

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021 年 10 月 21 日 (21.10.2021)



(10) 国际公布号
WO 2021/209072 A1

(51) 国际专利分类号:
C07K 1/113 (2006.01) *C07K 1/12* (2006.01)

(21) 国际申请号: PCT/CN2021/094480

(22) 国际申请日: 2021 年 5 月 19 日 (19.05.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010286336.4 2020年4月13日 (13.04.2020) CN
202011168150.5 2020年10月28日 (28.10.2020) CN

(71) 申请人: 南京大学(NANJING UNIVERSITY) [CN/
CN]; 中国江苏省南京市栖霞区仙林大道
163号, Jiangsu 210023 (CN)。

(72) 发明人: 龙亿涛(LONG, Yitao); 中国江苏省南京市栖霞区仙林大道163号, Jiangsu 210023 (CN)。
应佚伦(YING, Yilun); 中国江苏省南京市栖霞区仙林大道163号, Jiangsu 210023 (CN)。

(74) 代理人: 南京苏高专利商标事务所(普通合伙)(NANJING SUGAO PATENT AND TRADEMARK FIRM (ORDINARY PARTNERSHIP)); 中国江苏省南京市秦淮区中山东路198号龙台国际大厦1912室, Jiangsu 210005 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) Title: PROTEIN/POLYPEPTIDE SEQUENCING METHOD USING AEROLYSIN NANOCHANNELS

(54) 发明名称: 一种采用Aerolysin纳米孔道的蛋白质/多肽测序方法

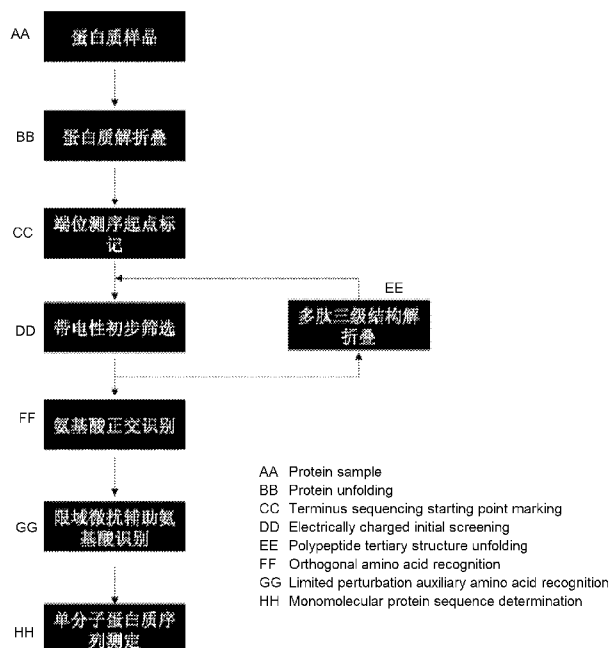


图 1

(57) Abstract: Provided in the present invention is an aerolysin nanochannel-based method for protein/polypeptide sequencing, enabling accurate acquisition of natural amino acids and post-translation-modified, specific-resolution, and monomolecular protein sequences of the amino acids, comprising the following steps: (1) protein unfolding; (2) terminus sequencing starting-point marking; (3) electrically charged initial screening; (4) polypeptide tertiary structure unfolding; (5) orthogonal amino acid recognition; (6) limited perturbation auxiliary amino acid recognition; (7) monomolecular protein sequence determination.



LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。
- 在修改权利要求的期限届满之前进行, 在收到该
修改后将重新公布 (细则48.2(h))。
- 包括关于请求恢复一项或多项优先权要求的信息
(细则26之二.3和48.2(b)(vii))。

(57) 摘要: 本发明提供了一种基于 Aerolysin 纳米孔道的用于蛋白质/多肽测序的方法, 实现对天然
氨基酸及其翻译后修饰的特异性分辨和单分子蛋白质序列的精准获取, 包括以下多个步骤: (1)
蛋白质解折叠; (2) 端位测序起点标记; (3) 带电性初步筛选; (4) 多肽三级结构解折叠; (5) 氨基酸
正交识别; (6) 限域微扰辅助氨基酸识别; (7) 单分子蛋白质序列测定。

一种采用 Aerolysin 纳米孔道的蛋白质/多肽测序方法

[0001]

技术领域

[0002] 本发明属于生物技术领域，具体涉及一种基于 Aerolysin 纳米孔道的蛋白质/多肽测序方法。

背景技术

[0003] 成千上万种不同的蛋白质维持着细胞的所有功能，对生物体内蛋白质氨基酸序列的精准测定可为理解蛋白质功能，其参与的生物学过程以及蛋白质与蛋白质（或其他生物分子）间相互作用提供关键信息。根据遗传信息的传递方向，蛋白质合成是由 DNA 转录、转录后加工、翻译以及翻译后修饰等过程合成的。然而，在转录时一个基因可以多种 mRNA 形式剪接，且同一个蛋白可以许多形式进行翻译后的修饰，因此蛋白质的基因型和表型之间具有较大差异。需要直接对蛋白质的氨基酸序列及翻译后修饰进行精准测定

近年来，随着单分子蛋白质检测需求的日益增长，单分子水平蛋白质分析技术得到快速发展，以尝试解读单个蛋白质分子序列信息，如蛋白质指纹分析荧光技术、隧穿电流蛋白检测技术等。然而，现有荧光测序方法缺少用于检测 20 种不同氨基酸且发射峰间无明显重叠的高效有机荧光团来特异性标记 20 种氨基酸。而隧穿电流检测技术所使用的亚纳米尺度隧穿测量界面难以稳定制备，这些挑战使得现有技术虽可实现几种氨基酸的区分，但尚不能实现 20 种氨基酸及其翻译后修饰的有效识别，更难以获取氨基酸序列信息。因此，目前单分子蛋白质测序仍然面临着巨大的挑战，亟需发展针对 20 种氨基酸序列信息灵敏检测的新原理，建立针对单个蛋白质分子氨基酸序列及翻译后修饰精准测量的创新方法。

[0004]

发明内容

[0005] 本发明所要解决的技术问题是：目前单分子蛋白质测序仍然面临着巨大的挑战，亟需发展针对 20 种氨基酸序列信息灵敏检测的新原理，建立针对单个蛋白质分子氨基酸序列及翻译后修饰精准测量的创新方法。

[0006] 为了解决上述技术问题, 本发明提供了一种基于 Aerolysin 纳米孔道的用于蛋白质/多肽测序的方法, 实现对天然氨基酸及其翻译后修饰的特异性分辨和单分子蛋白质序列的精准获取, 包括以下多个步骤:

包括以下步骤:

(1) 蛋白质解折叠; (2) 端位测序起点标记; (3) 带电性初步筛选; (4) 多肽三级结构解折叠; (5) 氨基酸正交识别; (6) 限域微扰辅助氨基酸识别; (7) 单分子蛋白质序列测定。

[0007] 所述步骤 (1) 蛋白质解折叠中, 单个蛋白质分子在纳米孔道测序前必须解开其高级结构以线性直链的方式进入单个纳米孔道, 单个蛋白质在经由温度调控和 pH 调控的方法将去折叠。

[0008] 步骤 (2) 端位测序起点标记利用特定序列的肽核酸、寡聚核苷酸、多肽链或有机功能团标记去折叠多肽链的 N 端或 C 端为测序起点, 从而获得离子流起点标签信号。

[0009] 步骤 (3) 带电性初步筛选中, 设计电性的初步筛选纳米孔道。

[0010] 步骤 (4) 多肽三级结构解折叠中, 为辅助电荷初筛, 进一步设计三级结构解折叠纳米孔道, 即在生物纳米孔道入口处构建