区分修饰的碱基,例如区分甲基化和未甲基化的核苷酸。

[0101] 与不具有本文所述的第一和第二带正电荷的取代的Cl<sub>y</sub>A纳米孔相比,本文描述的

修饰的ClyA纳米孔在低离子强度溶液中提供更高频率的捕获和/或多核苷酸穿过纳米孔的移位。

[0102] 如本文所用,术语“低离子强度溶液”是指具有小于2M的离子强度的溶液,包括例如小于1M,小于900mM,小于800mM,小于700mM,小于600mM,小于500mM,小于400mM,小于300mM,小于200mM,小于150mM或更小。在一些实施方案中,较低的离子强度溶液具有离子强度为至少约50mM,至少约100mM,至少约150mM,至少约200mM,至少约300mM,至少约400mM,至少约500mM,至少约600mM,至少约700mM,至少约800mM,至少约900mM,至少约1M或更高。上述参考范围的组合也包含在内。例如,低离子强度溶液可具有约100mM至约600mM,或约150mM至约300mM的离子强度。可以使用任何盐来产生具有适当离子强度的溶液。在一些实施例中,碱性盐(例如但不限于氯化钾或氯化钠)可用于低离子强度溶液中。

[0103] 修饰的ClyA纳米孔可以在一系列条件下区分不同的核苷酸。特别地,孔可以在有利于核酸表征,如测序的条件下区分核苷酸。修饰的ClyA纳米孔可以区分不同核苷酸的程度可以通过改变施加的电势,盐浓度,缓冲液,温度和添加剂如尿素,甜菜碱和DTT的存在来控制。这允许微调孔的功能,特别是在测序时。这在下文更详细地讨论。修饰的ClyA纳米孔也可以通过与一种或多种单体的相互作用而不是以核苷酸为基础的核苷酸来鉴定多核苷酸聚合物。

[0104] 在一些实施方案中,本文提供的修饰的ClyA纳米孔可用于表征核酸-蛋白质相互作用。在一些实施方案中,使用不同的感测模式,例如通过扫描和绘制沿核酸的结合位点的位置和/或通过探测蛋白质与核酸之间的相互作用的强度,纳米孔可以被用来鉴定蛋白质-核酸。在一些实施方案中,可利用核酸的天然电荷以将电泳力施加至核酸-蛋白质复合物。例如,在一些实施方案中,可以使用电压驱动的单个DNA分子穿过蛋白质纳米孔来评估DNA-蛋白质相互作用。在这样的实施方案中,施加于单个DNA蛋白质复合物(例如,DNA-外切核酸酶I复合物,DNA-解旋酶复合物,DNA-钳复合物)的电泳力可以拉开这两个分子,而同时离子电流变化可用来评估所述复合物的解离速率。在一些实施方案中,本文提供的修饰的ClyA纳米孔可用于检测和表征涉及核酸和其他核酸结合蛋白例如转录因子,酶,DNA包装蛋白等的核酸-蛋白质相互作用。

[0105] 修饰的ClyA纳米孔可以被分离,被基本上分离,纯化或基本上纯化。如果修饰的ClyA纳米孔完全不含任何其他组分,如脂质或其他孔,则修饰的ClyA纳米孔是被分离或纯化的。如果孔与不干扰其预期用途的载体或稀释剂混合,则所述孔基本上是被分离的。例如,如果孔以包含小于10%,小于5%,小于2%或小于1%的其他组分例如三嵌段共聚物,脂质或其他孔的形式存在,则所述孔基本上是分离的或基本上纯化的。或者,一个或多个修饰的ClyA纳米孔可以存在于膜中。下面讨论合适的膜。

[0106] 修饰的ClyA纳米孔可以作为单个或单独的孔而存在。或者,修饰的ClyA纳米孔可以存在于两个或更多个孔的同源或异源群体中。在一些实施方案中,修饰的ClyA纳米孔可以排列成微孔阵列,其中每个微孔含有至少一个纳米孔在膜中。

[0107] 同源多聚体ClyA纳米孔

[0108] 本文还提供了包含相同的修饰的ClyA纳米孔亚单位多肽的同源多聚体纳米孔。同源多聚体纳米孔可以包含本文描述的修饰的ClyA纳米孔亚单位多肽的任何实施方案。同源多聚体纳米孔可用于表征诸如测序多核苷酸和/或检测单链多核苷酸vs双链多核苷酸的存

在或不存在。本文所述的同源多聚体纳米孔可具有上述任何优点。

[0109] 同源多聚体孔可以含有任何数量的修饰的ClyA纳米孔亚单位多肽。该孔通常包含至少10个,至少11个,至少12个,至少13个或至少14个相同的修饰的ClyA纳米孔亚单位多肽,诸如12,13或14个相同的修饰的ClyA纳米孔亚单位多肽。

[0110] 以下更详细地讨论制备孔的方法。

[0111] 异源多聚体ClyA纳米孔

[0112] 本文还提供了包含至少一种修饰的ClyA纳米孔亚单位多肽的异源多聚体纳米孔。异源多聚体纳米孔可以用于表征诸如测序多核苷酸,和/或检测单链多核苷酸vs双链多核苷酸的存在或不存在。异源多聚体纳米孔可以使用本领域已知的方法制备(例如Protein Sci.2002Jul;11(7):1813-24)。

[0113] 异源多聚体孔含有足够的亚单位多肽以形成孔。亚单位多肽可以是任何类型的。该孔通常包含至少10个,至少11个,至少12个,至少13个或至少14个亚单位多肽,例如12,13或14个亚单位多肽。

[0114] 在一些实施方案中,所有亚单位多肽(诸如12,13或14个亚单位多肽)是修饰的ClyA纳米孔亚单位多肽,并且它们中的至少一个不同于其他亚单位多肽。在一些实施方案中,所述孔包含12或13个修饰的ClyA纳米孔亚单位多肽,并且它们中的至少一个不同于其他。它们可能都彼此不同。

[0115] 在一些实施方案中,至少一个亚单位多肽不是本文所述的修饰的ClyA纳米孔亚单位多肽。在该实施方案中,其余单体可以是本文所述的修饰的ClyA纳米孔亚单位多肽中的任一个。因此,所述孔可以包含12,11,10,9,8,7,6,5,4,3,2或1个修饰的ClyA纳米孔亚单位多肽。形成纳米孔的修饰的ClyA纳米孔亚单位多肽可以是相同的或不同的。

[0116] 以下更详细地讨论制备孔的方法。

[0117] 多核苷酸表征

[0118] 本发明的另一方面提供了表征靶多核苷酸的方法。该方法包括:(a)在约50mM至约1M的低离子强度溶液中提供根据本文所述的任何实施方案的修饰的ClyA纳米孔和膜,其中所述修饰的ClyA纳米孔存在于所述膜中;(b)在步骤(a)的低离子强度溶液中加入靶多核苷酸;和(c)在跨所述纳米孔施加电势期间,测量穿过所述修饰的ClyA纳米孔的离子流,其中离子流测量值代表所述靶多核苷酸的一个或多个特征。在一些实施方案中,将所述靶多核苷酸添加至所述低离子强度溶液的顺式侧。

[0119] 在一些实施方案中,低离子强度溶液可具有约50mM至约300mM,或约150mM至约300mM的离子强度。

[0120] 所述靶多核苷酸也可以被称为模板多核苷酸或感兴趣的多核苷酸。

[0121] 多核苷酸

[0122] 多核苷酸如核酸是含有两个或更多个核苷酸的大分子。多核苷酸或核酸可以包括任何核苷酸的任意组合。核苷酸可以是天然存在的或人造的。多核苷酸中的一个或多个核苷酸可被氧化或甲基化。多核苷酸中的一个或多个核苷酸可被损伤。例如,多核苷酸可以包含嘧啶二聚体。这种二聚体通常与紫外线损伤有关,并且是皮肤黑素瘤的首要原因。多核苷酸中的一个或多个核苷酸可以被修饰,例如用标记物或标签进行修饰。下面描述了合适的标记物。多核苷酸可以包含一个或多个间隔区。

- [0123] 核苷酸通常含有核碱基,糖和至少一个磷酸基团。核碱基和糖形成核苷。
- [0124] 核碱基通常是杂环的。核碱基包括但不限于嘌呤和嘧啶,更具体地腺嘌呤(A),鸟嘌呤(G),胸腺嘧啶(T),尿嘧啶(U)和胞嘧啶(C)。
- [0125] 糖通常是戊糖。核苷酸糖包括但不限于核糖和脱氧核糖。糖优选是脱氧核糖。
- [0126] 多核苷酸优选包含以下核苷:脱氧腺苷(dA),脱氧尿苷(dU)和/或胸苷(dT),脱氧鸟苷(dG)和脱氧胞苷(dC)。
- [0127] 核苷酸通常是核糖核苷酸或脱氧核糖核苷酸。核苷酸通常含有单磷酸,二磷酸或三磷酸。核苷酸可以包含多于三种的磷酸盐,例如4或5种磷酸盐。磷酸酯可以连接在核苷酸的5'或3'侧上。核苷酸包括,但不限于,单磷酸腺苷(AMP),单磷酸鸟苷(GMP),单磷酸胸苷(TMP),单磷酸尿苷(UMP),5-甲基单磷酸胞苷,5-羟甲基单磷酸胞苷,单磷酸胞苷(CMP),环状单磷酸腺苷(cAMP),环状单磷酸鸟苷酸(cGMP),脱氧单磷酸腺苷(dAMP),脱氧单磷酸鸟苷(dGMP),脱氧单磷酸胸苷(dTMP),脱氧单磷酸尿苷(dUMP),脱氧单磷酸胞苷(dCMP)和脱氧甲基单磷酸胞苷。核苷酸优选选自AMP,TMP,GMP,CMP,UMP,dAMP,dTMP,dGMP,dCMP和dUMP。
- [0128] 核苷酸可以是脱碱基(即缺乏核碱基)。核苷酸也可以缺乏核碱基和糖。
- [0129] 多核苷酸中的核苷酸可以以任何方式彼此连接。如在核酸中一样,核苷酸通常通过它们的糖和磷酸基团连接。如嘧啶二聚体中一样,核苷酸可以通过它们的核碱基连接。
- [0130] 多核苷酸可以是单链或双链的。至少一部分多核苷酸优选为双链。
- [0131] 多核苷酸可以是核酸,如脱氧核糖核酸(DNA)或核糖核酸(RNA)。多核苷酸可包含杂交到DNA的一条链的RNA的一个链。多核苷酸可以是本领域已知的任何合成核酸,诸如肽核酸(PNA),甘油核酸(GNA),苏糖核酸(TNA),锁核酸(LNA)或其它具有核苷酸侧链的合成聚合物。PNA骨架由通过肽键连接的重复的N-(2-氨基乙基)-甘氨酸单元组成。GNA骨架由通过磷酸二酯键连接在一起的重复的乙二醇单元组成。TNA骨架由通过磷酸二酯键连接在一起的重复苏糖组成。LNA由如上所述的核糖核苷酸组成,所述核糖核苷酸具有连接核糖部分中的2'氧和4'碳的额外桥。
- [0132] 多核苷酸最优选为核糖核酸(RNA)或脱氧核糖核酸(DNA)。
- [0133] 多核苷酸可以是任何长度。例如,多核苷酸的长度可以是至少10,至少50,至少100,至少150,至少200,至少250,至少300,至少400或至少500个核苷酸或核苷酸对的长度。多核苷酸可以是1000或更多个核苷酸或核苷酸对的长度,5000或更多个核苷酸或核苷酸对的长度,或100000或更多个核苷酸或核苷酸对的长度。
- [0134] 可以研究任何数量的多核苷酸。例如,本文所述的方法可能涉及表征2,3,4,5,6,7,8,9,10,20,30,50,100或更多个多核苷酸。如果表征两个或更多个多核苷酸,它们可以是不同的多核苷酸或相同多核苷酸的两种情况(two instances)。
- [0135] 多核苷酸可以是天然存在的或人造的。例如,该方法可以用于验证制备的寡核苷酸的序列。该方法通常在体外进行。
- [0136] 多核苷酸可以包含连接的物质,例如蛋白质或分析物。多核苷酸可以包含杂交的探针。
- [0137] 样本
- [0138] 每种分析物通常存在于任何合适的样品中。该方法可以在已知含有或怀疑含有分析物的两种或多种样品上进行。或者,可以在两个或多个样品上实施所述方法,以确认已知

的或预期在样品中的存在的两种或多种分析物的同一性。在一些实施方案中,可以在样品上实施该方法,以区分双链多核苷酸与单链多核苷酸。

[0139] 第一样品和/或第二样品可以是生物样品。本文所述的方法可以使用从任何生物或微生物获得或提取的至少一种样品在体外实施。第一样品和/或第二样品可以是非生物样品。非生物样品可以是流体样品。非生物样品的实例包括外科手术液,水如饮用水、海水或河水,以及用于实验室检测的试剂。

[0140] 第一样品和/或第二样品通常是在本文所述的方法中使用前被处理过,例如通过离心或通过膜滤掉不需要的分子或细胞,例如红血细胞。第一个样品和/或第二个样品可以在采集后立即测定。第一样品和/或第二样品通常也可以在测定之前被储存,优选低于-70℃储存。

[0141] 特征

[0142] 本方法可以包括测量两个,三个,四个或五个或更多个多核苷酸的特征。所述一个或多个特征优选地选自(i)多核苷酸的长度,(ii)多核苷酸的同一性,(iii)多核苷酸的序列,(iv)多核苷酸的二级结构和(v)多核苷酸是否被修饰。(i)至(v)的任意组合可以根据本文所述的方法测量,如{i},{ii},{iii},{iv},{v},{i,ii},{i,iii},{i,iv},{i,v},{ii,iii},{ii,iv},{ii,v},{iii,iv},{iii,v},{i,ii,iv},{i,ii,v},{i,iii,iv},{i,iii,v},{i,iv,v},{i,ii,iii,iv},{ii,iii,v},{ii,iv,v},{iii,iv,v},{v},{i,ii,iv,v},{i,iii,iv,v},{ii,iii,iv,v}或{i,ii,iii,iv,v}。可以测量第一多核苷酸相比于第二多核苷酸的(i)至(v)的不同组合,包括上述列出的任何组合。

[0143] 对于(i),多核苷酸的长度例如可以通过确定所述多核苷酸和所述孔之间的相互作用次数,或所述多核苷酸与所述孔之间相互作用的持续时间来测定。

[0144] 对于(ii),所述多核苷酸的同一性可以通过多种方式测定。所述多核苷酸的同一性可以联合所述多核苷酸的序列的测定来测定,或不联合多核苷酸的序列的测定进行测定。前者是直接进行的;对所述多核苷酸进行测序,并由此进行鉴定。后者可以以几种方式来完成。例如,可以测定多核苷酸中特定模序的存在(而无需测定该多核苷酸的其余序列)。或者,所述方法中测定的特定的电和/或光信号可鉴定来自特定来源的多核苷酸。

[0145] 对于(iii),多核苷酸的序列可以如以前所述确定。合适的测序方法,特别是那些使用电测量的方法,描述于Stoddart D et al.,Proc Natl Acad Sci,12;106(19):7702-7,Lieberman KR et al,J Am Chem Soc.2010;132(50):17961-72,和国际申请WO 2000/28312中。

[0146] 对于(iv),所述二级结构可以以多种方式测量。例如,如果该方法包含电测量,二级结构可以利用穿过孔时的停留时间的改变或电流变化进行测量。这使得单链和双链多核苷酸的区域能够得到区分。

[0147] 对于(v),可以测定任何修饰的存在或不存在。该方法优选包括确定所述多核苷酸是否通过甲基化、氧化、损伤、用一个或多个蛋白质或一个或多个标记物、标签或间隔器进行了修饰。特异性修饰将导致与孔的特异性相互作用,这可以使用下面描述的方法来测定。例如,甲基胞嘧啶可以基于其与每个核苷酸的相互作用过程中穿过孔时的电流,与胞嘧啶区别开。

[0148] 靶多核苷酸与本文所述的任何一种修饰的ClyA纳米孔接触。孔通常存在于膜中。

合适的膜在下面讨论。所述方法可以使用任何适于研究膜/孔系统(其中孔存在于膜中)的设备进行实施,该方法可以使用任何适合于感测跨膜孔的设备实施。例如,所述设备包括一个腔,所述腔包括水溶液和将该腔分割为两部分的屏障(barrier)。所述屏障通常具有开口,其中在开口中形成包括所述孔的膜。或者该屏障形成膜,所述膜中存在孔。

[0149] 该方法可以使用国际申请No.PCT/GB08/000562(WO 2008/102120)中描述的设备进行。

[0150] 多种不同类型的测量值可以得到。这包括但不限于:电测量值和光学测量值。可能的电测量包括:电流测量,阻抗测量,风洞测量(Ivanov AP et al., Nano Lett. 2011 Jan 12; 11(1): 279-85)和FET测量(国际申请WO 2005/124888)。光学测量可以与电测量组合(Soni GV et al., Rev Sci Instrum. 2010 Jan; 81(1): 014301)。所述测量值可以是跨膜电流测量值,如流过孔的离子电流的测量值。或者,测量值可以是代表穿过通道的离子流动的荧光测量值,如Heron et al., J. Am. Chem. Soc., 2009, 131(5), 1652-1653所公开的,或使用FET跨膜电压的测量值。

[0151] 电测量值可以使用在Stoddart D et al., Proc Natl Acad Sci, 12; 106(19): 7702-7, Lieberman KR et al., J Am Chem Soc. 2010; 132(50): 17961-72, 和国际申请WO 2000/28312中所描述的标准单通道记录设备获得。或者,电测量可以使用多通道系统进行,例如国际申请WO 2009/077734和国际申请号WO 2011/067559中描述的。

[0152] 所述方法可以跨膜施加电势实施。所施加的电势可以是电压电势。或者,所施加的电势可以是化学势。这方面的实例是跨膜,如两亲层,使用盐梯度。盐梯度在Holden et al., J Am Chem Soc. 2007 Jul 11; 129(27): 8650-5中公开。在一些情况下,随着多核苷酸相对于孔移动时,流经孔的电流被用来评估或确定多核苷酸的序列。这可以描述为链测序。

[0153] 所述方法可以包括随着多核苷酸相对孔移动,测量经过所述孔的电流。因此,所述方法中使用的设备还可以包括能够施加电势并测量穿过膜和孔的电信号的电路。所述方法可以使用膜片钳或电压钳实施。所述方法优选涉及使用电压钳。

[0154] 所述方法可以包括随着所述多核苷酸相对于所述孔移动,测量经过所述孔的电流。用于测量穿过跨膜蛋白孔的离子电流的合适的条件是本领域已知的,并且在实施例中公开。所述方法通常用跨所述膜和孔施加的电压进行实施。使用的电压通常为+5V至-5V,例如+4V至-4V, +3V至-3V或+2V至-2V。使用的电压通常为-600mV至+600mV或-400mV至+400mV。所使用的电压优选在具有以下下限和上限的范围内:所述下限选自-400mV, -300mV, -200mV, -150mV, -100mV, -50mV, -20mV和0mV的范围内,所述上限独立地选自+10mV, +20mV, +50mV, +100mV, +150mV, +200mV, +300mV, 和+400mV。使用的电压更优选在100mV至240mV的范围内,并且最优选在120mV至220mV的范围内。可通过施加提高的电势通过孔来提高对不同核苷酸的鉴别力。

[0155] 所述方法通常在任何载荷子,如金属盐,如碱金属盐,卤盐,如盐酸盐,如碱金属盐酸盐的存在下实施。载荷子可包括离子液体或有机盐,例如四甲基氯化铵,三甲基苯基氯化铵,苯基三甲基氯化铵,或1-乙基-3-甲基咪唑鎓氯化物。在上面论述的示例性设备中,所述盐存在于所述腔中的水溶液中。通常使用氯化钾(KCl),氯化钠(NaCl)、氯化铯(CsCl)或亚铁氰化钾和铁氰化钾的混合物。优选KCl、NaCl,和亚铁氰化钾与铁氰化钾的混合物。所述载荷子可以不对称的跨所述膜。例如在膜的每侧,所述载荷子的种类和/或浓度可以是不同

的。

[0156] 盐浓度可能处于饱和状态。盐浓度可以是3M或更低,通常为0.1至2.5M,0.3至1.9M,0.5至1.8M,0.7至1.7M,0.9至1.6M或1M至1.4M。盐浓度优选为150mM至1M。本方法优选使用至少0.3M,例如至少0.4M,至少0.5M,至少0.6M,至少0.8M,至少1.0M,至少1.5M,至少2.0M,至少2.5M或至少3.0M的盐浓度进行实施。高盐浓度提供高信噪比,并使得在正常电流波动的背景下,代表核苷酸存在的电流能被识别。虽然本文描述的修饰的ClyA纳米孔可用于表征高盐溶液中的多核苷酸,但所述修饰的ClyA纳米孔可允许甚至在如上所述的低离子强度溶液中有效捕获和/或移位多核苷酸(例如双链DNA或单链DNA)穿过纳米孔。

[0157] 本方法通常在缓冲剂存在下实施。在上面讨论的示例性设备中,缓冲剂存在于所述腔的水溶液中。本文所述的方法可以使用任何缓冲剂。通常,缓冲剂是磷酸盐缓冲剂。其他合适的缓冲液是HEPES和Tris-HCl缓冲液。本方法通常在pH为4.0至12.0,4.5至10.0,5.0至9.0,5.5至8.8,6.0至8.7或7.0至8.8或7.5至8.5的下进行。所使用的pH优选为约7.5或8.0。

[0158] 本方法可以在0℃至100℃,15℃至95℃,16℃至90℃,17℃至85℃,18℃至80℃,19℃至70℃或从20℃到60℃实施。本方法通常在室温下进行。本方法可选地在支持酶功能的温度下进行,如约37℃实施。

[0159] 多核苷酸结合蛋白

[0160] 在一些实施方案中,用于表征靶多核苷酸的方法可以包括在低离子强度溶液中添加多核苷酸结合蛋白,使得多核苷酸结合蛋白与靶多核苷酸结合并且控制靶多核苷酸穿过修饰的ClyA纳米孔的运动。

[0161] 所述多核苷酸结合蛋白可以是能够与多核苷酸结合并控制其穿过孔的运动的任何蛋白质。多核苷酸结合蛋白的例子包括但不限于解旋酶,聚合酶,外切核酸酶,DNA夹子等。可以以任何顺序使多核苷酸与多核苷酸结合蛋白和孔接触。优选地,当多核苷酸与多核苷酸结合蛋白(例如解旋酶)和孔接触时,所述多核苷酸首先与所述蛋白质形成复合物。当跨孔施加电压时,多核苷酸/蛋白质复合物然后与孔形成复合物并控制多核苷酸穿过孔的运动。

[0162] 使用多核苷酸结合蛋白的所述方法中的任何步骤通常在游离核苷酸或游离核苷酸类似物和促进多核苷酸结合蛋白的作用的酶辅因子的存在下进行。

[0163] 解旋酶(s)和分子制动器(s)

[0164] 在一个实施例中,本方法包括:

[0165] (a) 给多核苷酸提供一个或多个解旋酶和一个或多个与所述多核苷酸连接的分子制动器;

[0166] (b) 在包括存在于膜中的修饰的ClyA纳米孔的低离子强度溶液中添加所述多核苷酸,并跨所述孔施加电势使得所述一个或多个解旋酶与所述一个或多个分子制动器结合在一起并共同控制所述多核苷酸穿过孔的运动;

[0167] (c) 随着多核苷酸相对于孔移动,在跨纳米孔施加电势过程中,测量穿过所述修饰的ClyA纳米孔的离子流,其中离子流测量值代表所述多核苷酸的一个或多个特征,并由此表征所述多核苷酸。这种类型的方法在国际申请No.PCT/GB2014/052737中详细讨论。

[0168] 膜

[0169] 本文所述的修饰的ClyA纳米孔可以存在于膜中。在表征多核苷酸的方法中,多核苷酸通常与膜中的修饰的ClyA纳米孔接触。可以使用任何膜。合适的膜在本领域中是公知的。该膜优选为两亲性层。两亲性层是由两性分子如磷脂形成的层,其具有亲水性和亲油性。两亲分子可以是合成的或天然存在的。非天然存在的两亲物质和形成单分子层的两亲物质是本领域已知的,包括例如嵌段共聚物(Gonzalez-Perez et al., Langmuir, 2009, 25, 10447-10450)。嵌段共聚物是其中两个或多个单体亚单元聚合在一起而形成单一聚合物链的聚合物材料。嵌段共聚物具有的性质通常由每个单体亚单元贡献。然而,嵌段共聚物可具有从各亚单元形成的聚合物不具备的独特性质。嵌段共聚物可以被改造使得单体亚单元之一是疏水性或亲油性的,而其他亚单元在水性介质中是亲水性的。在这种情况下,嵌段共聚物可具有两亲特性,并可形成模仿生物膜的结构。嵌段共聚物可以是二嵌段共聚物(由两个单体亚单元组成),但也可以由两个以上的单体亚单元构成,形成更复杂的结构的两亲物质。共聚物可以是三嵌段,四嵌段或五嵌段共聚物。该膜优选为三嵌段共聚物膜。

[0170] 古细菌双极四醚脂质是天然存在的脂质,其构建为使得脂质形成单层膜。这些脂质通常存在于在恶劣生物环境中存活的极端微生物,嗜热菌,嗜盐菌和嗜酸菌中。它们的稳定性被认为是来源于最终双分子层的融合性质。很容易通过生成具有亲水-疏水-亲水一般基序的三嵌段聚合物,来构建模拟这些生物实体的嵌段共聚物的材料。这种材料可以形成单层膜,其行为类似于脂质双分子层并且包括从囊泡到层状膜的一系列相行为。由这些三嵌段共聚物形成的膜相对于生物脂质膜具有几个优点。因为三嵌段共聚物是合成的,所以可以仔细控制精确的构造以提供正确的链长和性质以形成膜并与孔和其他蛋白质相互作用。

[0171] 嵌段共聚物也可以由不归类为脂质副材料(lipid sub-materials)的亚单元构成,例如疏水聚合物可以由硅氧烷或其他非烃基单体制成。嵌段共聚物的亲水性亚部分也可具有低蛋白质结合性质,这允许形成暴露于未加工生物样品时高度耐受的膜。该头基单位(head group unit)也可能来自非经典的脂质头基(head-groups)。

[0172] 与生物脂质膜相比,三嵌段共聚物膜也具有增加的机械和环境稳定性,例如更高得多的操作温度或pH范围。所述嵌段共聚物的合成性质为对于定制基于聚合物的膜的广泛应用提供了平台。

[0173] 膜最优选为国际申请号PCT/GB2013/052766或PCT/GB2013/052767中公开的膜之一。

[0174] 两亲分子可以被化学修饰或官能化以促进多核苷酸的结合。

[0175] 两亲层可以是单层或双分子层。两亲层通常是平面的。两亲层可以是弯曲的。两亲性层可以被支撑的。

[0176] 两亲性膜通常是自然流动的,基本上表现为二维流体,其脂质扩散速率约为 $10^{-8} \text{ cm s}^{-1}$ 。这意味着孔和结合的多核苷酸通常在两亲膜内移动。

[0177] 该膜可以是脂质双分子层。脂质双分子层为细胞膜的模型,可作为一系列实验研究的优秀平台。例如,脂质双分子层膜可用于通过单通道记录器进行膜蛋白的体外研究。或者,脂质双分子层可用作生物传感器来检测一系列物质的存在。脂质双分子层可以是任何脂质双分子层。合适的脂质双分子层包括但不限于,平面脂质双分子层,支撑双分子层或脂质体。脂质双分子层优选是平面脂质双分子层。合适的脂质双分子层在国际申请No. PCT/

GB08/000563 (公布号为W0 2008/102121), 国际申请No. PCT/GB08/004127 (公布号为W0 2009/077734) 和国际申请No. PCT/GB2006/001057 (公布号为W0 2006/100484) 中公开。

[0178] 在一些实施方案中, 可将多核苷酸与包含本文所述的任何一种修饰的ClyA纳米孔的膜连接。所述方法可以包括将多核苷酸与包含本文所述的任何一种修饰的ClyA纳米孔的膜连接。优选使用一个或多个锚将多核苷酸与膜连接。使用任何已知的方法将多核苷酸与膜连接。

[0179] 双链多核苷酸测序

[0180] 在一些实施方案中, 多核苷酸可以是双链的。如果多核苷酸是双链的, 则本方法可以进一步包括在接触步骤之前将发夹适配器连接到多核苷酸的一端。然后随着多核苷酸与如本文所述的修饰的ClyA纳米孔接触或相互作用, 或在所述多核苷酸与如本文所述的修饰的ClyA纳米孔接触或相互作用之前, 所述多核苷酸的两条链可以被分离。随着多核苷酸结合蛋白例如解旋酶或分子制动器控制下的多核苷酸穿过孔的运动, 所述两条链可以被分开。这在国际申请号PCT/GB2012/051786 (公布为W02013/014451) 中描述。以这种方式连接和询问 (interrogating) 双链构建体上的两条链提高了表征的效率和准确性。

[0181] 在拐角处测序

[0182] 在一个优选的实施方案中, 目标双链多核苷酸在一端提供有发夹环适配器, 并且本方法包括使所述多核苷酸与本文所述的任何一种修饰的ClyA纳米孔接触, 使得多核苷酸的两条链移动穿过孔, 并且随着所述多核苷酸的两条链相对于所述孔移动, 获取一个或多个测量值, 其中测量值代表所述多核苷酸的链的一个或多个特征, 从而表征所述目标双链多核苷酸。上面讨论的任何实施例同样适用于该实施例。

[0183] 前导序列

[0184] 在接触步骤之前, 本方法优选包括在多核苷酸连接前导序列, 所述前导序列优先旋入孔。所述前导序列有助于本发明的方法。引导序列被设计为优选地旋入本文所述的修饰的ClyA纳米孔中的任何一个中, 并因此有助于多核苷酸穿过纳米孔的移动。前导序列也可以用于将多核苷酸连接至如上所述的一个或多个锚。

[0185] 修饰的多核苷酸

[0186] 在表征之前, 在聚合酶使用靶多核苷酸作为模板形成修饰的多核苷酸的条件下, 通过使多核苷酸与聚合酶和游离核苷酸群接触来修饰靶多核苷酸, 其中当形成修饰的多核苷酸时, 所述聚合酶用不同核苷酸物质取代靶多核苷酸中的一个或多个核苷酸物质。然后对所述修饰的多核苷酸提供连接到所述多核苷酸的一个或多个解旋酶以及连接到所述多核苷酸的一个或多个分子制动器。国际申请No. PCT/GB2015/050483中描述了这种类型的修饰。可使用本文讨论的任何聚合酶。

[0187] 在聚合酶使用模板多核苷酸作为模板形成修饰的多核苷酸的条件下, 使模板多核苷酸与聚合酶接触。这样的条件在本领域中是已知的。例如, 多核苷酸通常在商业可获得的聚合酶缓冲液 (例如来自New England Biolabs®的缓冲液) 中与聚合酶接触。通常使用引物或3'发夹作为聚合酶延伸的成核点。

[0188] 使用跨膜孔的多核苷酸表征, 例如测序, 通常包括分析由k个核苷酸组成的聚合物单元, 其中k是正整数 (即“k-聚体 (k-mers)”)。这在国际申请No. PCT/GB2012/052343 (公开为W0 2013/041878) 中讨论。虽然期望对于不同的k-mer具有清楚区分的电流测量值, 但这

些测量值中的一些通常是重叠的。尤其是k-mer中具有大数量的聚合物单元,即高值的k,分辨由不同的k-mer产生的测量值变得困难,不利于所述多核苷酸信息的获得,例如评估所述多核苷酸的潜在序列(underlying sequence)。

[0189] 通过用修饰的多核苷酸中的不同核苷酸物质取代靶多核苷酸中的一个或多个核苷酸物质,所述修饰的多核苷酸含有不同于所述靶多核苷酸的k-mer。所述修饰的多核苷酸中的不同k-mer能够产生与靶多核苷酸中的k-mer的不同的电流测量值,因此修饰的多核苷酸提供了与所述靶多核苷酸的不同的信息。来自修饰多核苷酸的附加信息可以使得表征靶多核苷酸更容易。在一些情况下,修饰的多核苷酸本身可能更容易地表征。例如可设计修饰的多核苷酸包含在它们电流测量值之间具有增强的分离或清楚的分离的k-mer或具有减少噪音的k-mer。

[0190] 当形成所述修饰的多核苷酸时,聚合酶优选不同的核苷酸物质取代靶多核苷酸中的两种或多种核苷酸物质。聚合酶可以用不同的核苷酸物质取代靶多核苷酸中的两种或多种核苷酸物质中的每一种。聚合酶可以用相同的核苷酸物质取代靶多核苷酸中的两种或多种核苷酸物质中的每一种。

[0191] 如果靶多核苷酸是DNA,则修饰的多核苷酸中的不同核苷酸物质通常包含不同于腺嘌呤,鸟嘌呤,胸腺嘧啶,胞嘧啶或甲基胞嘧啶的核碱基,和/或包含与脱氧腺苷,脱氧鸟苷,胸苷,脱氧胞苷或脱氧甲基胞苷不同的核苷。如果靶多核苷酸是RNA,则修饰的多核苷酸中的不同核苷酸物质通常包含不同于腺嘌呤,鸟嘌呤,尿嘧啶,胞嘧啶或甲基胞嘧啶的核碱基和/或包含不同于腺苷,鸟苷,尿苷,胞苷或甲基胞苷的核苷。不同的核苷酸物质可以是上面讨论的任何通用核苷酸。

[0192] 聚合酶可以用不同核苷酸物质取代一种或多种核苷酸物质,所述不同核苷酸物质包含不存在于所述一种或多种核苷酸物质中的化学基团或原子。所述化学基团可以是丙炔基,硫基,氧基,甲基,羟甲基,甲酰基,羧基,羰基,苄基,炔丙基或炔丙胺基。

[0193] 聚合酶可以用不同核苷酸物质取代一种或多种核苷酸物质,所述不同核苷酸物质缺少存在于所述一种或多种核苷酸物质中的化学基团或原子。聚合酶可以用具有改变的电负性(altered electronegativity)的不同核苷酸物质取代一种或多种核苷酸物质。具有改变的电负性的不同核苷酸物质优选包含卤素原子。

[0194] 本方法优选进一步包括选择性地从修饰的多核苷酸中的一种或多种不同的核苷酸物质中去除核碱基。

[0195] 其他的表征方法

[0196] 在另一个实施方案中,通过检测随着聚合酶将核苷酸并入多核苷酸时释放的标记物质来表征多核苷酸。所述聚合酶使用多核苷酸作为模板。每种标记物质对各核苷酸都是特异的。使多核苷酸与本文所述的修饰的ClyA纳米孔、聚合酶和标记的核苷酸接触,使得当通过所述聚合酶将核苷酸添加到多核苷酸时,磷酸盐标记的物质被顺序地释放,其中磷酸盐物质含有对各核苷酸特异的标记。所述聚合酶可以是上面讨论的任何一种。使用孔检测磷酸盐标记物质并由此表征多核苷酸。这种类型的方法在欧洲申请13187149.3(公开为EP 2682460)中公开。上面讨论的任何实施例同样适用于本方法。

[0197] 试剂盒

[0198] 本发明的另一方面还提供了用于表征靶多核苷酸的试剂盒。该试剂盒包含本文所

述的任何一种修饰的ClyA纳米孔和膜的组件。膜优选由所述组件形成。孔优选存在于所述膜中。试剂盒可以包含上述任何膜的组件,例如两亲性层或三嵌段共聚物膜。

[0199] 所述试剂盒可以进一步包含多核苷酸结合蛋白。

[0200] 所述试剂盒可以进一步包含一个或多个用于将多核苷酸与膜偶联的锚。

[0201] 本发明的试剂盒可以另外包括能使上面描述的任何实施方式实施的一种或多种其他试剂或仪器。所述试剂或仪器包括以下的一种或多种:合适的缓冲剂(水溶液),用于从受体获得样本的工具(诸如导管或含有针的仪器),用于扩增和/或表达多核苷酸的工具,或者电压钳或膜片钳装置。存在于所述试剂盒中的试剂可以是干态的,液体样本使所述试剂重悬。所述试剂盒还可选地包括使该试剂盒能在本发明任一方法中使用或关于该方法可以用于哪些有机体的细节的说明书。

[0202] 装置

[0203] 本文描述的另一方面还提供了用于表征靶多核苷酸的装置。该装置包含多个如本文所述的修饰的ClyA纳米孔和多个膜。在一些实施方案中,所述多个修饰的ClyA纳米孔存在于所述多个膜中。在一些实施方案中,修饰的ClyA纳米孔和膜的数量相等。在一个实施方案中,单个修饰的ClyA纳米孔存在于每个膜中。

[0204] 该装置可以进一步包括用于执行如本文所述的任何方法的指令。该装置可以是用于多核苷酸分析的任何常规装置,例如阵列或芯片。本发明方法(例如用于表征靶多核苷酸的方法)的任何以上讨论的实施方式同样适用于本文所述的装置。该装置可以进一步包括本文所述的试剂盒中存在的任何特征。

[0205] 在一些实施方案中,所述装置设置成执行本文所述的任何方法,例如用于表征靶多核苷酸。

[0206] 在一个实施方案中,该装置包括:(a)传感器设备,其能够支撑多个修饰的ClyA纳米孔和膜并且可操作的使用纳米孔和膜执行多核苷酸表征;和(b)至少一个用于输送执行所述表征的材料的端口。

[0207] 或者,该装置可以包括:(a)传感器设备,其能够支撑多个修饰的ClyA纳米孔和膜并且可操作的使用纳米孔和膜执行多核苷酸表征;和(b)至少一个用于保存执行所述表征的材料的存储器。

[0208] 在另一个实施方案中,该装置可以包括:(a)传感器设备,其能够支撑膜和多个修饰的ClyA纳米孔和膜,并且可操作的使用孔和膜执行多核苷酸表征;(b)至少一个用于保存执行所述表征的材料的存储器;(c)流体系统,所述流体系统构造为可控地将来自所述至少一个储存器的材料供应到所述传感器设备;和(d)用于接收相应样品的一个或多个容器,所述流体系统被配置为将来自一个或多个容器的所述样品选择性地供应到所述传感器设备。

[0209] 该装置可以是任何国际申请PCT/GB08/004127(公开为W02009/077734),PCT/GB10/000789(公开为W02010/122293),国际申请PCT/GB10/002206(公开为W02011/067559)或国际申请号PCT/US99/25679(公为W0 00/28312)中描述的那些装置。

[0210] 无需进一步详述,相信本领域技术人员可以基于以上描述最大程度地利用本发明。因此,以下具体实施方式应理解为仅示例性,而不以任何方式限制本发明的其余部分。出于本文引用的目的或主题,本文列举的所有公开物都以引用的方式并入本文。

[0211] 实施例

[0212] 实施例1. 纳米孔的精确纳米级设计使DNA在生理离子强度下移位

[0213] 生物学中的许多重要过程涉及生物聚合物移位穿过纳米尺寸的孔,例如核酸穿过核孔转运,蛋白质穿过膜通道移位,以及病毒DNA注入靶细胞。此外,嵌入绝缘膜中的生物和人造纳米孔提供了研究这一过程的有用工具,并可用于快速DNA或蛋白质测序、单分子DNA测序和分析以及生物标志物感测。已经具体研究了DNA移位穿过纳米孔的机制。几种门(portal)噬菌体蛋白质的晶体结构显示,在DNA包装和注射期间,dsDNA移位穿过窄的纳米孔( $\sim 3.5\text{nm}$ ),所述窄的纳米孔具有被正电荷的环所修饰的强负电荷表面。所述纳米孔的带负电荷内表面被认为促进了带负电荷的DNA的滑动,虽然正电荷的作用被认为有助于该过程。在该实施例中,发现只有正电荷的两个环被设计在纳米孔的宽入口和中间部分,才能够在生理离子强度下观察到DNA电泳移位穿过ClyA纳米孔,所述ClyA纳米孔具有相同折叠,尺寸和整体内部电荷的门蛋白质(portal proteins)。令人惊讶的是,纳米孔的具有强带负电的 $3.3\text{nm}$ 内部收缩部不需要修饰。研究结果表明,设计的正电荷对于对齐DNA以克服DNA移位穿过窄的收缩部的熵和静电能垒是重要的。不希望受理论束缚,为了移位穿过具有负电荷密度的窄纳米孔,应当将DNA分子定向。

[0214] 流经重组成脂质膜的生物纳米孔的离子电流被用来识别小分子或折叠蛋白质,并监测单分子水平的化学或酶反应。DNA穿过重组成人造膜的纳米孔的电泳移位,对于DNA测序和生物标记物识别等实际应用具有很大潜力。 $\phi 29$ 门蛋白质,其不是膜蛋白本身,被发现插入到黑色脂质双分子层中,并且这种纳米孔在 $1.0/0.5\text{mM NaCl}$ 下电泳移位dsDNA。然而,允许膜插入的纳米孔的确切疏水性修饰是未知的。事实上, $\phi 29$ 纳米孔偶尔从脂质膜释放,因此在实际应用中受到限制。dsDNA已显示移位穿过在固态膜上制备的人造纳米孔,所述人造纳米孔除了原子-薄材料(例如石墨烯或二硫化钼的双分子层)之外,多数具有负的内表面电荷。在这种半径与溶液的德拜长度相当的纳米孔中,由内部纳米孔壁上的双电层(EDL)产生的表面电势重叠,导致对于DNA进入纳米孔的大的静电能垒。因此,在生理离子强度下使用大纳米孔( $10\text{nm}$ )使DNA移位穿过固态纳米孔,或在 $340\text{nM}$ 盐浓度下或在不对称盐浓度下使用小纳米孔( $\sim 3.5\text{nm}$ ),使DNA移位穿过固态纳米孔。另外,在生理离子强度下DNA移位穿过直径与DNA尺寸(dsDNA的B型约为 $2.2\text{nm}$ ,ssDNA约为 $1\text{nm}$ )相当的固态纳米孔仍有待观察。

[0215] ClyA纳米孔,具有 $\sim 3.3\text{nm}$ 的内部收缩部的十二聚体蛋白(图1,版A),已被用作研究折叠蛋白的工具。尽管在 $2.5\text{M NaCl}$ 溶液中观察到dsDNA移位穿过纳米孔,但孔的强带负电内部(图1,版A)阻止了在较低离子强度下的DNA移位。在该实施例中,ClyA纳米孔被设计,使其能够在生理离子强度下移位DNA。这在许多应用中很有用,其中分子和DNA之间的静电相互作用很重要,例如在DNA测序或制图中酶被用于控制DNA移位穿过纳米孔或研究DNA-蛋白相互作用。在纳米孔的更宽的顺式侧添加两个正电荷环之后观察到DNA移位,而纳米孔的更收缩的反式入口的修饰没有提高DNA移位的效率。此外,这些修饰并未改变纳米孔的离子选择性并反映了 $\phi 29$ 门蛋白质的电荷分布。进一步地,设计的孔允许DNA仅从纳米孔的宽的一侧移位。有趣的是,许多在DNA上滑动的蛋白质表现出类似于人工ClyA纳米孔的表面电荷,表明正电荷和负电荷的交替可能提供了用于改善DNA穿过纳米孔移位的一般机制。该实施例表明,纳米孔的形状和内部表面电荷的精确设计对于DNA移位和滑动穿过直径与DNA相似的纳米级孔是重要的。

[0216] 结果

[0217] 设计ClyA纳米孔以捕获DNA

[0218] ClyA-AS (图1, 版A; 图11, 版A) 是来自伤寒沙门菌 (*Salmonella typhi*) 的溶细胞素A的设计版本, 因其在平面脂质双分子层中的有利特性而被选择, 在所述平面脂质双分子层中ssDNA或dsDNA的移位仅在2.0M以上的NaCl离子强度下观察到。最有可能的是, 在低离子强度下, 纳米孔内部的强负静电势 (图11, 版B) 阻止了DNA进入和移位, 而在高离子强度下, 纳米孔表面的电荷被有效地屏蔽。为了诱导纳米孔在生理离子强度下捕获DNA, ClyA-AS纳米孔的内部电荷被修饰 (表1和2和图1, 版A; 图11, 版A)。有时ClyA变体显示出开孔电导 (门控) 的瞬时减少。作为门控的测量值, 触发电压 (VG) 被限定为施加电压, 在所述施加电压下, 使用保持打开30秒钟时间间隔的典型纳米孔 (表1)。通过添加1 $\mu$ M 90mer的3'-生物素化ssDNA分子 (图1, 版A, 表3), 随后添加等摩尔浓度的其互补链 (图1, 版B, 表3), 并最后添加中性抗生物素蛋白 (1.2 $\mu$ M, 单体), 在VG下测试DNA穿过修饰的纳米孔的移位。

[0219] 在ClyA-AS的顺式入口引入精氨酸残基形式的正电荷的单环 (S110R, ClyA-R, 图1, 版A; 图11, 版A), 然后将纳米孔的三部分: 顺式入口, 中间部分和反式收缩部进行修饰 (图1, 版A; 图11, 版A)。在ClyA-R的顺式开口处用正电荷残基取代中性残基, 在150mM NaCl中未显示出DNA移位 (表1, 表2)。在顺式开口处添加正电荷, 显示在150mM NaCl既没有通道插入平面脂质双分子层 (ClyA-R-E106R和ClyA-R-D114R), 也没有DNA移位 (ClyA-R-D122R和ClyA-R-D129R)。当位置64的带负电的谷氨酸残基被精氨酸 (D64R, ClyA-RR) 取代时, 但在附近位置的中性侧链被精氨酸取代时 (Q56R) 时, ClyA-R纳米孔中间部分的精氨酸环诱导ssDNA (图1, 版C) 和dsDNA (图1, 版D) 的移位。用精氨酸取代任一邻近位置的中性侧链 (Q56R), 跨膜区中的带负电荷的残基的去除 (ClyA-R-E11S), 或带正电荷的残基的添加 (ClyA-R-Q8K), 在150mM NaCl溶液中未诱导DNA移位事件 (图6)。令人惊讶的是, 在ClyA-R的中间部分和反式入口中用带正电荷残基取代中性残基 (ClyA-R-Q56R-Q8K) 也不诱导DNA移位事件 (图6)。除了ClyA-R-D129R外的所有测试的突变体均降低了触发电压 (gating voltage) (表1)。ClyA-RR是唯一的ClyA突变体, 在将ssDNA或dsDNA添加到纳米孔的顺式侧之后, 显示了DNA诱导的电流事件 (+70mV, 图1版C-D和图6)。尽管观察到只有ClyA-RR允许DNA移位, 但ClyA-RR, ClyA-R和ClyA-AS都显示出相同的离子选择性 (各自为 $P_{Na^+}/P_{Cl^-} = 1.9 \pm 0.7, 2.0 \pm 1.6, 1.9 \pm 0.9$ , 表4), 表明纳米孔的离子选择性由纳米孔的跨膜区的电荷分布支配, 并且不是通过穿过纳米孔的增强的电渗流诱导。

[0220] 更一般地, 区域A (如图20所示) 中的第一氨基酸的取代可至少具有 $\Delta 1$ 的增加的正电荷 (即通过带正电荷的氨基酸取代中性氨基酸) 和区域B的取代可具有至少 $\Delta 2$ 的增加的正电荷 (即通过带正电的氨基酸取代带负电的氨基酸)。

[0221] 为了更深入地了解由两个额外的精氨酸环引起的静电势的变化, 使用始于大肠杆菌ClyA晶体结构的VMD (Humphrey et al. J. Mol. Graphics (1996) 14:33-38) 和NAMD (Phillips et al., J. Comput. Chem. (2005) 26:1781-1802) 构建ClyA-AS和ClyA-RR的全原子同源性模型。使用例如Baker等, PNAS (2001) 98:10037-10041; Dolinsky等, Nucleic Acids Res. (2004) 32:W665-W667; 和Dolinsky等, Nucleic Acid Res. (2007) 35:W522-W525) 描述的自适应泊松-波尔兹曼解算器 (APBS) 计算150mM NaCl中的两种孔的电势分布 (图11, 版B)。在ClyA-AS中, 孔中心的电势被发现, 从顺式入口移动穿过中间部分并到达反式入口越来越负 (分别平均为-2.6, -4.8和-15.2mV)。在ClyA-RR的情况下, 可以观察到在孔

的顺式入口和中间部分处电势的升高(分别平均为-0.3和-1.1mV)。在反式收缩部的电势似乎进一步下降至平均-17.3mV。应该注意的是,这些值是在没有施加外部偏差时计算的。

[0222] 表1:设计的ClyA纳米孔变体的电学性质。通过将~0.1ng寡聚蛋白添加到顺式侧腔来测试纳米孔的活性。负活性表明没有观察到通道插入。VG是触发电压,代表在30秒时间范围内未观察到门控事件的最高施加电压。DNA移位表明可以形成dsDNA轮烷。每个数据点是至少三次实验的平均值,误差是标准偏差。实验在0.15M NaCl,15mM Tris HCl,pH 7.5的溶液中进行。

[0223] 表1

孔变体	双分子 层活性	IO+100mV IO1-00Mv(pA)	整流比	VG (mV)	DNA 捕获 (顺式)	DNA 移位 (顺式)
ClyA-AS	+	+190±13 -138±6	1.4±0.1	+100	-	-
[0224] ClyA-AS- S110R(ClyA-R)	+	+198±1 -127±2	1.6±0.0	+100	-	-
ClyA-R- E106R	-	-	-	-	-	-
ClyA-R- D114R	-	-	-	-	-	-
ClyA-R- D122R	+	+207±2 -99.8±2	2.1±0.1	+50	-	-
ClyA-R- E129R	+	+171±25 -161±24	1.1±0.2	+100	-	-
ClyA-R- D64R(ClyA-RR)	+	+198±8 -110±4	1.8±0.1	+70	+	cis
[0225] ClyA-R- Q56R	+	+202±8 -128±3	1.6±0.1	+50	-	-
ClyA-R- Q8K	+	+202±15 -147±18	1.4±0.2	+50	-	-
ClyA-R- E11S	+	+194±4 -154±0	1.3±0.03	+70	-	-
ClyA-R- Q56R-Q8K	+	+207±20- 150±15	1.4±0.2	+50	-	-

[0226] 作为DNA移位证据的DNA轮烷

[0227] 轮烷是哑铃状分子,由大环绕被两个塞子(stoppers)锁定的链形成。在该实施例中,在150mM NaCl溶液中形成的两个纳米孔/DNA轮烷证明ssDNA和dsDNA移位穿过纳米孔。使用100mer的5'-生物素化ssDNA分子作为起始链(2a,表3)添加到顺式侧腔形成第一轮烷。使用3'生物素化的59个碱基对的dsDNA分子形成第二轮烷,所述dsDNA分子用31个碱基的3'生物素突出端(1a/1c,表3)延伸。通过在纳米孔的另一侧添加另一生物素化的ssDNA分子2b

(50mer, 5'-生物素化的) 或 1d (31mer, 3'-生物素化的), 其被设计成分别与 2a 或 1a/1c 的突出端杂交来锁定轮烷。顺式和反式溶液均含有中性抗生物素蛋白 (NA, 0.3  $\mu$ M), 其与生物素复合并阻止 DNA 链完全移位穿过纳米孔。

[0228] 在 150mM NaCl 和 +50mV 下, ssDNA 和 dsDNA/ssDNA 穿过纳米孔 (分别 IRES+5092  $\pm$  0.02 和 0.84  $\pm$  0.07, N=3), 但当施加的电势反转至 -50mV 时, 其从所述孔排出 (图 2, 版 A-B)。随后将 DNA: 中性抗生物素蛋白塞子添加到反式溶液中诱导了在两种电势下的永久性阻断, 表明组装了 DNA 轮烷, 并且显示两种链都移位纳米孔。在负施加电势下, 对于两种轮烷, 阻断的离子电流都高于开孔电流 (对于 ssDNA 和 dsDNA/ssDNA 链各自 IRES-50 = 1.16  $\pm$  0.03 和 1.11  $\pm$  0.06, N=3 个独立的纳米孔实验, 图 2, 版 A-B; 图 12, 版 A-B)。这种效应之前在低离子强度下 DNA 移位穿过 10nm 固态纳米孔观察到并用被 DNA 封闭的孔内的抗衡离子的积聚来解释。相比之下, 在正施加电势下, 开孔电流高于阻断电流 (图 1, 版 C-D 和图 2, 版 A-B, 图 12, 版 A-B), 表明在这种构型中, 中性抗生物素蛋白可能与纳米孔的内腔相互作用并且在纳米孔的顺式和反式侧 DNA 上抗衡离子的积聚不同。

[0229] DNA 捕获/穿入和移位依赖于溶液的离子强度

[0230] 捕获率  $k_{on}$  是事件间时间  $\tau_{on}$  的倒数 (表 7, +70mV, 1  $\mu$ M DNA), 其随着用于 ssDNA 和 dsDNA 的溶液的德拜长度 ( $\lambda_D$ ) 而增加 (图 3, 版 B-C; 图 13, 版 A-B)。然而, 尽管 dsDNA 捕获率随着  $\lambda_D$  线性增加 (图 13, 版 A), 但 ssDNA 捕获率随着  $\lambda_D$  呈指数增加 (图 13, 版 B)。因此, 这表明 dsDNA 和 ssDNA 的不同捕获机制。顺式侧添加的 dsDNA 移位的频率随着溶液的德拜长度线性增加 (+70mV, 图 3 版 A, 图 7 和图 8), 表明 DNA 和纳米孔之间的静电相互作用对于 DNA 进入和移位是重要的。如之前对固态纳米孔报道的, 随着溶液的离子强度降低 (例如, 从 2.5M NaCl 中的 0.78  $\pm$  0.09 到 150mM NaCl 中的 0.92  $\pm$  0.02), 被 DNA 封闭的纳米孔的剩余电流增加。有趣的是, 发现了 DNA 阻断的 IRES 与溶液的德拜长度之间的线性关系 (图 3, 版 B-C)。对于与中性抗生物素蛋白复合的 dsDNA, 剩余电流比游离 DNA 移位过程低  $\sim$  10%, 表明中性抗生物素蛋白对阻断的整体离子电流有贡献, 最可能是通过与纳米孔腔相互作用而有贡献。

[0231] ssDNA 移位的频率随着溶液的德拜长度以指数方式增长 ( $R^2=0.99$ ) 而不是以线性方式 ( $R^2=0.78$ ) (图 3, 版 C), 表明除了 C1yA 腔中的被设计的正电荷和 DNA 之间的相互作用之外的其他因素对进入和/或移位纳米孔起重要作用。在 150mM NaCl 时, 与中性抗生物素蛋白复合的 ssDNA 分子显示了对 C1yA-RR 纳米孔的永久性阻断, 而在 1M NaCl 或更高时, 阻断是瞬时的 (图 3, 版 D, 图 10)。对这些数据可能的解释是, 在高离子强度下, ssDNA 从顺式侧进入并离开孔。在  $\geq$  1M 的离子强度下, 存在和不存在中性抗生物素蛋白时, ssDNA 的 IRES 值相同 (图 3, 版 C; 图 10), 表明在这些条件下, ssDNA 可能没有完全穿过纳米孔, 阻止了中性抗生物素蛋白与 C1yA 的内腔相互作用。

[0232] DNA 单向进入 C1yA 纳米孔

[0233] 在 150mM NaCl 溶液中以及负施加电势 (高达 -100mV) 下, 向 C1yA-RR 腔的顺式侧和反式侧添加 1  $\mu$ M ssDNA 或 dsDNA 不会诱导 DNA 阻断, 表明 DNA 不能从所述纳米孔的反式入口进入所述纳米孔 (图 4, 版 A)。在正施加偏置下, 在高于  $\sim$  +50mV 的电势下出现电流阻断, 表明 ssDNA 从纳米孔的顺式侧移位的电压阈值的存在。然而, 在 1M NaCl 溶液中观察到的 DNA 从反式侧腔进入 (图 4, 版 B) 和移位 (图 9) 表明阻止在 150mM NaCl 下, 阻止从反式侧腔移位的能垒本质上是静电。

[0234] 为了观察在生理离子强度下DNA从反式侧腔进入,C1yA-RR纳米孔的跨膜区的电荷被重构(表5和图10)。发现在负施加电势下,将1 $\mu$ M的dsDNA1添加到反式侧腔时,纳米孔跨膜区域中带负电荷的残基的取代不引起电流阻断(图10),表明在这些条件下从C1yA-RR纳米孔的顺式侧和反式侧的DNA移位的相对较大的非对称势垒。

[0235] 讨论

[0236] 精确的纳米孔工程支持生理离子强度下的DNA移位

[0237] 在该实施例中,C1yA纳米孔被设计以允许在生理离子强度下的DNA的电泳移位。当在C1yA纳米孔的入口和中间部分引入两组正电荷时,观察到DNA移位(图11,版A)。令人惊讶的是,为DNA移位提供最高熵和静电能垒的纳米孔的反式入口不需要修饰(图11,版A-B)。此外,尽管对C1yA的反式入口进行了很大程度的重构(表1-2),但只有当从纳米孔的更宽的顺式入口引发时才能观察到DNA移位。此外,dsDNA移位穿过C1yA-RR纳米孔的频率随溶液的德拜长度而增加(图13,版A),表明dsDNA与C1yA-RR的顺式入口的有利的静电相互作用相对于DNA与纳米孔收缩部的不利的静电排斥占主导地位。应该指出的是,在测试的离子强度范围内,dsDNA的刚度没有显著变化。此外,随着离子强度降低,增加的电渗流不能解释增加的DNA移位的频率,因为电渗流与DNA进入和移位相反。因此,这些数据表明,纳米孔的顺式侧腔对引发DNA移位穿过纳米孔的收缩部是重要的。

[0238] DNA分子受电驱动力移位穿过纳米孔,并且由纳米孔内的电渗流 EOF 产生的流体动力粘滞阻力与DNA的移位相反。C1yA和大多数固态纳米孔具有负表面电荷,其通过与通常称为双电层(EDL)的表面直接接触的阳离子的层进行静电平衡。在施加的电场下,EDL中离子的移动诱导抗衡离子的优先移位,这反过来产生EOF并获得纳米孔离子选择性(例如,C1yA-AS PNa/Pk=1.9,表2)。由于通过电解质的筛选,EDL力在扩散层上以指数方式衰减。该力的范围由德拜长度给出,其强度由表面电势给出。在窄纳米孔中,特别是在低盐浓度的情况下,包括扩散层的EDL的厚度可以与纳米孔的尺寸相当,从而产生重叠的EDL。在这种情况下,接近这种纳米孔的DNA分子(直径2.2nm)将经历强表面电势,对于具有负表面电荷的纳米孔将阻止DNA进入所述纳米孔。

[0239] dsDNA和ssDNA移位穿过C1yA纳米孔的机制

[0240] C1yA可以近似为圆柱形顺式侧腔(5.5nm直径和10nm长度)接着较小且带负电荷的反式收缩部(3.3nm直径和3.0nm长度,图1),其被预期为对抗DNA移位的主要的电泳的和熵能垒。令人惊讶的是,当将一组正电荷添加至纳米孔(C1yA-RR)的顺式侧腔时,观察到DNA移位穿过C1yA纳米孔;而所述纳米孔的收缩部不需要任何修饰。尽管对C1yA的反式入口进行了广泛的修饰(表2),但只有当从纳米孔的更宽的顺式侧引发时才可以观察到DNA移位,表明纳米孔的顺式侧腔对于引发DNA移位穿过纳米孔是重要的。双链DNA移位穿过C1yA-RR纳米孔的频率随着溶液的德拜长度线性增加(图3,版A),表明dsDNA与C1yA-RR中设计的电荷的静电相互作用更有利,而不是对抗移位过程更有利。提出了一种模型,其中当dsDNA链通过纳米孔的顺式侧腔预对齐时,获得在生理离子强度下dsDNA穿过反式收缩部的移位(图5,版A)。在该视图中,dsDNA最初与顺式入口处的电荷相互作用,然后进入孔的腔中,在那里其进一步与所述纳米孔的中间部分的精氨酸残基相互作用(图1,版A)。这些静电相互作用“抓住”DNA的磷酸盐基团,阻止DNA退回到顺式侧溶液。在该构象中,dsDNA被对齐以进入反式收缩部,在所述反式收缩部电泳力是最强的,允许DNA平滑移位穿过所述纳米孔(图5,版A)。

[0241] 观察到ssDNA阻断的德拜长度依赖性很好地符合单一指数(图3,版A),而不是对于dsDNA所观察到的线性函数,表明与dsDNA相比,附加因素影响ssDNA的移位。在实验中,DNA伸直长度,即当其完全伸展时的总长度,低于dsDNA( $\sim 50\text{nm}$ )的持续长度,表明dsDNA分子以刚性棒的形式移位(图5,版A)。相反,ssDNA具有卷曲结构(持续长度 $\sim 1.5\text{nm}$ ),其回转半径为 $\sim 6\text{nm}$ ,该回转半径是该聚合物卷材中任意点与其质心的平均平方距离。由于所述回转半径类似于纳米孔顺式入口的直径(图5,版B),ssDNA很可能以部分卷曲结构进入纳米孔的顺式侧(图5,版B)。随着ssDNA从腔顺式侧移动到反式侧,它然后必然逐渐展开,以便穿过所述纳米孔的反式收缩部,然后在反式侧重新卷曲(图5,版B)。这种与转变过程中的DNA的构象变化相关的熵展开和重新卷曲力(其在高离子强度下促进固定化的ssDNA从纳米孔逆着施加的电势的顺式排出(图3,版D)),随着溶液的离子强度的降低而降低,由于所述溶液的离子强度降低增加了DNA移位的效率。应该指出,被DNA封闭的纳米孔内的离子浓度和德拜长度是未知的。尽管如此,两者都与纳米孔电流相关,而纳米孔电流又与电解质主体的浓度相关(图3,版B和C)。

[0242] DNA移位机制:dsDNA捕获是扩散受限的,ssDNA捕获是反应受限的

[0243] 在不同盐浓度下的DNA移位实验显示了dsDNA和ssDNA的两种不同的捕获机制(分别为图13,版A-B和图14,版A-B)。dsDNA的行为与扩散限制的捕获过程一致。这是因为在这项工作中使用的dsDNA短于其持续长度(150bp),并表现为刚性均匀带电荷的棒。在捕获半径内(对于 $0.5\text{nm}$ 的 $\lambda_D$ ,距离纳米孔中心约 $50\text{nm}$ ),电场吸引DNA朝向孔并沿着场线使其排列,使得它用一端撞到孔入口(图14,版A,i)。一旦进入孔中,设计的电荷与DNA相互作用,防止DNA退回至顺式溶液(图14,版A,ii-iv)。因此,DNA捕获的动力学可以通过电泳起源的纯引力势中的扩散粒子的动力学来近似。在这种情况下,dsDNA的电泳迁移率与溶液的德拜长度成比例并且相应的漂移-扩散方程可以被精确地求解,这将在下面进一步详细描述。通过用圆柱体长度 $l=13\text{nm}$ 和捕获直径 $d=6\text{nm}$ 近似ClyA纳米孔的几何形状(图11,版A),捕获频率可以通过以下来估计:

$$[0244] \quad k_{\text{on}} \sim 14\lambda_D (\text{s nm } \mu\text{M})^{-1}$$

[0245] 这与 $\lambda_D$ 的实验数据非常吻合(在高盐浓度下,图13,版A)。这是惊人的,因为没有使用拟合参数。然而,在这种比较中应该小心一些,因为ClyA的几何形状明显偏离完美的圆柱体,所以孔参数的选择在一定程度上是武断的。在低盐浓度( $0.15\text{MNaCl}$ , $\lambda_D=0.8\text{nm}$ )下,捕获率高于以上公式所预测的(图13,版A)。很可能,在ClyA-RR入口处的正电荷(在所述模型中没有考虑到)加快了低盐浓度下的捕获速度,而在较高的盐浓度下,这些电荷更有效地被屏蔽。

[0246] 对于ssDNA, $k_{\text{on}}$ 和 $\lambda_D$ 之间的关系呈指数关系,这与能垒穿越一致(反应限制的过程)。在溶液中,ssDNA假定以卷曲构象,随DNA接近纳米孔,被电泳力拉向所述纳米孔(图14,版B,i)。然而,在所述孔的入口附近,如果链的一端面向所述孔的入口并且如果ssDNA是展开的(图14,版B,iii,iv),才能发生成功的移位事件(图14,版B,ii)。熵源的这种额外排斥力有效地产生必须在移位之前穿过的能垒。已经讨论了这种能垒限制的移位的理论,并且基于一般理由,捕获率由下式给出: $k_{\text{on}} = \omega e^{-\Delta F_b/k_B T}$

[0247] 这里, $\Delta F_b$ 是能垒高度, $\omega$ 是对于能垒穿越的特征尝试速率。指数因子给出了成功穿越事件的概率。估计可以完成来自模型输入的 $\Delta F_b$ ;结果表明成功移位的概率包含与电泳

迁移率成比例的项,其又与 $\lambda_D$ 成比例。这将解释 $k_{on}$ 对 $\lambda_D$ 的指数依赖性(图13,版B)。应该注意的是,尽管 $k_{on}$ 是从事件间隔时间的倒数获得的,但并非所有测量的电流阻断都必然描述移位事件。部分的这些阻断可能是由于DNA链进入然后撤回至顺式侧(图14,版B,iii至i)。尽管如此,轮烷的形成表明至少一些分子成功地移位。无论如何,导致 $k_{on}$ 对 $\lambda_D$ 的指数依赖的理由仍然有效。

#### [0248] 生物学意义

[0249] 有趣的是,允许DNA移位穿过ClyA纳米孔的修饰也在蛋白质中观察到,所述蛋白质的生物功能是沿DNA滑动。在噬菌体中,通过包装蛋白质将DNA转移到衣壳中,所述包装蛋白质对齐并推动DNA穿过门蛋白质,所述门蛋白质具有与ClyA类似的尺寸、化学计量、内部表面电荷和与内部构造尺寸。负的内部表面电荷似乎对于平滑移位DNA穿过门蛋白质很重要,正如包围和沿着DNA滑动的其他蛋白质(例如 $\beta$ -夹子蛋白质)中观察到的那样。门蛋白质和 $\beta$ -夹子蛋白质也具有带正电荷的环,其已经被提出在DNA包装中通过与移位DNA的带负电荷的磷酸骨架相互作用而起直接作用。当在纳米孔的顺式入口和中间部分引入两个带正电荷的残基环时,将DNA对齐穿过窄的和带负电荷的收缩部的通道,可以观察到DNA电泳移位穿过ClyA纳米孔。在没有这种相互作用的情况下,也就是说,在从反式侧穿入期间,不能观察到DNA移位。因此,本文呈现的结果表明,在连接蛋白中,这种正电荷的环对于启动DNA从衣壳排出进入感染细胞中可能是重要的。

[0250] 本实施例中提出了一种设计的ClyA十二聚体纳米孔ClyA-RR,在引入两个正电荷环后,在生理离子强度下移位dsDNA和ssDNA。ClyA-RR可用于研究单分子水平的蛋白质-DNA相互作用,并可用于DNA制图和测序应用,其中一种酶控制着核酸穿过纳米孔的移位。已发现在纳米孔的较宽入口(顺式侧)引入正电荷的环(吸引力相互作用)对于诱导DNA移位穿过窄的且带负电荷的反式收缩部是重要的。令人惊讶的是,收缩部本身并不需要修饰。这些结果表明,在纳米孔的入口和中间的吸引力相互作用对于“抓取”并定向DNA以有效地电泳驱动滑动其穿过窄的且带负电荷的反式收缩部是重要的。有趣的是,ClyA-RR中的电荷分布反映在病毒门蛋白质中,表明生物纳米孔的精确设计对于在病毒衣壳内外的DNA的有效包装和排出是重要的。此外,dsDNA和ssDNA各自移位频率的线性和指数离子强度依赖性表明可能机制即dsDNA捕获遵循扩散限制的过程,而ssDNA捕获是反应限制的过程。还显示ssDNA以卷曲结构进入纳米孔,所述卷曲结构需要展开以便移位穿过纳米孔的收缩部。这些发现可用于帮助固态纳米孔的设计。例如,具有有利表面电荷和与放置在纳米孔上方的DNA的回转半径相似的直径的纳米级腔应该有利于DNA的移位,尤其是在低离子强度下。此外,发现允许DNA移位的ClyA纳米孔的修饰反映在病毒门蛋白质中,表明生物纳米孔的精确设计对于病毒衣壳内外的DNA的有效包装和排出是重要的。

#### [0251] 材料和方法

[0252] DNA购自Integrated DNA Technologies(IDT)。中性抗生物素蛋白从Thermo Fisher获得,并从Avanti Polar Lipids获得1,2-二植烷酰-sn-丙三基-3-磷酸胆碱。 $\beta$ -十二烷基麦芽糖苷(DDM)购自GLYCON Biochemicals GmbH。除非另有说明,酶购自Fermentas并且所有其他材料购自Sigma。

[0253] 蛋白质纯化。通过使用Soskine等人,J. Am. Chem. Soc. 1995, 116, 2975-2852和Miyazaki等人,Methods Enzymol. (2011) 498:399-406所述的“大引物”方法对ClyA-AS基因

进行单点突变。通过使用pT7质粒在E. *cloni*®EXPRESS BL21 (DE3) 细胞中表达ClyA。如Waduge等, ACS Nano (2015) 9:7352-7359中所述, 通过使用Ni-NTA亲和层析纯化单体并在0.5%β-十二烷基麦芽糖苷 (GLYCON Biochemicals GmbH) 存在下寡聚。在大肠杆菌BL21细胞中表达单体 (含有C端寡聚组氨酸标签), 并使用Ni-NTA亲和缓冲液 (150mM NaCl, 15mM Tris HCl, pH 7.5, 0.2%DDM和1mM EDTA) 纯化可溶解组分并在4℃下储存。

[0254] DNA制备。通过首先混合等摩尔浓度的1a和1b来制备dsDNA 1 (表3)。混合物温度升至95℃, 并且该温度以规律时间间隔下降。使用含有固定在琼脂糖珠上的单体抗生物素蛋白的生物素结合柱 (Thermo Scientific Pierce), 用亲和层析从过量的ssDNA纯化DNA。然后根据制造商的说明书从所述柱中洗脱dsDNA。将洗脱级分浓缩并使用PCR快速纯化试剂盒 (QIAGEN) 进一步纯化。典型地, 获得0.2μg/mL的DNA浓度。1a/1c双链体按所说明的退火1小时, 但未纯化。

[0255] 离子渗透性。通过在对称条件下 (顺式和反式溶液中均为150mM NaCl, 15mM Tris-HCl pH 7.5) 将ClyA添加到顺式侧腔中来收集不对称条件下的I-V曲线 (表6)。然后平衡电极, 并且通过将等分的5M NaCl储备溶液添加到顺式侧腔中, 将顺式侧中的电解质浓度提高到1M。通过添加与加入到顺式侧中溶液等体积的缓冲液来调节反式侧腔的体积, 所使用的缓冲液与顺式侧中溶液 (150mM NaCl) 的缓冲液相同。

[0256] 利用Goldman-Hodgkin-Katz方程 (以下) 使用测量的反向电势 ( $V_r$ ) 值 (其从I-V曲线外推), 计算渗透率比率 ( $P_{Na^+}/P_{Cl^-}$ , 表4)。

$$[0257] \quad P_{Na^+}/P_{Cl^-} = \frac{[a_{Cl^-}]_{trans} - [a_{Cl^-}]_{cis} e^{V_r F/RT}}{[a_{Na^+}]_{trans} e^{V_r F/RT} - [a_{Na^+}]_{cis}}$$

[0258] R是通用气体常数 ( $8.314 \text{ J K}^{-1} \text{ mol}^{-1}$ ), T是开氏温度, F是法拉第常数 ( $96485 \text{ C mol}^{-1}$ ),  $P_{Na^+}$  和  $P_{Cl^-}$  是  $Na^+$  或  $Cl^-$  离子的相对膜透过率, 并且  $a_{Na^+}$  和  $a_{Cl^-}$  是它们各自的活性。顺式侧腔接地 (ground)。使用含2.5%NaCl的2.5%琼脂糖桥的Ag/AgCl电极进行所有实验。

[0259] 电子记录。通过从双植烷酰-sn-丙三基-3-胆碱磷酸 (Avanti Polar Lipids, Alabaster, AL) 形成的平面双分子层记录来测量离子电流。如Ho等人 Sci. Adv. (2015) 1, e1500905; 和Maglia等人, Methods Enzymol. (2010) 475:591-623中描述的, 使用膜片钳放大器 (Axopatch 200B, Axon Instruments, Foster City, CA), 用浸没在琼脂糖 (3%w/v低熔点琼脂糖的2.5M NaCl缓冲液) 中的Ag/AgCl电极测量电流。通过测量电流与施加电压的关系 (I-V曲线, 21s中以10mV梯度从-100至+100mV施加电势, 图6, 10和表5) 来表征单通道。在0.15M NaCl中, 通过施加2kHz低通贝塞尔滤波器并使用10kHz采样率记录离子电流。在较高盐浓度下, 离子电流在50kHz采样, 并且低通贝塞尔滤波器设定为10kHz。采集后用4kHz贝塞尔数字滤波器对0.3和0.5M NaCl下的电流迹线进行过滤 (图16, 17)。使用不同的滤波频率影响检测到的事件的总数。例如, 将2kHz数字高斯滤波器应用于50kHz下采样的迹线, 而应用10kHz贝塞尔滤波器增加约50%的事件间时间 (从221秒到311毫秒, 0.17μMdsDNA, 1M NaCl, 平均停留时间为0.12毫秒)。因此, 为了测试过度滤波对DNA捕获频率的德拜长度依赖性的影响, 在对所有电流迹线施加1kHz高斯滤波器之后, 绘制图13版A中描述的数据 (图18)。发现ssDNA和dsDNA阻断分别很好地符合指数和线性回归 (图18)。

[0260] 数据分析。通过使用Clampfit软件 (Molecular Devices) 的“单通道搜索”功能, 使

用0.05ms的数据采集阈值,单独地收集电流阻断事件。从高斯拟合到全点直方图计算得到开口和阻塞的孔电流。通过分开开孔电流值的阻塞孔电流值来计算剩余电流。从来自DNA阻断停留时间的事件直方图的双指数拟合来计算DNA移位停留时间( $\tau_{\text{off}}$ )值,而使用来自事件间时间的对数直方图的指数对数概率拟合来计算( $\tau_{\text{on}}$ )值(图13,表7和图16,17)。误差表明来自至少三个独立纳米孔实验的平均值的标准偏差,其数目由N表示。

[0261] 根据本文所述的一个实施方案的关于制备修饰的ClyA纳米孔亚单位多肽的额外信息

[0262] 通过使用“大引物(mega primer)”方法对ClyA-AS基因进行单点突变。通常,实施两个PCR循环以制备新的DNA构建体:在第一个PCR反应中,用两种引物扩增质粒DNA:正向引物是长度为20-30个碱基的寡核苷酸,其具有碱基取代,反向引物是T7启动子或T7终止子。对于跨膜区的突变,反向引物是与蛋白质序列中部的片段互补的25mer寡聚体(表3)。将含有大引物(200-300bp)的PCR产物加载到琼脂糖凝胶(2%琼脂糖/TAE和结晶紫)中,将所述大引物切掉并使用PCR快速纯化试剂盒(QIAGEN)纯化。将5 $\mu$ L纯化的大引物加载到2%琼脂糖/TAE凝胶上以检查纯度,并且使用5-10 $\mu$ L大引物进行第二次PCR反应。然后首先用DpnI(1-2小时,37 $^{\circ}$ C,快速消化DpnI,Fermenthas)消化第二次PCR产物以消除ClyA-AS模板DNA,然后 $\sim$ 1 $\mu$ L被用来与电感受细胞E. *cloni*®EXPRESS BL21(DE3)(制造者)进行转化。

[0263] 关于DNA制备的额外信息

[0264] 通过培育1a,3'-生物素化的ssDNA分子(表3)和20%过量的互补ssDNA 1b(表3)形成dsDNA 1。将温度升至95 $^{\circ}$ C保持1分钟,然后逐步降至室温。在约70 $^{\circ}$ C的估计退火温度下,温度以2 $^{\circ}$ C步幅降至21 $^{\circ}$ C。每一步保持1分钟。然后使用含有固定在琼脂糖珠上的单体抗生物素蛋白的生物素结合柱(Thermo Scientific Pierce),通过亲和层析从过量的ssDNA纯化DNA。然后根据制造商说明书用生物素阻断/洗脱缓冲液洗脱dsDNA。将洗脱级分浓缩并使用PCR快速纯化试剂盒(QIAGEN)进一步纯化。通常,获得0.2 $\mu$ g/mL的DNA浓度。通过使用2%琼脂糖凝胶的TAE缓冲液检查dsDNA的尺寸和纯度,并进行光谱定量分析。将纯化的dsDNA在1mM EDTA存在下在-20 $^{\circ}$ C下保存。通过用等摩尔浓度的1c培育3'-生物素化的ssDNA分子(1a,表3)而形成1a:1c。温度升至95 $^{\circ}$ C保持1分钟,然后逐步降至室温。在估计的退火温度约70 $^{\circ}$ C左右,温度以2 $^{\circ}$ C步幅降低,每一步保持1分钟。

[0265] 有关电子记录和数据分析的额外信息

[0266] 如上所述制备人造平面脂质双分子层。如果没有特别说明,在处理后用10-kHz贝塞尔滤波器以50kHz的采样率采集信号。通过用1-2 $\mu$ L 10%的1,2-二植烷酰-丙三基-3-胆碱磷酸的戊烷溶液预处理在聚四氟乙烯膜(Goodfellow,UK)上的小孔( $\sim$ 100 $\mu$ m)而形成所述脂质双分子层。通过使用浸没在琼脂桥(3%w/v低熔点琼脂糖的2.5MNaCl缓冲液)中的Ag/AgCl电极施加电势。所施加的电势是指连接到装置的反式侧腔的工作电极的电势。将ClyA纳米孔溶液(0.01-0.1ng/mL)添加到与接地电极连接的顺式侧腔中。在单孔插入之后,通过几次灌注循环去除过量的蛋白质。电记录在22 $^{\circ}$ C下0.15-2.5M NaCl,15mM Tris HCl,pH8.0中进行。在0.15M NaCl中,通过使用2kHz低通贝塞尔滤波器并使用10kHz采样率来记录数据。在较高的盐浓度下,数据在50kHz下采样,并且低通贝塞尔滤波器设为10kHz。在采集后用4kHz贝塞尔数字滤波器对0.3和0.5M NaCl下的电流痕迹进行过滤。通过使用Clampfit软件(Molecular devices)的“单通道搜索”功能,使用0.05ms的数据采集阈值,单

独收集电流阻断事件。从开口和阻塞孔电流的全点直方图的高斯拟合分别计算 $I_0$ 和 $I_b$ 值。通过从阻塞孔电流事件( $t_{off}$ )的事件直方图的单指数标准拟合来计算DNA移位停留时间 $\tau_{off}$ 。通过使用阻塞孔孔电流事件之间的事件间时间( $t_{on}$ )的对数直方图的指数对数概率拟合来计算事件间时间 $\tau_{on}$ 。误差表示来自至少三次独立重复的平均值的标准偏差,其数量用“n”表示。

[0267] 从顺式侧腔插入的孔显示在正施加电势下的较高电导,有助于评估插入通道的定向。通过测量电流与施加电压的关系(I-V曲线,电势在21秒内以10mV步幅从-100到+100mV施加)来表征单通道。从+100mV的开孔电流和在-100mV的开孔电流之比( $I_{0+100mV}/I_{0-100mV}$ )得到孔隙整流。通过在给定的施加电势下连续测量开孔电流来评估纳米孔的门控倾向。然后通过所述施加电势下给出 $V_{MAX}$ ,在所述施加电势下,30秒时间范围内没有观察到门控事件(gating events)。在高于 $V_{MAX}$ 的施加电压下观察离子电流的自发可逆门控。在施加电势等于 $V_{MAX}$ 下,通过添加1 $\mu$ M的3'末端生物素化ssDNA1a,随后添加互补的ssDNA 1b(表3),然后添加中性抗生物素蛋白(1.2 $\mu$ M,单体)到顺式侧腔中测试DNA进入和移位穿过孔。

[0268] 有关离子渗透性的额外信息

[0269] 通过测量不对称盐条件(150mMNaCl反式,1M NaCl顺式)下的反转电势来计算ClyA纳米孔的渗透率。将蛋白质纳米孔添加到顺式侧腔,并且首先在对称条件(顺式和反式溶液中均为150mM NaCl,15mM Tris HCl pH 7.5)下表征单通道。在电极平衡后,通过将等分的5M NaCl储备溶液添加到顺式侧腔中,将顺式中的电解质浓度增加到1M。通过添加与加入到顺式侧中溶液等体积的缓冲液来调节反式侧腔的体积,所使用的缓冲液与顺式侧中溶液(150mMNaCl)的缓冲液相同。通过电流-电压(IV)曲线(表6)获得反转电势( $V_r$ ,表3),其是用于获得零电流的电势。通过使用Goldman-Hodgkin-Kats (GHK) 方程从 $V_r$ 计算离子选择性( $P_{Na^+}/P_{Cl^-}$ )。根据GHK方程,观察到的ClyA纳米孔的 $V_r$ 的正值显示了阳离子优先移动穿过孔,表明所述孔是阳离子选择性通道。将顺式侧腔接地,用含有2.5M NaCl的2.5%琼脂糖桥的Ag/AgCl电极进行所有实验。

$$[0270] \quad P_{Na^+}/P_{Cl^-} = \frac{[a_{Cl^-}]_{trans} - [a_{Cl^-}]_{cis} e^{V_F/RT}}{[a_{Na^+}]_{trans} e^{V_F/RT} - [a_{Na^+}]_{cis}}$$

[0271] 其中 $V_r$ 是膜电势,R是通用气体常数(8.314J.K<sup>-1</sup>.mol<sup>-1</sup>),T是开尔文温度,F是法拉第常数(96485C.mol<sup>-1</sup>), $P_x$ 为Na<sup>+</sup>和Cl<sup>-</sup>的膜相对透性, $[a_x]_{cis}$ 是顺式侧腔中的Na<sup>+</sup>和Cl<sup>-</sup>的活性, $[a_x]_{trans}$ 是反式侧腔中的Na<sup>+</sup>和Cl<sup>-</sup>的浓度, $a_x$ 是Na<sup>+</sup>和Cl<sup>-</sup>的活性(J.F.Zemaitis, Handbook of aqueous electrolyte thermodynamics:theory and application,1986; Ludwig Molecular Microbiology 1999;Li-Qun Gu PNAS 2000;Petr G.Merzlyak Biophysics2005)。

[0272] 详细介绍方程(1)和(2)的推导过程。所述方程(1)和(2)分别描述了dsDNA和ssDNA的捕获率。

[0273] dsDNA捕获

[0274] 本方法涉及格罗斯伯格(Grosberg)和拉宾(Rabin)开发的方法。ClyA纳米孔膜被描述为平面介电表面,其厚度为l,具有直径为d的圆柱形孔。ClyA孔的特征距离为 $l=13$ nm和 $d=6$ nm。使用 $\Delta V$ 来表示膜的顺式和反式侧之间的势差,可以看出,顺式侧的电势由下式

给出：

$$[0275] \quad V(r) = \frac{d^2}{8lr} \Delta V \quad (1)$$

[0276] 其在顺式侧远离孔衰减为 $1/r$ /(按惯例,顺式侧电极处的电势被设为零)。坐标原点( $r=0\text{nm}$ )是孔的中部(图19)。

[0277] dsDNA被近似为带电质点,其进行具有扩散常数 $D$ 的扩散运动,和以电泳迁移率 $\mu$ 为特征的电泳漂移。所得到的dsDNA浓度( $c(r, t)$ )的径向坐标下的漂移扩散方程由下式给出：

$$[0278] \quad \frac{\partial c}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left[ Dr^2 \frac{\partial c}{\partial r} - \mu r^2 c \frac{\partial V}{\partial r} \right] \quad (2)$$

[0279] 其中电泳电流前面的负号是因为DNA带负电荷。在本约定中,正迁移系数 $\mu > 0$ 保持不变,因此由于施加的电场引起的漂移速度为 $v = -\mu E$ 。注意爱因斯坦关系不适用于这个系统(即, $D \neq \mu k_B T$ ),因此不能简单地将 $D$ 和 $\mu$ 关联起来。

[0280] 等式(2)的定态解( $\partial c / \partial t = 0$ )是：

$$[0281] \quad c(r) = c_0 \frac{1 - e^{-r^*(1/R - 1/r)}}{1 - e^{-r^*/R}} \quad (3)$$

[0282] 其中边界条件为：无限远处的 $c(R) \rightarrow c_0$ ,和 $c(R) = 0$ ,其中 $R$ 是孔径数量级的微观距离。距离 $r^*$ 定义为：

$$[0283] \quad r^* = \frac{\mu d^2 \Delta V}{8Dl} \quad (4)$$

[0284] 这允许我们将电泳电势(1)改写为：

$$[0285] \quad V(r) = \frac{D}{\mu} \frac{r^*}{r} \quad (5)$$

[0286] 从求解(3)和前面的关系可以得到径向粒子电流密度：

$$[0287] \quad j(r) = -D \frac{\partial c}{\partial r} + \mu c \frac{\partial V}{\partial r} = \frac{Dr^* c_0}{r^2} \frac{1}{1 - e^{-r^*/R}} \quad (6)$$

[0288] 速率是通过对半径为 $r$ 的半球壳上的电流密度进行积分得到的(计算顺式侧可用的表面)：

$$[0289] \quad k_{on} = 2\pi r^2 j(r) = \frac{2\pi Dr^* c_0}{1 - e^{-r^*/R}} \approx 2\pi Dr^* c_0 \quad (7)$$

[0290] 其中近似值 $r^* \gg R$ ,有效性在稍后检查。最终结果与没有外部电势时扩散粒子的Smoluchowski扩散限制的反应速率形式上类似。这里 $r^*$ 可以解释为dsDNA被孔不可逆地捕获的距离。该捕获半径在较高的施加电势下或为了提高电泳迁移率(4)而增加。

[0291] 结合(4)和(7)得到：

$$[0292] \quad k_{on} = \frac{\pi d^2 \Delta V c_0 \mu}{4l} \quad (8)$$

[0293] 请注意,由于 $r^*$ 与 $D$ 成反比例,所以从以前的等式消除 $D$ 。

[0294] 继续进一步估计 $\mu$ 。长度为 $L$ 的dsDNA分子上的总电荷为 $Q = -2\alpha eLa/$ ,其中 $a = 0.34\text{nm}$ ,是两个碱基之间的距离, $\alpha < 1$ 是一个数字系数,反映了并非所有磷酸基团都被电离的事实。将DNA近似为表面积为 $A$ 的圆柱体,阻力估计为 $(\eta A \lambda_D /)$ ,其中 $\eta = 10^{-3} \text{kg m}^{-1} \text{s}^{-2}$ 是水粘度, $\lambda_D$ 是德拜长度。使用定义 $v = -\mu E$ 得到:

$$[0295] \quad \mu = \frac{2\alpha\lambda_D}{\eta\pi ab} \quad (9)$$

[0296] 其中 $b = 2\text{nm}$ 是双螺旋直径。基于 $\zeta$ -电势的计算,格罗斯伯格和拉宾给出了这个等式的替代推导。现在结合等式(7)和(9)并使用与实验相关的数值( $\Delta V = +70\text{mV}$ ,  $c_0 = 1\mu\text{M}$ )并设定 $\alpha = 1$ ,即完全电离,获得:

$$[0297] \quad k_{\text{on}} = 14\lambda_D (\text{s nm } \mu\text{M})^{-1} \quad (10)$$

[0298] 这是上述的等式(1)。

[0299] 最终计算捕获半径 $r^*$ 。为此目的,使用斯托克斯定律(Stokes' law)估计扩散系数:

$$[0300] \quad D = \frac{k_B T}{6\eta\pi R_H} \quad (11)$$

[0301] 其中 $R_H$ 是流体动力学半径。考虑到dsDNA是半径为 $1\text{nm}$ 和长度为 $34\text{nm}$ (100bp)的圆柱体,使用Hansen等人(J.Chem.Phys. (2004) 121:9111-9115)给出的表达式,估计 $R_H \approx 6\text{nm}$ 。结合(11)和(4):

$$[0302] \quad r^* = \frac{3d^2 \lambda_D R_H e \Delta V}{2l ab k_B T} \approx 50 \text{ nm} \quad (12)$$

[0303] 其中 $\lambda_D = 0.5\text{nm}$ 和 $k_B T \approx 25\text{meV}$ 。捕获半径比Debye长度大两个数量级,并且比孔半径大得多,因此在方程(7)中使用的近似值是有道理的。

[0304] ssDNA捕获

[0305] 关于ssDNA捕获的讨论受到Rowghanian等人(Phys.Rev.E (2013) 87:042723)开发的用于能垒限制的过程的方法的启发。这种情况比扩散限制的情况复杂得多,理论不太成熟。本模型是基于使用单个“反应”坐标 $r$ 的漂移扩散方程, $r$ 是一端与孔入口的距离。足够远离孔的ssDNA仅经受如等式(1)描述的有吸引力的电泳力。在孔附近,距离 $\leq R_g$ 处( $R_g$ 是回旋平衡半径),还有熵源的额外排斥力:当末端被迫接近孔入口时,ssDNA卷曲减少其构象熵。如果链足够长,则熵斥力控制静电引力,导致能垒(图19)。

[0306] 用 $U(r)$ 表示熵电势下面的径向电流密度:

$$[0307] \quad j(r) = -D \frac{\partial c}{\partial r} + \mu c \frac{\partial V}{\partial r} - \tilde{\mu} c \frac{\partial U}{\partial r} \quad (13)$$

[0308] 其中 $\mu$ 是电泳迁移率,而 $\tilde{\mu}$ 是与通用非电力相关的迁移率,在这种情况下是熵斥力。当 $\mu$ 不满足爱因斯坦关系( $D \neq \mu k_B T$ )时,通用迁移率 $\tilde{\mu}$ 满足这个关系( $\tilde{\mu} = D/k_B T$ )。公式(13)中的粒子电流可以重写如下:

[0309] 
$$j(r) = -D \left( \frac{\partial c}{\partial r} - \frac{c}{k_B T} \frac{\partial F_b}{\partial r} \right) \quad (14)$$

[0310] 其中：

[0311] 
$$F_b(r) = U(r) - \frac{\mu}{\tilde{\mu}} V(r) \quad (15)$$

[0312] 因此,问题在于电势 $F_b$ 中的粒子的扩散运动。由于违反了爱因斯坦关系,该电势还包含动力学参数,如电泳迁移率 $\mu$ 和来自 $\tilde{\mu} \sim \eta^{-1}$ 的溶剂粘度 $\eta$ 。所述电势具有接近距离 $R_g$ 的最小值和接近孔入口定义的能垒高度( $\Delta F_b \equiv F_b^{\max} - F_b^{\min}$ )的最大值。根据克拉梅斯理论(Kramers' theory),捕获率 $k_{\text{on}}$ 取决于能垒高度:

[0313] 
$$k_{\text{on}} = \omega e^{-\Delta F_b / k_B T} \quad (16)$$

[0314] 通过增加施加的电压 $\Delta V$ 可以降低能垒,从而增强对孔的静电吸引力。式(15)意味着通过增加 $\mu$ ,ssDNA的电泳迁移率可以获得类似的效果。修改 $\mu$ 的一个显而易见的方法是通过改变溶液的离子强度来改变德拜长度。如式(9)所示,电泳迁移率与 $\lambda_D$ 成正比。请注意,盐浓度也影响ssDNA的持续长度,从而影响对能垒的熵贡献 $U(r)$ ,然而这种影响预计是较弱的。预计盐浓度变化对能垒高度的主要影响是包含在 $\lambda_D$ 中的线性项:

[0315] 
$$\Delta F_b = a - b\lambda_D \quad (17)$$

[0316] 用 $a, b > 0$ 与(15)一起解释在实验中观察到的对 $\lambda_D$ 的 $k_{\text{on}}$ 的指数增长。

[0317] 表2从反式侧的孔设计的DNA移位。每个数据点是至少三次实验的平均值,误差是标准偏差。实验在0.15MNaCl,15mM Tris HCl,pH 7.5的溶液中进行。通过将0.01-0.1ng寡聚蛋白添加至反式侧腔来测试纳米孔的活性。负活性表明没有观察到通道插入。 $V_G$ 表示最大施加电压,在该电压下在30秒内未观察到门控事件。DNA捕获表明只有添加与中性抗生物素蛋白复合的生物素化dsDNA时才能观察到瞬时电流阻断。DNA移位表明可以形成dsDNA轮烷。

[0318] 表2

[0319]

孔变体	$I_{\text{O}+100\text{mV}}$ $I_{\text{O}100\text{Mv}}(\text{pA})$	整流比	$V_G, (\text{mV})$	DNA 捕获	DNA 移位
ClyA-RR-E7S	+186±2 -110±2	1.7±0.0	- 70	+	-
ClyA-RR-E11S	+214±27 -124±14	1.7±0.3	- 100	-	-
ClyA-RR-D21S	+193±9 -113±9	1.7±0.2	- 70	-	-
ClyA-RR-D21K	+149±0 -112±0	1.3±0.0	- 50	-	-
ClyA-RR-D32N	+196±5 -104±5	1.9±0.1	-150	-	-

[0320]

ClyA-RR-E7S-D32N	+182±4 -104±6	1.8±0.1	- 70	+	-
ClyA-RR-E7S-D21S	+182±5 -121±3	1.5±0.1	- 70	+	-
ClyA-RR-E129R	无活性	无活性	无活性	无	无
ClyA-RR-1'R	+184±8 -101±3	1.8±0.1	-150	+	-
ClyA-RR-1'R-E7S	+176±5 -109±3	1.6±0.1	- 50	+	-
ClyA-RR-1'R-D21S	+179±9 -108±5	1.7±0.1	- 50	+	-

[0321] 表3:本工作中使用的DNA分子。1是通过将1a与20%过量的1b一起培育并用如所述方法中所述的亲和层析纯化而形成。1\*是通过将1a与20%过量的1b一起培育但没有进一步纯化而形成。两条DNA链中的互补序列以斜体显示。后缀bio表示生物素部分。

[0322] 表3

[0323]

名称	DNA 序列
1a	5'- GGATGA CCT GAT CCA GAT ATT TAT TAT ACA GGT CCA GCG CAC CGT CAG CCC AAT CGC ACT TTT CAC AAA AAG AGA GAG AGATCG ATT ACC /3Bio/-3' (SEQ ID NO: 5)
1b	5'- GGT AAT CGA TCT CTC TCT CTT TTT GTG AAA AGT GCG ATT GGG CTG ACG GTG CGC TGG ACC TGT ATA ATA AAT ATC TGG ATC AGG TCA TCC-3' (SEQ ID NO: 6)
1c	5'- GGT AAT CGATCT CTC TCT CTT TTT GTG AAA AGT GCG ATT GGG CTG ACG GTG CGCTGG AC-/3Bio/-3' (SEQ ID NO: 7)
1d	5'-CTG TAT AAT AAA TAT CTG GAT CAG GTC ATC C /3Bio/-3' (SEQ ID NO: 8)
2a	5'- /5Bio/CCG TAGTTT GGG ATG ACCTGA TCC AGATAT TTATTATAC AGGTCC AGC GCA CCGTCA GCC CAA TCG CACTTT TCA CAA AAA GAG AGA GAG ATC GAT TAC C-3' (SEQ ID NO: 9)
2b	5'- /5Bio/GGT AAT CGATCT CTC TCT CTT TTT GTG AAA AGT GCG ATTGGG CTG ACG GT-3' (SEQ ID NO: 10)

[0324] 表4:选定的ClyA纳米孔的离子选择性。以平均值±标准偏差报道的ClyA变体纳米孔的渗透率( $P_{Na^+}/P_{Cl^-}$ )和反转电势( $V_r$ )。测量每个变体的四个或更多个单通道。使用的缓冲液是:15mM TRIS.HCl pH7.5,顺式侧腔中1M NaCl和反式侧腔中150mMNaCl。

[0325] 表4

[0326]

孔变体	$V_r$ , mV	$P_{Na^+}/P_{Cl^-}$
ClyA-AS	+11.5±0.7	1.92±0.08
ClyA-R	+11.9±1.6	1.97±0.08
ClyA-RR	+11.4±0.9	1.91±0.10

[0327] 表5:ClyA突变体的IV曲线。在22℃下在0.15M NaCl,15mM Tris HCl,pH 7.5中实施电记录。每个数据点是至少三次实验的平均值,误差是标准偏差。

[0328] 表5

[0329]

电压 (mV)	ClyA-AS	ClyA-AS-S110R( ClyA-R)	ClyA-R-D56R	ClyA-R-Q8K	ClyA-R-D64R (ClyA-RR)
-100	-138±6	-128±2	-128±2	-147±18	-111±2
-90	-126±6	-118±1	-119±2	-134±15	-104±2
-80	-115±5	-107±1	-108±2	-120±12	-96.2±1.8
-70	-102±5	-96.2±1.3	-97.3±1.3	-107±10	-87.5±1.4
-60	-89.1±4.3	-84.6±1.1	-85.9±0.9	-93.2±7.9	-78.1±1.3
-50	-75.8±3.6	-72.4±0.9	-73.2±0.3	-78.1±4.9	-67.4±1.1
-40	-61.8±2.9	-59.3±0.9	-60.4±0.4	-63.9±4.2	-56±1
-30	-47.1±2.2	-45.6±0.6	-46.2±0.2	-48.6±2.8	-43.6±0.7
-20	-31.9±1.5	-31.2±0.4	-31.7±0.3	-32.3±2	-30±1
-10	-16.2±0.7	-15.9±0.2	-16.2±0.3	-16.7±0.9	-15.4±0.2
0	0	0	0	0	0
+10	16.8±0.9	16.8±0.1	17.1±0.1	17±2	16.5±0.1
+20	34.2±1.7	34.4±0.3	35.2±0.3	35.3±3.1	34±1
+30	52.1±2.5	52.5±0.6	53.5±0.9	54.4±4.3	52.3±0.3
+40	70.5±3.4	71.6±0.7	72.9±1.1	73.8±5.7	71.5±0.3
+50	89.0±4.5	91.3±0.8	93±2	94.3±6.8	91.8±0.5
+60	108±5	112±1	114±3	115±8	112±1
+70	128±7	132±1	135±4	137±10	13±1
+80	148±8	154±1	157±5	157±12	156±1
+90	168±10	175±2	179±6	181±14	179±1
+100	190±13	198±1	202±8	202±16	202±1
电压 (mV)	ClyA-R E11S	ClyA-R-D122R	ClyA-R-E129R	ClyA-R-D56R-Q 8K	
-100	-165±19	-99.8±2.1	-161±24	-150±15	
-90	-150±17	-93.8±2.1	-145±23	-135±14	
-80	-136±15	-87.3±1.8	-130±20	-123±13	
-70	-120±13	-78.8±2.6	-114±18	-110±11	

[0330]

-60	-105±11	-70.7±2.2	-98.3±15.1	-94.9±10.8	
-50	-88.7±9.3	-62.5±1.3	-81.2±11	-81±9	
-40	-71.9±7.3	-52.3±0.9	-65.4±8.9	-65.3±7.9	

[0331]

-30	-54.6±5.3	-41.1±0.5	-49.2±6.5	-49.9±6.4	
-20	-36.9±3.6	-28.6±0.4	-32.9±4.4	-33±5	
-10	-18.6±1.8	-14.9±0.2	-16.6±2.4	-16.5±2.5	
0	0	0	0	0	
+10	19±2	16±0	16.8±2.5	17.9±2.3	
+20	38.4±3.4	33.3±0.5	33.8±5.1	35.4±5.6	
+30	58±5	51.6±0.5	50.7±7.6	54±9	
+40	77.7±6.9	71.1±0.8	67.6±10.2	72.5±12.5	
+50	97.8±8.2	91.6±0.8	84.6±12.6	91.7±15.9	
+60	119±10	113±1	101±15	114±17	
+70	140±11	135±1	118±18	133±20	
+80	159±13	158±2	136±20	154±23	
+90	181±15	182±2	153±23	182±20	
+100	201±13	207±2	171±26	207±20	

[0332] 表5继续

[0333]

电压 (mV)	ClyA-RR-E7S	ClyA-RR-E11S	ClyA-RR-D21S	ClyA-RR-D21K	ClyA-RR-D32N
-100	-111±4	-128±11	-113±9	-120±1	-108±1
-90	-103±3	-119±10	-106±8	-109±0	-101±1
-80	-95.4±3.4	-109±9	-96.6±7.9	-99.1±0.3	-93.2±0.9
-70	-87±3	-98.9±8.2	-87.6±6.5	-88.1±0.1	-84.7±0.7
-60	-77.3±2.8	-87.9±7.5	-77.9±5.4	-76.6±0.2	-75.5±0.6
-50	-66.3±2.2	-75.5±6.2	-67.1±4.5	-65±0	-65.2±0.5
-40	-54.7±1.9	-62.5±5.2	-55.2±3.8	-53.2±0.1	-54.2±0.5
-30	-42.3±1.4	-48.3±4	-42.3±3.3	-40.5±0.1	-42.1±0.4
-20	-28.9±1	-33.2±2.8	-29±2	-27.5±0	-29±0
-10	-15±0	-17±1	-15.1±0.9	-13.9±0	-15±0
0	0	0	0	0	0
+10	15±1	18±1	16±1	14.3±0	15.6±0.6
+20	31.2±2.4	36.8±3.1	32.5±2.2	29.2±0	32.2±1.3
+30	49.1±2	56.3±4.6	50.1±3.1	44.3±0.1	49.6±2.1
+40	66.2±3.7	76.8±6.6	67.2±5.1	60±0	67.9±2.8
+50	85.7±3.3	98±8	87.3±5.5	75.6±0.5	87±4
+60	105±4	120±11	107±6	92.2±0.1	107±5
+70	125±4	142±13	127±7	109±0	127±5
+80	145±5	165±15	148±8	125±0	149±6
+90	166±5	189±17	170±9	142±0	171±7
+100	188±6	214±19	193±11	160±0	198±4

[0334]

电压(mV)	ClyA-RR-E7S-D21S	ClyA-RR-E7S-D32N	ClyA-RR-1R (ClyA-3R)	ClyA-3R-E7S	ClyA-3R-D21S
-100	-120±4	-104±6	-112±9	-109±3	-108±5
-90	-109±1	-96.9±5.6	-103±12	-101±2	-101±4
-80	-99.9±2.1	-89.2±4.7	-93.5±11.8	-93.2±2.9	-93.1±4
-70	-88.6±1	-80.9±4.2	-86±10	-84.8±1.1	-83.4±2
-60	-80.4±1.7	-71.9±3.5	-77.6±8.2	-74.3±0.9	-73.2±1.3
-50	-67.4±2.7	-62±3	-65.9±7.1	-63.7±1	-63.3±1
-40	-56.2±2	-51.4±2.3	-54.6±5.6	-52.8±1	-52.3±0.8
-30	-43.4±1.1	-39.8±1.7	-42.5±3.6	-40.6±1	-40.6±0.6
-20	-28.9±1.6	-27.4±1.2	-30±3	-28.1±0.5	-28±0
-10	-13.5±1	-14.2±0.5	-14.9±1.7	-14.3±0.3	-14.4±0.2
0	0	0	0	0	0
+10	16.4±2.4	15±0	16.1±0.8	15.1±0.5	15.3±0.3
+20	32.2±1.4	30.7±1	32.4±2.1	30.9±1.1	31.5±1.1
+30	50.3±2.1	47.4±1.5	50±2	47.5±1.5	48.3±1.8
+40	68±2	64.7±1.9	68.7±2.7	64.5±2	65.8±2.4
+50	85.8±1.1	82.7±2.4	88±3	81.7±2.5	83.7±3.3
+60	103±1	101±3	106±5	99.5±2.9	102±4
+70	122±4	121±3	126±5	118±3	121±6
+80	143±0	140±3	149±9	137±4	140±6
+90	163±2	160±3	170±5	156±4	160±7
+100	184±0	182±4	191±11	176±5	179±9

[0335] 表6: 不对称盐浓度下的ClyA变体的IV曲线。测量每个变体的四个或多个单通道。每个数据以平均值±标准偏差报告。使用的缓冲液是15mM TRIS.HCl pH 7.5, 而顺式侧腔含有1M NaCl和反式侧腔150mM。在22℃下在0.15M NaCl, 15mM Tris HCl, pH 7.5中实施电记录。通过应用2kHz低通贝塞尔滤波器并使用100μs (10kHz) 采样率来记录数据。

[0336] 表6

[0337]

IV	ClyA 离子渗透性 开孔电流, pA (平均值±标准偏差)		
电压(mV)	ClyA-AS	ClyA-R	ClyA-RR
20	26.5±0.7	36.7±14.1	32.4±4.3
19	23.1±0.9	32.5±13	28.6±2.7
18	19.8±0.5	27.6±12.8	25.2±2.5
17	16.6±0.9	23.4±10.8	21.3±1.8
16	13.6±1.8	17.9±10.7	16.6±1.2

[0338]

15	9.9±1.9	14.4±9.7	12.8±1.9
14	7.2±1.2	7.8±9.3	10.2±1.2
13	4.4±1.1	2.9±8.9	7.2±1.4
12	2.5±1.4	-1.7±9.4	1.8±1.3
11	-0.7±2.5	-5.6±8.6	-0.8±1.9
10	-3.8±2.7	-12.7±7.3	-4±1
9	-7.3±2.7	-15.6±8	-9.3±3.1
8	-10.7±1.5	-22.3±6.3	-11.9±0.5
7	-13.4±4.3	-24.5±5.9	-15.5±2.6
6	-16.2±0.9	-31.2±5.8	-19.8±3.2
5	-18±2	-35.2±4.5	-23.1±3.1
4	-22.3±2.5	-40.4±5.2	-25.7±1.7
3	-25±2	-43.7±3.5	-30.1±2.9
2	-27.8±3.1	-51.2±4.1	-33.8±4.5
1	-30.5±3.1	-55.6±2.1	-36.9±4.8
0	-34.9±2.7	-60.7±2	-40.7±5
-1	-37.1±3.5	-65.1±2.6	-44.4±3.8
-2	-41±3	-68.8±3.8	-48.2±4.2
-3	-42.1±3.9	-74.8±3.6	-51.3±6.5
-4	-46.2±4.2	-79.8±2.1	-54.8±7.9
-5	-48.8±4.2	-85.2±2.5	-57.8±6.5

[0339]

IV	ClyA 离子渗透性 开孔电流, pA (平均值±标准偏差)		
电压 (mV)	ClyA-AS	ClyA-R	ClyA-RR
-6	-51.5±3.5	-90±3	-61.2±7.8
-7	-55.1±6.3	-94±4	-66.2±7.3
-8	-57.8±5.5	-100±3	-68.6±10.7
-9	-61±4	-103±3	-73.1±8.3
-10	-62.8±4.9	-109±4	-76.5±7.9
-11	-66.2±5	-114±4	-80.1±9.1
-12	-69.7±6	-117±4	-83.6±9.9
-13	-74.7±5.5	-123±4	-86.6±8.7
-14	-74.8±6.1	-129±5	-91.1±11.2
-15	-78.3±5.7	-134±7	-93.5±10.7
-16	-80.2±6.2	-137±8	-96.7±9.7
-17	-84.2±6.4	-144±7	-100±13
-18	-87.6±7.6	-148±8	-104±12
-19	-90.4±7.7	-153±8	-108±12

[0340]

-20	-92.4±7.3	-158±8	-112±11
-----	-----------	--------	---------

[0341] 表7:ssDNA (1a) 和dsDNA (1) 移位穿过ClyA-RR纳米孔。每个条件测量三个或更多个单通道。数据以平均值±标准偏差报告。在22℃下15mM Tris-HCl, pH 7.5中实施电记录。通过应用10kHz低通贝塞尔滤波器并使用20μs (50kHz) 采样率来记录数据。

[0342] 表7

[0343]

ssDNA (1a)			
[NaCl] (M)	I <sub>RES</sub>	$\tau_{off}$ (ms)	$\tau_{on}$ (ms)
0.15	0.92±0.00	0.54±0.28	8.5±1.1
0.3	0.89±0.01	0.18±0.04	44±1
0.5	0.88±0.02	0.12±0.02	112±14
1	0.82±0.01	0.13±0.01	232±36
2	0.84±0.01	0.12±0.02	393±17
2.5	0.78±0.01	0.18±0.02	500±50

[0344]

dsDNA (1)			
[NaCl] (M)	I <sub>RES</sub>	$\tau_{off}$ (ms)	$\tau_{on}$ (ms)
0.15	0.92±0.00	0.29±0.07	40±13
0.6	0.83±0.03	0.26±0.09	162±31
1	0.76±0.01	0.26±0.09	214±18
2	0.75±0.04	0.33±0.07	532±52
2.5	0.75±0.01	0.60±0.48	641±37

[0345] 序列表:

[0346]

介绍	序列
伤寒沙门菌 ( <i>S. typhi</i> ) ClyA(ClyA-WT)的 蛋白质序列 SEQ ID NO: 1	MTGIFAEQTVEVVKSAIETADGALDLYNKYLDQVIPWKTFDETIKELSRFKQE YSQEASVLVGDIKVLLMDSQDKYFEATQTVYEWCGVVTQLLSAYIQLFDEYNE KKASAQKDILIRILDDGVKKLNEAQKSLTSSQSFNNASGKLLALDSQLTNDP SEKSSYFQSQVDRIKEAYAGAAAGIVAGPFGLIISYSIAAGVIEGKLIPELN NRLKTVQNFFTSLSATVKQANKDIDAAKLKLATEIAAIGEIKTETETTRFYVD YDDLMLSLKGAACKMINTCNEYQQRHGKKTLEFVDPV
ClyA-AS 的蛋白质 序列 SEQ ID NO: 2	MTGIFAEQTVEVVKSAIETADGALDLYNKYLDQVIPWKTFDETIKELSRFKQE YSQEASVLVGDIKVLLMDSQDKYFEATQTVYEWAGVVTQLLSAYIQLFDGYNE KKASAQKDILIRILDDGVKKLNEAQKSLTSSQSFNNASGKLLALDSQLTNDP SEKSSYFQSQVDRIKEAYAGAAAGIVAGPFGLIISYSIAAGVIEGKLIPELN NRLKTVQNFFTSLSATVKQANKDIDAAKLKLATEIAAIGEIKTETETTRFYVD YDDLMLSLKGAACKMINTSNEYQQRHGKKTLEFVDPVGVSSYHHHHH*
伤寒沙门菌 ClyA (ClyA-WT)的核苷 酸序列 SEQ ID NO: 3	CCTGCGTAGATAAGCAGGAAGCAGGCAGTATTTCCAGCTTCTGGAATGTTAAA GCTACAAAAGTTGTCTGGAGGTAATAGGTAAGAATACTTTATAAAACAGGTAC TTAATTGCAATTTATATATTTAAAGAGGCAAATGATTATGACCGGAATATTTG CAGAACAACTGTAGAGGTAGTTAAAGCGCGATCGAAACCGCAGATGGGGCA TTAGATCTTTATAACAAATACCTCGACCAGGTCATCCCCTGGAAGACCTTTGA TGAAACCATAAAAGAGTTAAGCCGTTTTAAACAGGAGTACTCGCAGGAAGCTT CTGTTTTAGTTGGTGATATTAAAGTTTTGCTTATGGACAGCCAGGACAAGTAT TTTGAAGCGACACAACTGTTTATGAATGGTGTGGTGTGCTGACGCAATTACT

	<p>CTCAGCGTATATTTTACTATTTGATGAATATAATGAGAAAAAGCATCAGCCC AGAAAGACATTCTCATTAGGATATTAGATGATGGTGTCAAGAACTGAATGAA GCGCAAAATCTCTCCTGACAAGTTCACAAAGTTTCAACAACGCTTCCGGAAA ACTGCTGGCATTAGATAGCCAGTTAACTAATGATTTTTCGGAAAAAGTAGTT ATTTCCAGTCACAGGTGGATAGAATTCGTAAGGAAGCTTATGCCGTTGCTGCA GCCGGCATAGTCGCCGGTCCGTTTGGATTAATTATTTCTATTCTATTGCTGC GGGCGTGATTGAAGGGAAATTGATTCCAGAATTGAATAACAGGCTAAAAACAG TGCAAAATTTCTTTACTAGCTTATCAGCTACAGTGAAACAAGCGAATAAAGAT ATCGATGCGGCAAAATTGAAATTAGCCACTGAAATAGCAGCAATTGGGGAGAT AAAAACGGAAACCGAAACAACAGATTCTACGTTGATTATGATGATTTAATGC TTTCTTTATTAAAAGGAGCTGCAAAGAAAATGATTAACACCTGTAATGAATAC CAACAAAGACACGGTAAGAAGACGCTTTTCGAGGTTCTTGACGCTCTGATACAT TTTCATTGATCTGTGTACTTTTAACGCCCGATAGCGTAAAGAAAATGAGAGA CGGAGAAAAAGCGATATTCAACAGCCCGATAAACAGAGTCGTTACCGGGCTG ACGAGGTTATCAGGCGTTAAGCTGGTAG</p>
[0347]	<p>ClyA-AS 的核苷酸 序列</p> <p>SEQ ID NO: 4</p> <p>ATGACGGGTATCTTTGCGGAACAGACGGTGGAAGTTGTGAAAAGTGCGATTGA AACGGCTGACGGTGCGCTGGACCTGTATAATAAATATCTGGATCAGGTCATCC CGTGGAACCTTTGACGAAACGATTAAAGAACTGAGCCGTTTCAAACAGGAA TACAGTCAAGAAGCGTCCGTCCTAGTGGGCGATATCAAAGTGCTGCTGATGGA TTCTCAGGACAAATATTTTGAAGCTACCCAAACGGTTTACGAATGGGCGGGTG TGGTTACCCAGCTGCTGTCCGCATATATTCAGCTGTTCGATGGATACAATGAG AAAAAAGCGAGCGCGCAGAAAGACATTCTGATCCGCATTCTGGATGACGGCGT GAAAAAAGTGAATGAAGCCAGAAATCGCTGCTGACCAGCTCTCAATCATTTA ACAATGCCTCGGGTAACTGCTGGCACTGGATAGCCAGCTGACGAACGACTTT TCTGAAAAAGTTCCTATTACCAGAGCCAAGTCGATCGTATTCGTAAAGAAGC CTACGCAGGTGCCGCAGCAGGTATTGTGGCCGGTCCGTTCCGTTGATTATCT CATATTCAATTGCTGCGGGCGTTGTGCAAGGTAACTGATTCCGGAAGTGAAC AATCGTCTGAAAACCGTTTCAAGAACTTTTTCACCACTGCTGCTGCTACGGTCAA ACAAGCGAATAAAGATATCGACGCCGCAAACTGAACTGGCCACGGAAATCG CTGCGATTGGCGAAATCAAACCGAAACGGAACCGCGCTTTTATGTTGAT TACGATGACCTGATGCTGAGCCTGCTGAAAGGTGCCGCAAGAAATGATTAA TACCTCTAATGAATATCAGCAGCGTCACGGTAGAAAAACCTGTGTTGAAGTCC CGGATGTGGGAGCAGCTACCACCATCATCACCCTAAAAAGCTT</p>

[0348] 其他实施方式

[0349] 本说明书中公开的所有特征可以以任何组合结合。本说明书中公开的每个特征可以由相同，等同或类似目的的选择性特征代替。因此，除非另有明确说明，否则所公开的每个特征仅仅是一系列等同或类似特征的一个例子。

[0350] 根据以上描述，本领域技术人员可以容易地确定本发明的基本特征，并且在不脱离其精神和范围的情况下，可以对本发明进行各种改变和修饰以使其适应各种用途和条件。因此，其他实施方式也在权利要求内。

[0351] 等同物

[0352] 尽管本文已经描述和说明了若干发明实施方式，但本领域普通技术人员将容易想到其他手段和/或结构，用于执行本文描述的功能和/或获得结果和/或一个或多个优点，并且每个这样的变化和/或修饰被认为是在发明实施方式的范围内。更一般地，本领域技术人员将容易理解，本文所述的所有参数、尺寸、材料和配置都是示例性的，并且实际参数、尺寸、材料和/或配置将取决于使用本发明的特定应用或教导的应用。本领域的技术人员将认识到，或仅仅使用常规实验就能够确定本文所述的具体发明实施方式的许多等同物。因此，应该理解的是，前述实施方式仅以示例的方式呈现，并且在所附权利要求及其等同物的范围内，可以以与具体描述和要求保护不同的方式来实践发明实施方式。本发明创造性的实施方式涉及本文所述的每个单独的特征、系统、物品、材料，试剂盒和/或方法。此外，如果这

些特征、系统、物品、材料、试剂盒和/或方法不是相互不一致,则两个或更多个这样的特征、系统、物品、材料、试剂盒和/或方法的任何组合都包括在本发明的保护范围内。

[0353] 如本文所定义和使用的定义应理解为由字典定义、通过引用并入的文献中的定义和/或所定义的术语的普通含义控制。

[0354] 本文公开的所有参考文献,专利和专利申请通过引用每个被引用的主题并入,在一些情况下可以涵盖整个文档。

[0355] 除非有明确的相反指示,否则本文在说明书和权利要求书中使用的不定冠词“一”和“一个”应理解为意指“至少一个”。

[0356] 在本文说明书和权利要求书中使用的短语“和/或”应理解为意指这样结合的元素的一个或两个,即在一些情况下元素结合地存在和在其他情况下元素分离地存在。用“和/或”列出的多个元素应该以相同的方式解释,即如此结合的元素中的一个或多个。除了由“和/或”句子明确限定的元素之外,其他元素可以可选地存在,不管与具体限定的那些元素相关还是不相关。因此,作为非限制性示例,对“A和/或B”的引用,当与如“包括”的开放式语言结合使用时,在一个实施例中可以仅指A(可选地包括除B之外的元素);在另一个实施例中,仅指B(可选地包括除A之外的元素);在又一个实施例中,指A和B(可选地包括其他元素)等等。

[0357] 如在本说明书和权利要求书所使用的,“或”应被理解为具有与如上所定义的“和/或”相同的含义。例如,当分离列表中的项目时,“或”或“和/或”应被解释为是包含的,即包含(多个元素或元素列表并且可选地其他未列出项目中的)至少一个元素,但也包括不只一个元素。只有术语明确表示相反,那么如“仅一个”或“恰好一个”,或者当在权利要求中使用时的“由...组成”将指包含多个元素或元素列表中的恰好一个元素。一般而言,当之前有排他性术语(如“或者”“一个”“只有一个”或“恰好一个”)时,在本文中使用的术语“或”应当仅被解释为指示排他性替代(例如“一个或另一个,但不是两个”);当“基本上由.....组成”在权利要求中使用时,应具有在专利法领域中使用的普通含义。

[0358] 如本文的说明书和权利要求书所使用的,关于一个或多个元素的列表的短语“至少一个”应该理解为意指选自该元素列表中的任何一个或多个元素的至少一个元素,但不一定包括元素列表内具体列出的每个元素中的至少一个,并且不排除元素列表中的元素的任何组合。该定义还允许,除了短语“至少一个”所指的元素列表内的具体指定的元素以外,元素可以可选地存在,不管与具体指定的那些元素相关还是不相关。因此,作为非限制性示例,“A和B中的至少一个”(或者等同地,“A或B中的至少一个”,或者等同地“A和/或B中的至少一个”)可以指在一个实施例中的至少一个,可选地包括多于一个,A,不存在B(并且可选地包括除了B之外的元素);在另一个实施例中,指至少一个,可选地包括多于一个,B,不存在A(并且可选地包括除了A之外的元素);在又一个实施例中,指至少一个,任选地包括多于一个,A和至少一个,可选地包括多于一个,B(和可选地包括其他元素)等等。

[0359] 还应该理解,除非清楚相反指示,否则在本文所限定的任何包括多于一个步骤或行为的方法中,所述方法的步骤或行动的顺序不一定限于所述方法记载的步骤或行动的顺序。

## 序列表

&lt;110&gt;鲁汶大学研究与发展中心

&lt;120&gt;修饰的纳米孔，包含其的组合物及其应用

&lt;130&gt; O0366.70046WO00

&lt;140&gt;还未分配

&lt;141&gt; 2016-12-08

&lt;150&gt; US 62/264,709

&lt;151&gt; 2015-12-08

&lt;160&gt; 10

&lt;170&gt; PatentIn version 3.5

&lt;210&gt; 1

&lt;211&gt; 303

&lt;212&gt; PRT

&lt;213&gt;伤寒沙门菌 (S. typhi)

&lt;400&gt; 1

Met Thr Gly Ile Phe Ala Glu Gln Thr Val Glu Val Val Lys Ser Ala  
1 5 10 15

[0001]

Ile Glu Thr Ala Asp Gly Ala Leu Asp Leu Tyr Asn Lys Tyr Leu Asp  
20 25 30

Gln Val Ile Pro Trp Lys Thr Phe Asp Glu Thr Ile Lys Glu Leu Ser  
35 40 45

Arg Phe Lys Gln Glu Tyr Ser Gln Glu Ala Ser Val Leu Val Gly Asp  
50 55 60

Ile Lys Val Leu Leu Met Asp Ser Gln Asp Lys Tyr Phe Glu Ala Thr  
65 70 75 80

Gln Thr Val Tyr Glu Trp Cys Gly Val Val Thr Gln Leu Leu Ser Ala  
85 90 95

Tyr Ile Leu Leu Phe Asp Glu Tyr Asn Glu Lys Lys Ala Ser Ala Gln  
100 105 110

Lys Asp Ile Leu Ile Arg Ile Leu Asp Asp Gly Val Lys Lys Leu Asn  
115 120 125

Glu Ala Gln Lys Ser Leu Leu Thr Ser Ser Gln Ser Phe Asn Asn Ala

	130	135	140	
	Ser Gly Lys Leu Leu Ala Leu Asp Ser Gln Leu Thr Asn Asp Phe Ser			
	145	150	155	160
	Glu Lys Ser Ser Tyr Phe Gln Ser Gln Val Asp Arg Ile Arg Lys Glu			
		165	170	175
	Ala Tyr Ala Gly Ala Ala Ala Gly Ile Val Ala Gly Pro Phe Gly Leu			
		180	185	190
	Ile Ile Ser Tyr Ser Ile Ala Ala Gly Val Ile Glu Gly Lys Leu Ile			
		195	200	205
	Pro Glu Leu Asn Asn Arg Leu Lys Thr Val Gln Asn Phe Phe Thr Ser			
		210	215	220
	Leu Ser Ala Thr Val Lys Gln Ala Asn Lys Asp Ile Asp Ala Ala Lys			
	225	230	235	240
[0002]	Leu Lys Leu Ala Thr Glu Ile Ala Ala Ile Gly Glu Ile Lys Thr Glu			
		245	250	255
	Thr Glu Thr Thr Arg Phe Tyr Val Asp Tyr Asp Asp Leu Met Leu Ser			
		260	265	270
	Leu Leu Lys Gly Ala Ala Lys Lys Met Ile Asn Thr Cys Asn Glu Tyr			
		275	280	285
	Gln Gln Arg His Gly Lys Lys Thr Leu Phe Glu Val Pro Asp Val			
		290	295	300
	<210> 2			
	<211> 312			
	<212> PRT			
	<213>伤寒沙门菌 (S. typhi)			
	<400> 2			
	Met Thr Gly Ile Phe Ala Glu Gln Thr Val Glu Val Val Lys Ser Ala			
	1	5	10	15
	Ile Glu Thr Ala Asp Gly Ala Leu Asp Leu Tyr Asn Lys Tyr Leu Asp			
		20	25	30

	Gln Val Ile Pro Trp Lys Thr Phe Asp Glu Thr Ile Lys Glu Leu Ser	
	35 40 45	
	Arg Phe Lys Gln Glu Tyr Ser Gln Glu Ala Ser Val Leu Val Gly Asp	
	50 55 60	
	Ile Lys Val Leu Leu Met Asp Ser Gln Asp Lys Tyr Phe Glu Ala Thr	
	65 70 75 80	
	Gln Thr Val Tyr Glu Trp Ala Gly Val Val Thr Gln Leu Leu Ser Ala	
	85 90 95	
	Tyr Ile Gln Leu Phe Asp Gly Tyr Asn Glu Lys Lys Ala Ser Ala Gln	
	100 105 110	
	Lys Asp Ile Leu Ile Arg Ile Leu Asp Asp Gly Val Lys Lys Leu Asn	
	115 120 125	
	Glu Ala Gln Lys Ser Leu Leu Thr Ser Ser Gln Ser Phe Asn Asn Ala	
	130 135 140	
[0003]	Ser Gly Lys Leu Leu Ala Leu Asp Ser Gln Leu Thr Asn Asp Phe Ser	
	145 150 155 160	
	Glu Lys Ser Ser Tyr Tyr Gln Ser Gln Val Asp Arg Ile Arg Lys Glu	
	165 170 175	
	Ala Tyr Ala Gly Ala Ala Ala Gly Ile Val Ala Gly Pro Phe Gly Leu	
	180 185 190	
	Ile Ile Ser Tyr Ser Ile Ala Ala Gly Val Val Glu Gly Lys Leu Ile	
	195 200 205	
	Pro Glu Leu Asn Asn Arg Leu Lys Thr Val Gln Asn Phe Phe Thr Ser	
	210 215 220	
	Leu Ser Ala Thr Val Lys Gln Ala Asn Lys Asp Ile Asp Ala Ala Lys	
	225 230 235 240	
	Leu Lys Leu Ala Thr Glu Ile Ala Ala Ile Gly Glu Ile Lys Thr Glu	
	245 250 255	
	Thr Glu Thr Thr Arg Phe Tyr Val Asp Tyr Asp Asp Leu Met Leu Ser	
	260 265 270	

Leu Leu Lys Gly Ala Ala Lys Lys Met Ile Asn Thr Ser Asn Glu Tyr  
275 280 285

Gln Gln Arg His Gly Arg Lys Thr Leu Phe Glu Val Pro Asp Val Gly  
290 295 300

Ser Ser Tyr His His His His His  
305 310

<210> 3

<211> 1194

<212> DNA

<213> 伤寒沙门菌 (S. typhi)

<400> 3

[0004]

```
cctgcgtaga taagcaggaa gcaggcagta ttccagctt ctggaatgt aaagctacaa      60
aagttgtctg gaggtaatag gtaagaatac ttataaaac aggtacttaa ttgcaattta      120
tataattaaa gaggcaaatg attatgaccg gaatatttgc agaacaaact gtagaggtag      180
ttaaagcgc gatcgaaacc gcagatgggg cattagatct ttatacaaa tacctcgacc      240
aggtcatccc ctggaagacc ttgatgaaa ccataaaaga gttaagccgt tttaacagg      300
agtactcgca ggaagcttct gtttagttg gtgatattaa agtttgcct atggacagcc      360
aggacaagta tttgaagcg acacaaactg ttatgaatg gtgtggtgtc gtgacgcaat      420
tactctcagc gtatatttta ctatttgatg aatataatga gaaaaagca tcagcccaga      480
aagacattct cattaggata ttagatgatg gtgtcaagaa actgaatgaa gcgcaaaaat      540
ctctcctgac aagtcacaa agttcaaca acgcttccgg aaaactgctg gcattagata      600
gccagttaac taatgatttt tcggaaaaaa gtagttattt ccagtcacag gtggatagaa      660
ttcgtaaagg agcttatgcc ggtgctgcag ccggcatagt cgccgggtccg ttggattaa      720
ttatttccta ttctattgct gcgggcgtga ttgaaggga attgattcca gaattgaata      780
acaggctaaa aacagtgcaa aatttcttta ctagcttacc agctacagt aaacaagcga      840
ataaagatat cgatgcggca aaattgaaat tagccactga aatagcagca attggggaga      900
taaaaacgga aaccgaaaca accagattct acgttgatta tgatgattta atgctttctt      960
tattaaaagg agctgcaaag aaatgatta acacctgtaa tgaataccaa caaagacacg     1020
gtaagaagac gcttttcgag gttcctgacg tctgatacat ttcatcga tctgtgtact     1080
ttaacgccc gatagcgtaa agaaaatgag agacggagaa aaagcgatat tcaacagccc     1140
```

	gataaacaag agtcgttacc gggctgacga gggtatcagg cgtaaagctg gtag	1194
	<210> 4	
	<211> 945	
	<212> DNA	
	<213>伤寒沙门菌 (S. typhi)	
	<400> 4	
	atgacgggta tctttgcgga acagacgggtg gaagttgtga aaagtgcgat tgaacggct	60
	gacggtgctg tggacctgta taataaatat ctggatcagg tcatcccggtg gaaaaccttt	120
	gacgaaacga ttaaagaact gagccgttgc aaacaggaat acagtcaaga agcgtccgtc	180
	ctagtgggctg atatcaaagt gctgctgatg gattctcagg acaaatattt tgaagctacc	240
	caaacgggtt acgaatgggc ggggtgtggt acccagctgc tgcgcgata tattcagctg	300
	ttcgatggat acaatgagaa aaaagcgagc gcgcagaaag acattctgat ccgcattctg	360
	gatgacggcg tgaaaaaact gaatgaagcc cagaaatcgc tctgaccag ctctcaatca	420
	ttaacaatg cctcgggtaa actgctggca ctggatagcc agctgacgaa cgacttttct	480
	gaaaaaagt cctattacca gagccaagtc gatcgtatc gtaaagaagc ctacgcaggt	540
	gccgcagcag gtattgtggc cgggtccgtc ggtctgatta tctcatattc aattgctgcg	600
[0005]	ggcgttctcg aaggtaaact gattccggaa ctgaacaatc gtctgaaaac cggtcagaac	660
	ttttcacca gtctgtctgc tacggtcaaa caagcgaata aagatatga cgcgcgaaaa	720
	ctgaaactgg ccacggaaat cgctgcgatt ggcgaaatca aaaccgaaac ggaaccacg	780
	cgcttttatg ttgattacga tgacctgatg ctgagcctgc tgaagggtgc cgcaagaaa	840
	atgattaata cctctaataa atatcagcag cgtcacggta gaaaaacct gtttgaagtc	900
	ccggatgtgg gcagcagcta ccaccatcat caccactaaa agctt	945
	<210> 5	
	<211> 90	
	<212> DNA	
	<213>人工序列 (Artificial Sequence)	
	<220>	
	<223>合成多核苷酸 (Synthetic Polynucleotide)	
	<220>	
	<221>misc_binding	
	<222> (90)..(90)	
	<223>用 3' /3Bio/修饰	
	<400> 5	
	ggatgacctg atccagatat ttattatata ggtccagcgc accgtcagcc caatgcact	60

	ttcacaaaa agagagagag atcgattacc	90
	<210> 6 <211> 90 <212> DNA <213>人工序列 (Artificial Sequence)	
	<220> <223>合成多核苷酸	
	<400> 6 ggtaatcgat ctctctctct tttgtgaaa agtgcgattg ggctgacggt gcgctggacc	60
	tgtataataa atatctggat caggtcatcc	90
	<210> 7 <211> 59 <212> DNA <213>人工序列 (Artificial Sequence)	
	<220> <223>合成多核苷酸	
[0006]	<220> <221>misc_binding <222> (59)..(59) <223>用 3' /3Bio/修饰	
	<400> 7 ggtaatcgat ctctctctct tttgtgaaa agtgcgattg ggctgacggt gcgctggac	59
	<210> 8 <211> 31 <212> DNA <213>人工序列 (Artificial Sequence)	
	<220> <223>合成多核苷酸	
	<220> <221>misc_binding <222> (31)..(31) <223>用 3' /3Bio/修饰	
	<400> 8 ctgtataata aatatctgga tcaggtcatc c	31
	<210> 9 <211> 100 <212> DNA	

<213>人工序列 (Artificial Sequence)

<220>

<223>合成多核苷酸

<220>

<221>misc\_binding

<222> (1)..(1)

<223>用 5' /5Bio/修饰

<400> 9

ccgtagtgtg ggatgacctg atccagatat ttattataca ggtccagcgc accgtcagcc 60

caatcgcaact ttccacaaaa agagagagag atcgattacc 100

[0007]

<210> 10

<211> 50

<212> DNA

<213>人工序列 (Artificial Sequence)

<220>

<223>合成多核苷酸

<220>

<221>misc\_binding

<222> (1)..(1)

<223>用 5' /5Bio/修饰

<400> 10

ggtaatcgat ctctctctct tttgtgaaa agtgcgattg ggctgacggt 50

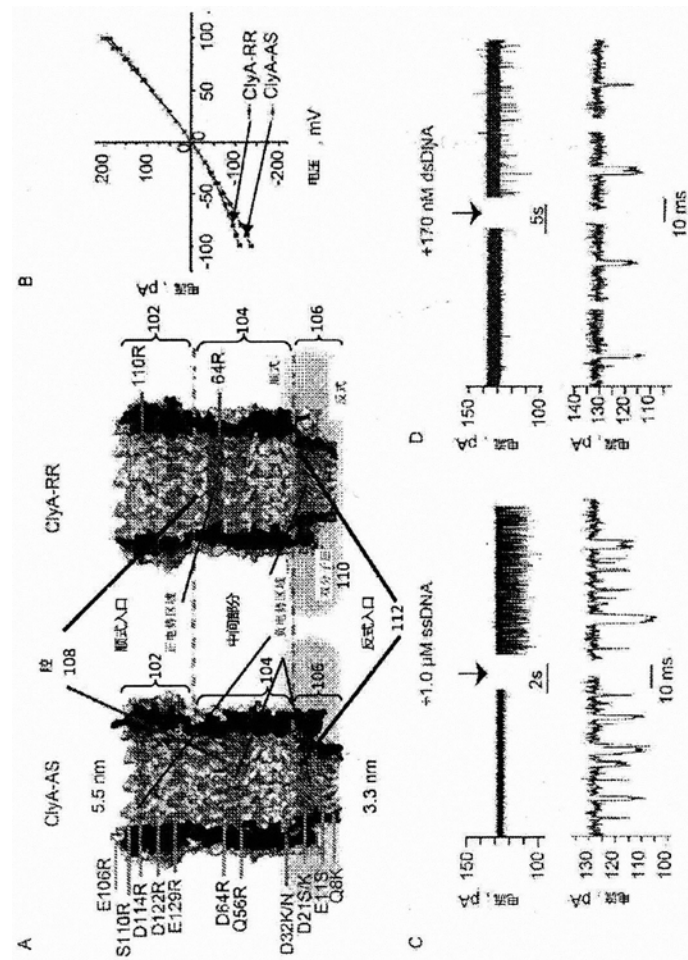


图1

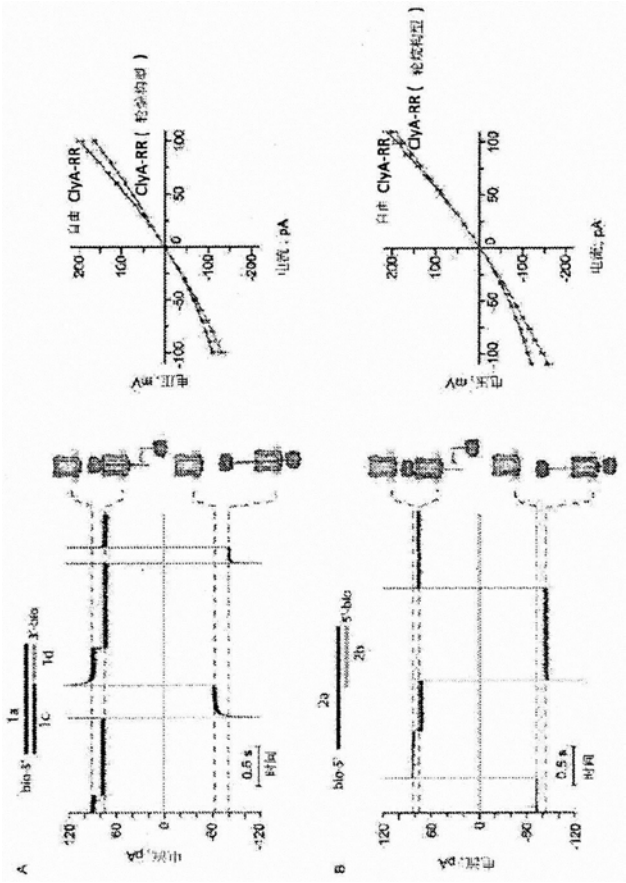


图2

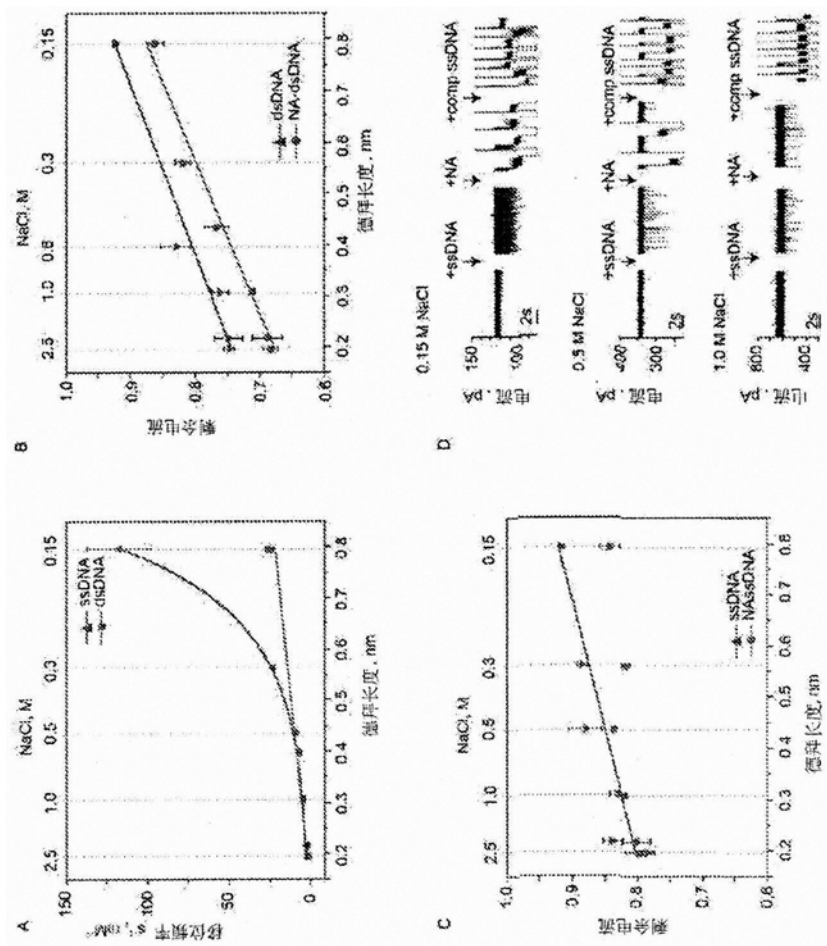


图3

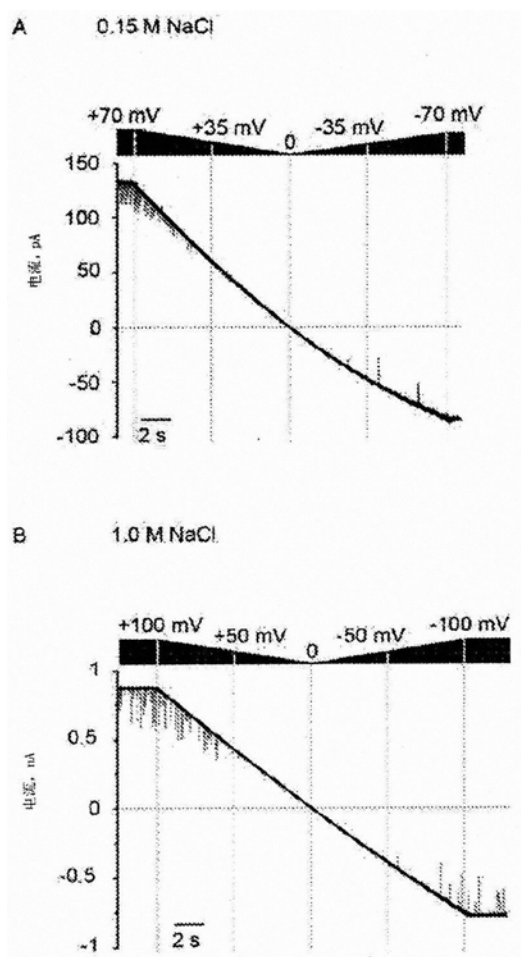


图4

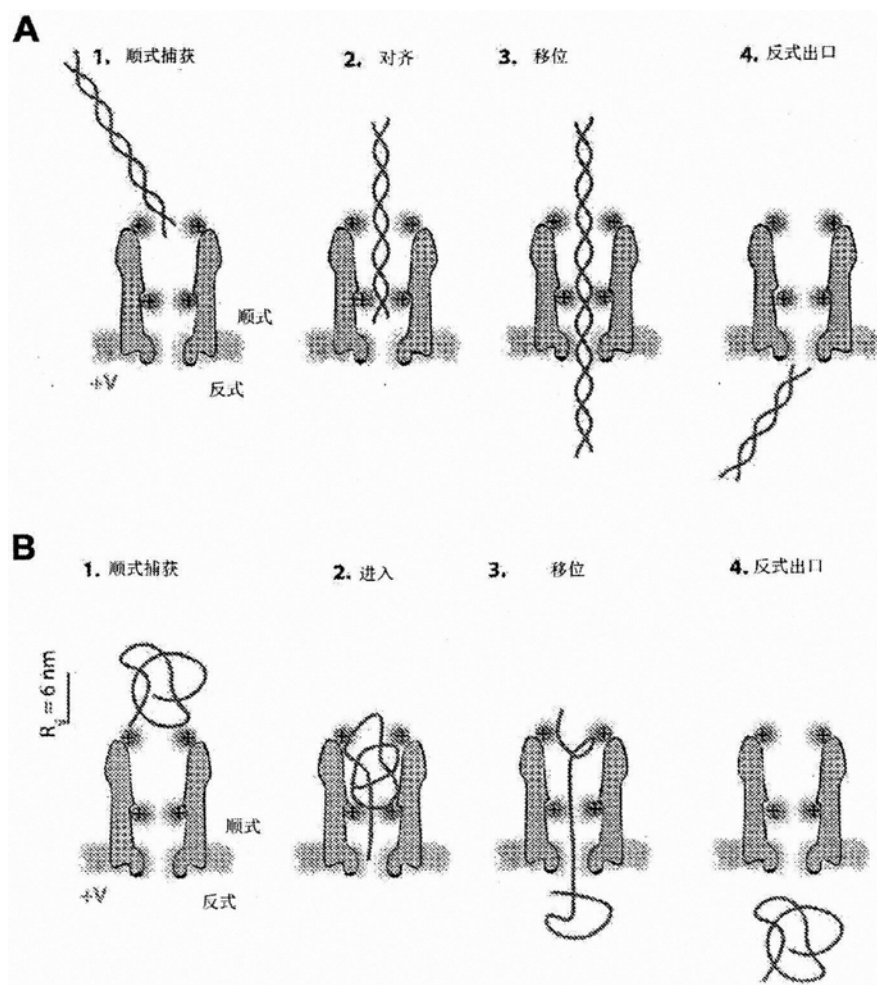


图5

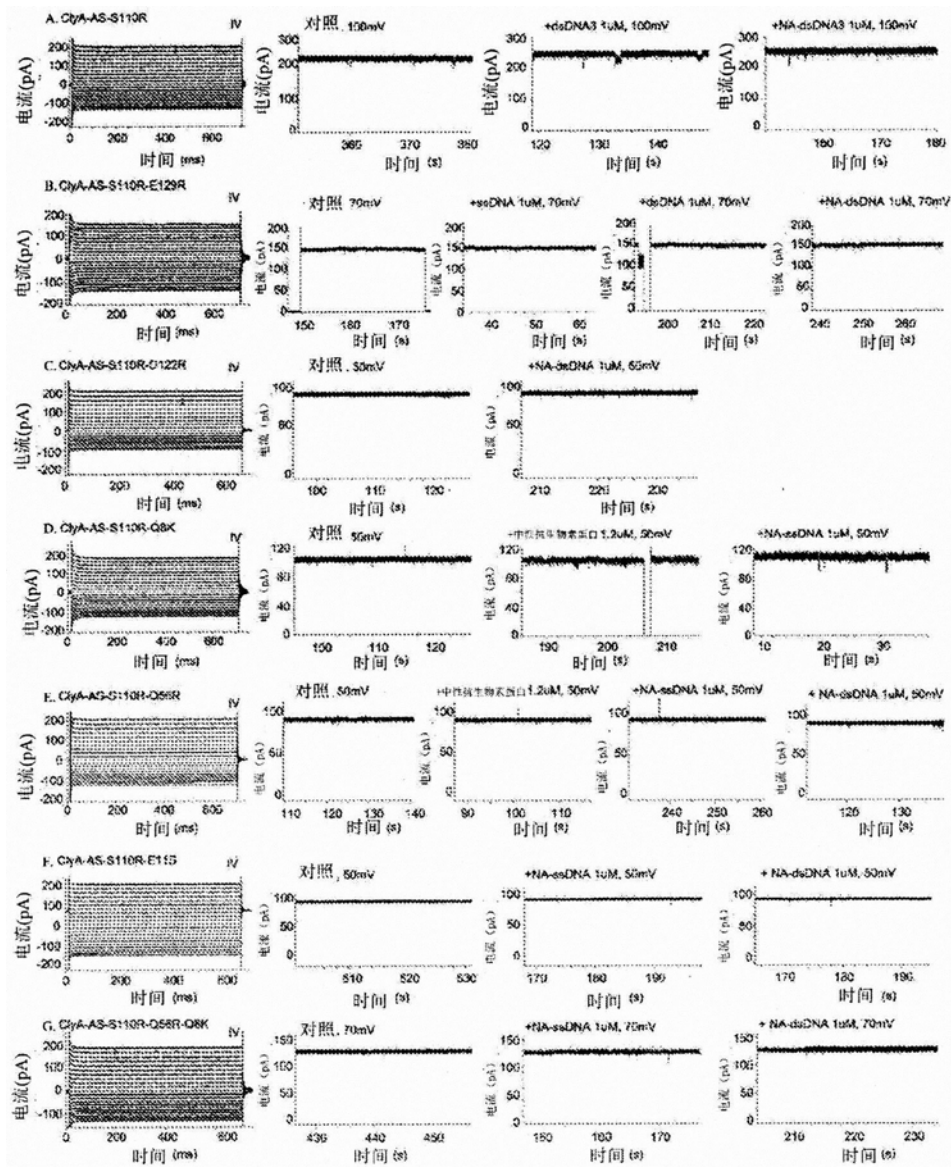


图6

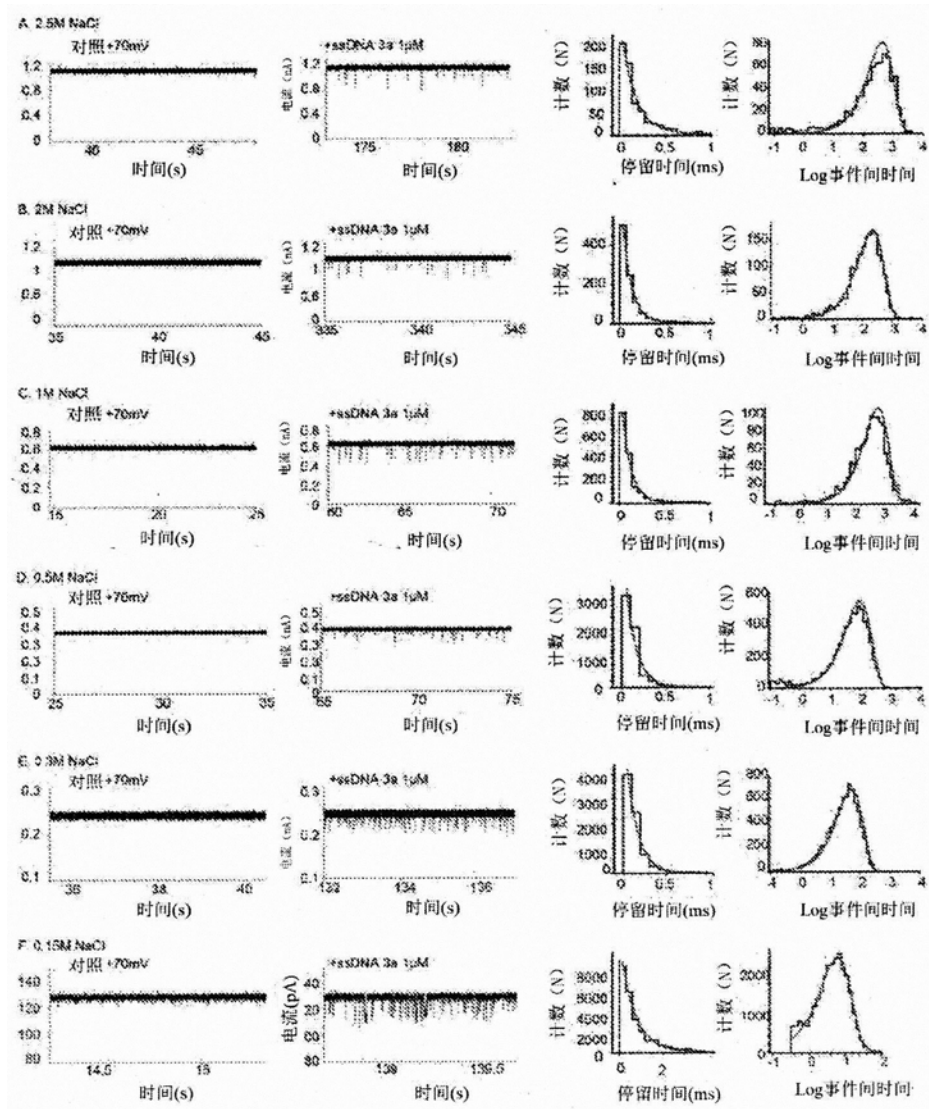


图7

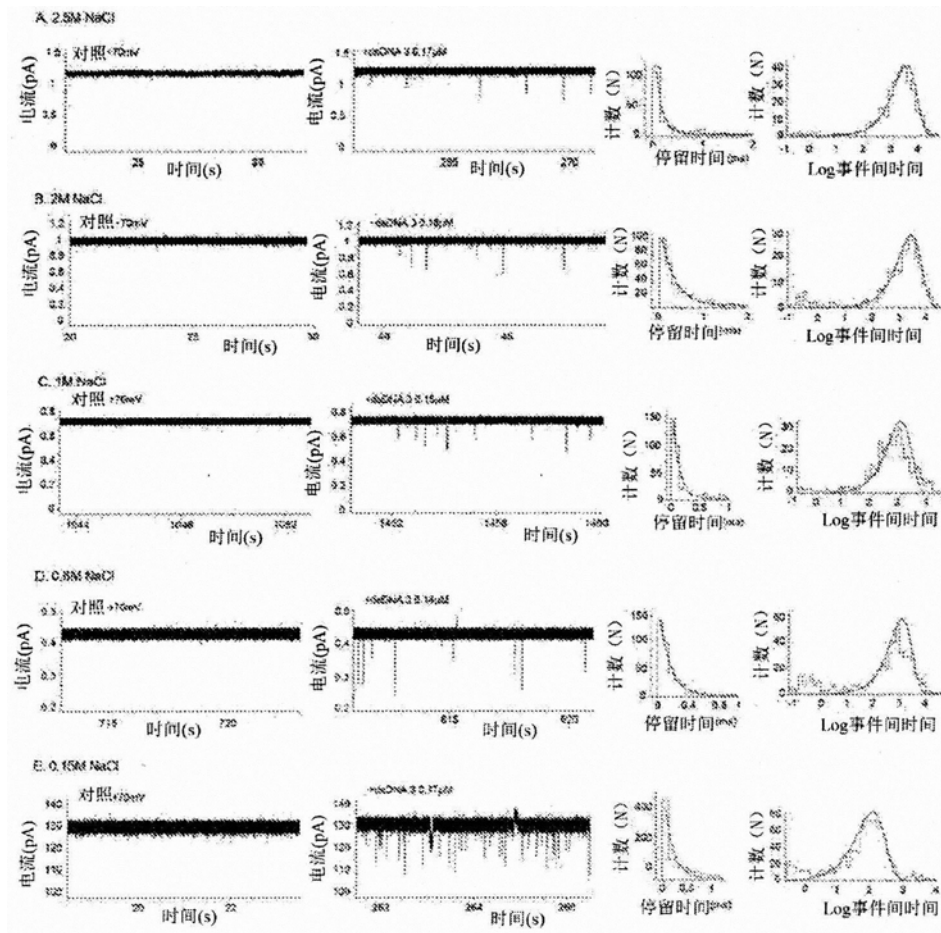


图8

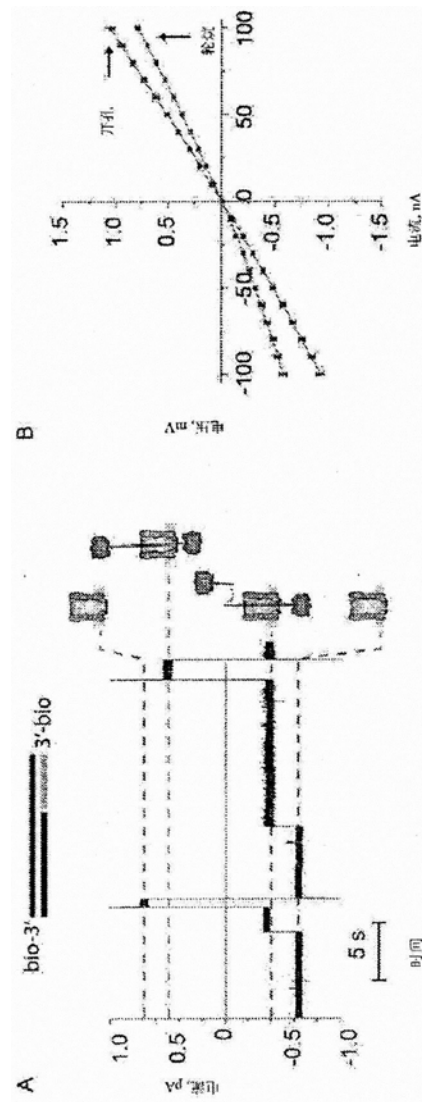


图9

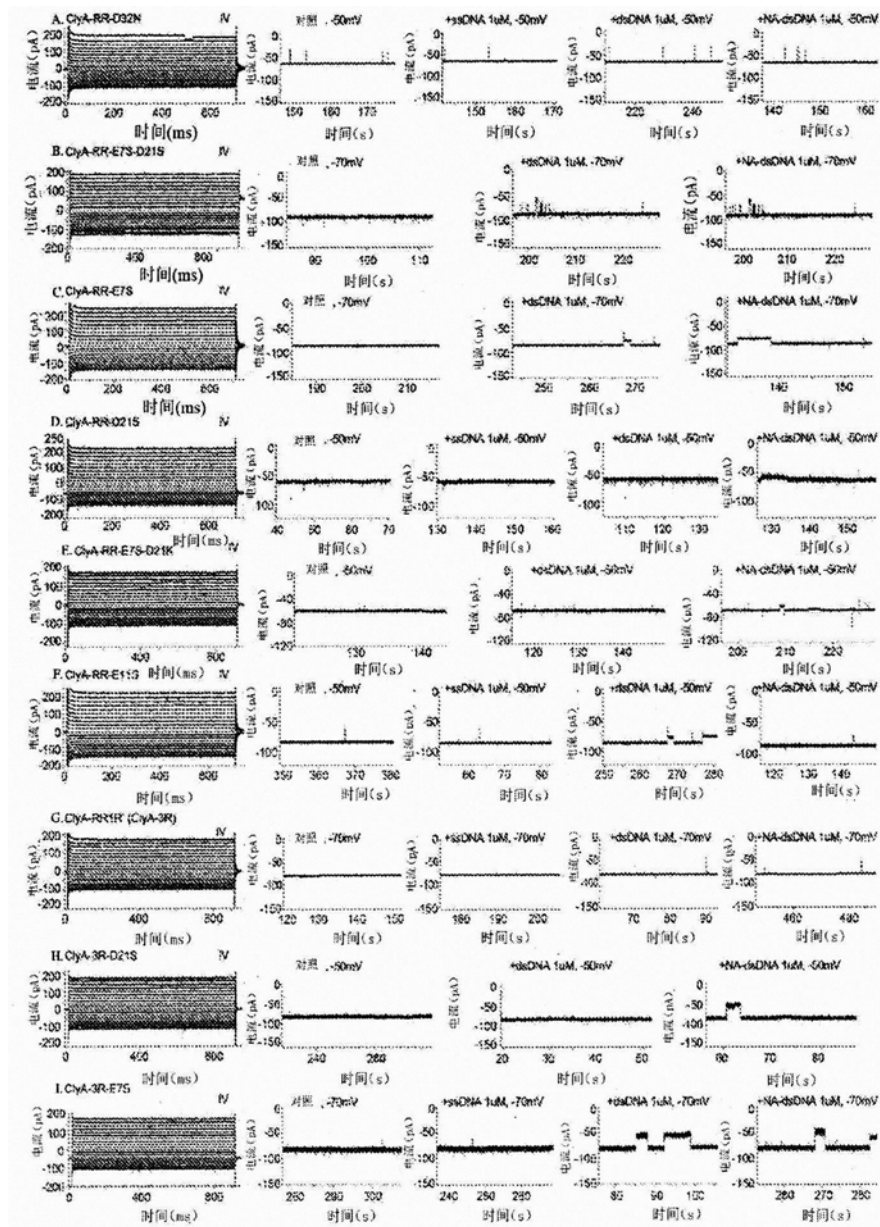


图10

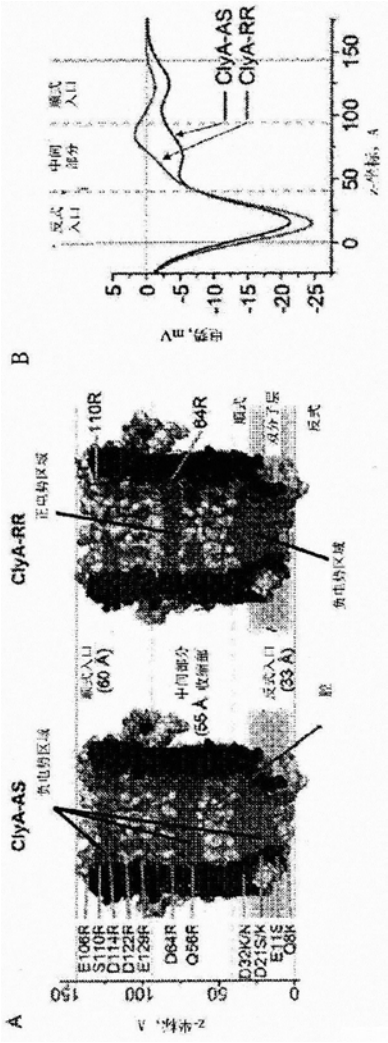


图11

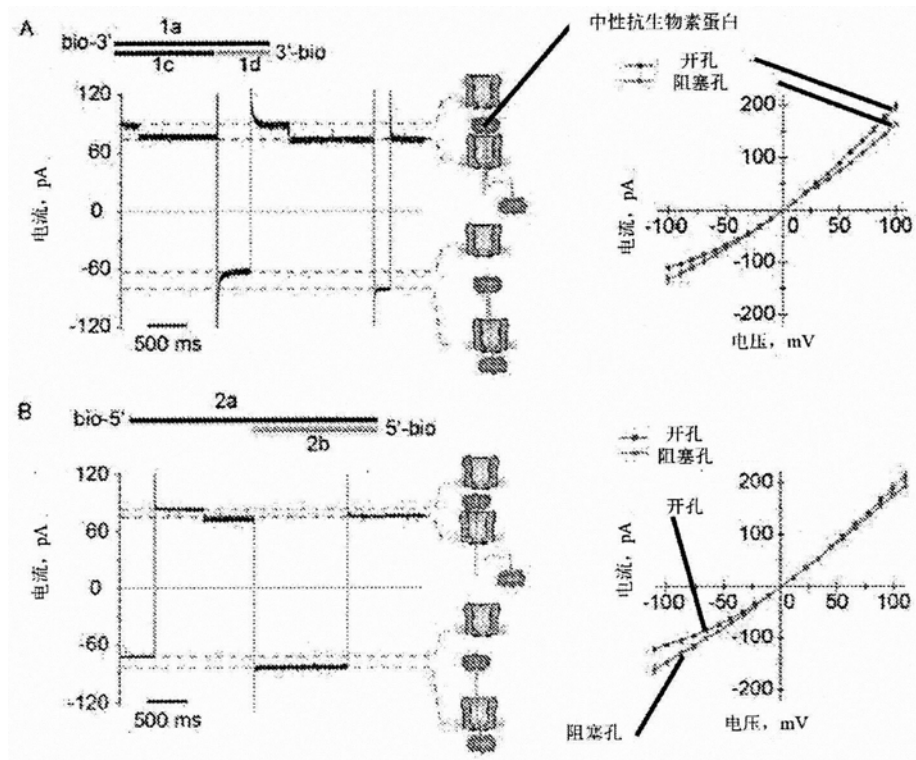


图12

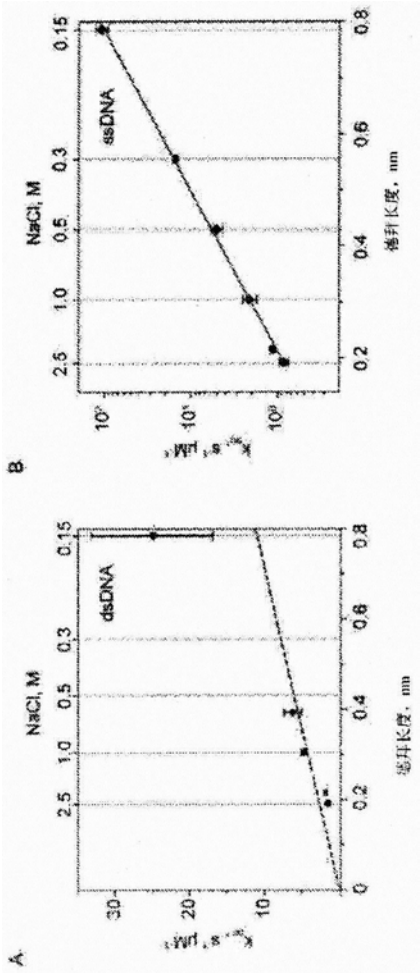


图13

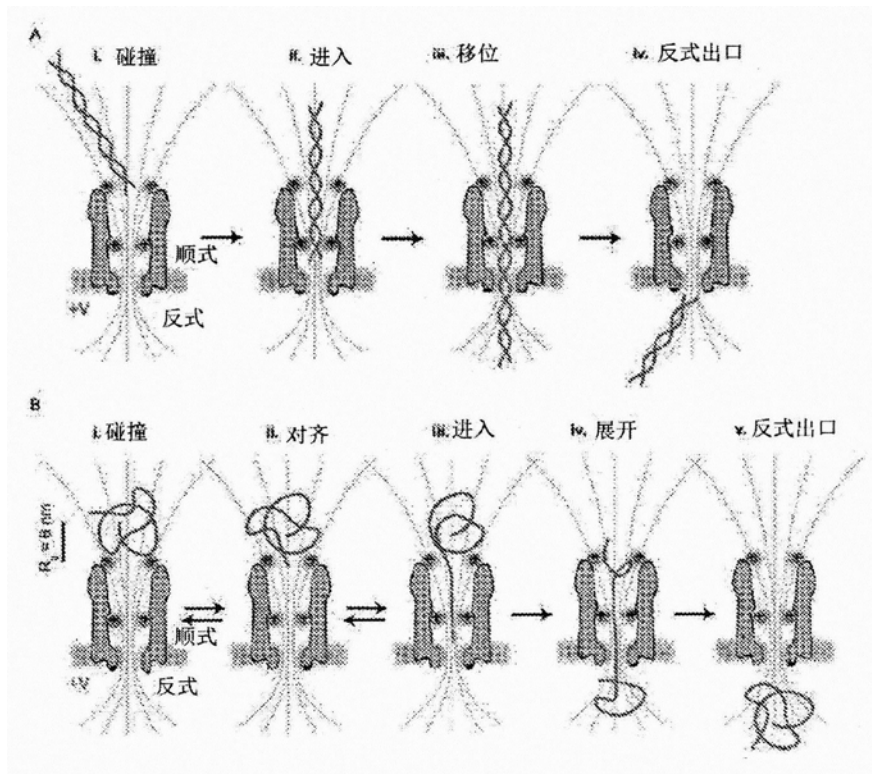


图14

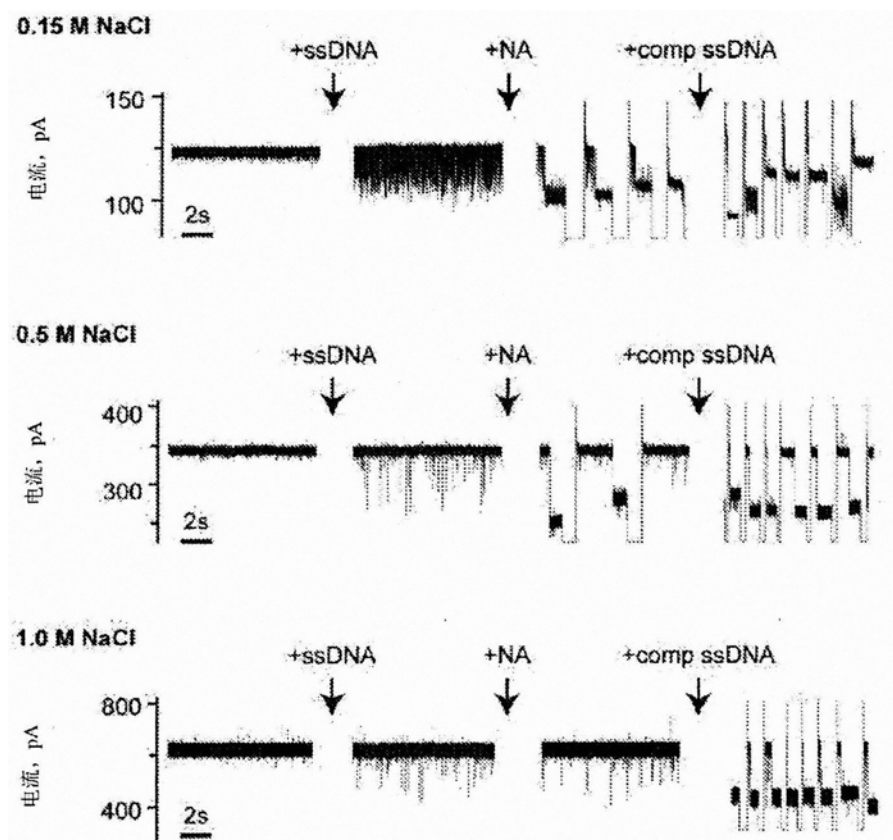


图15

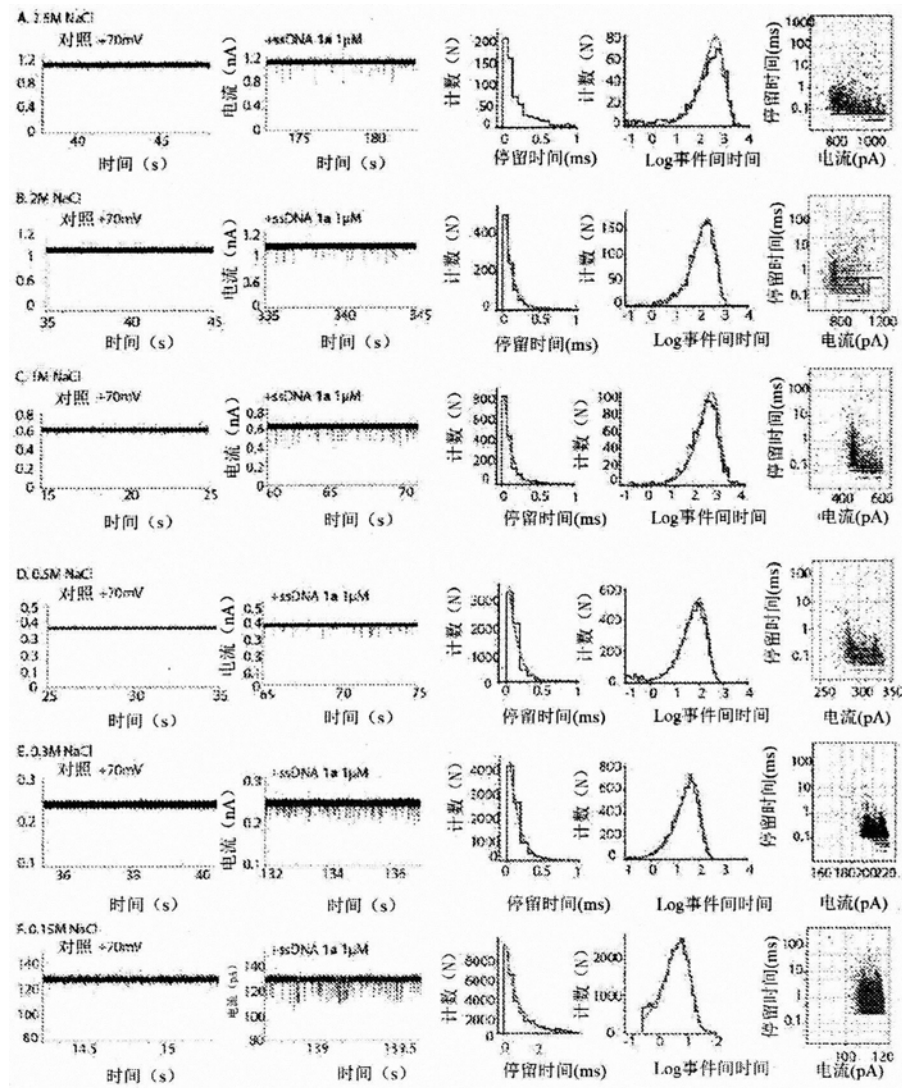


图16

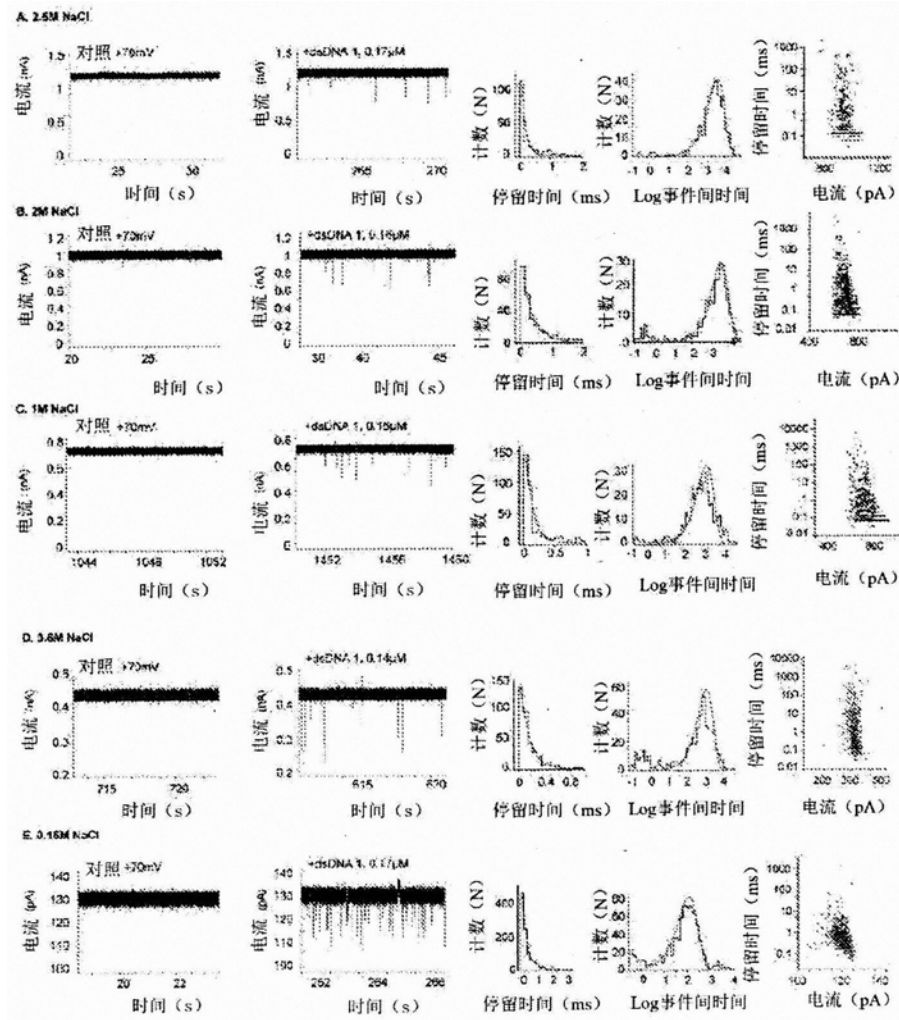


图17

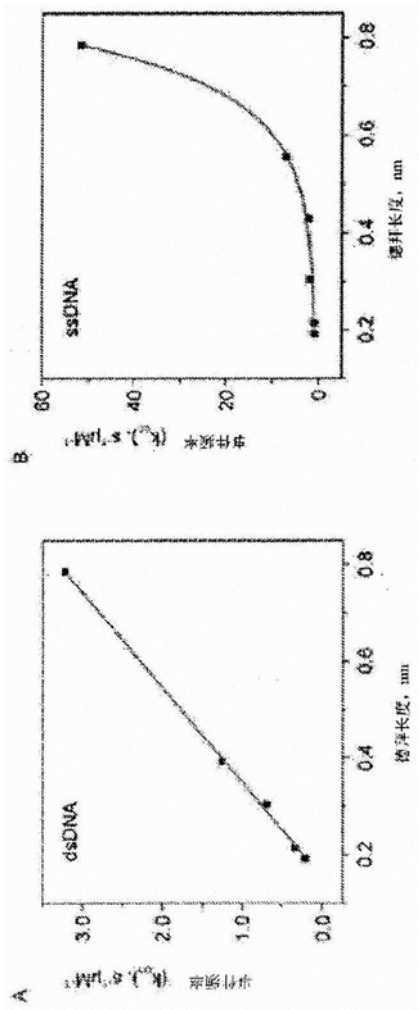


图18

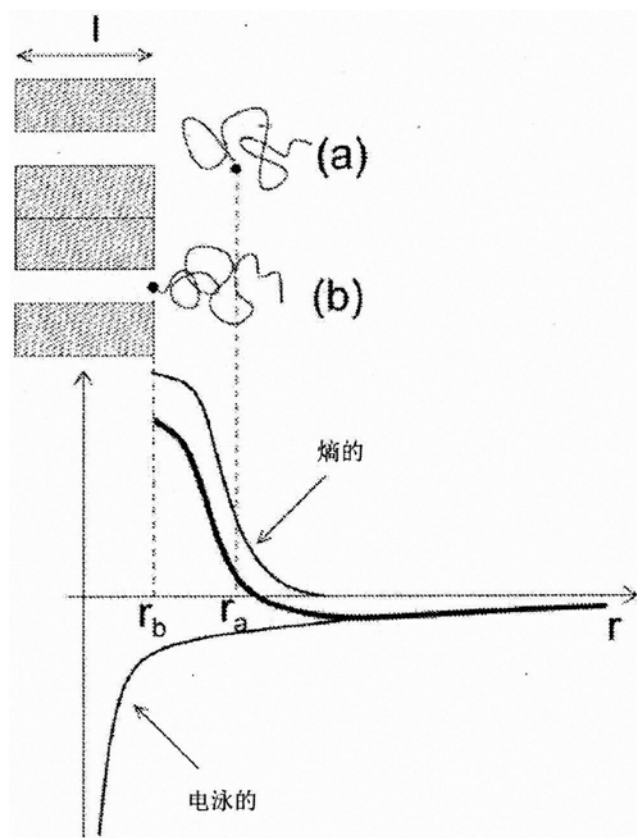


图19

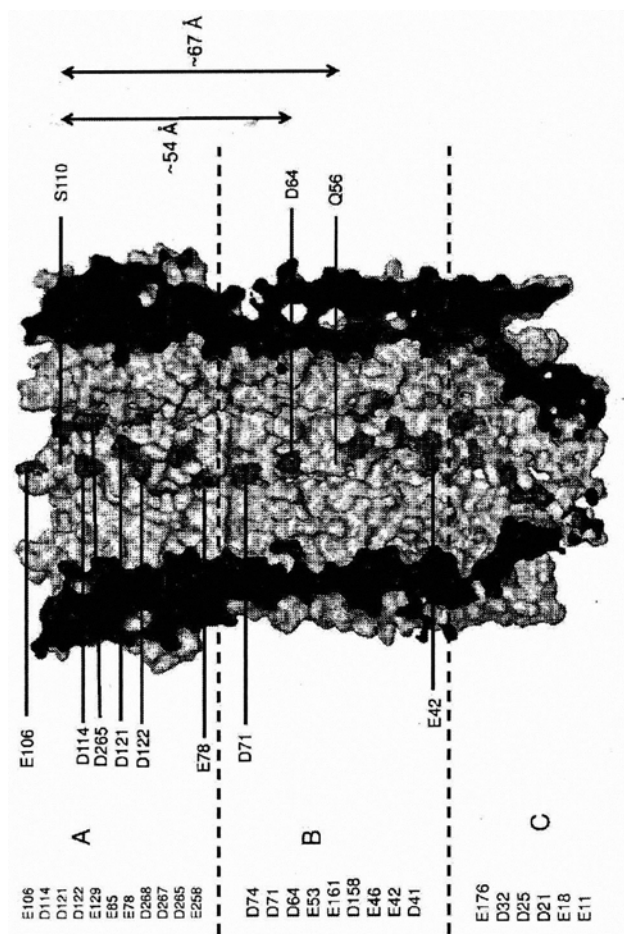


图20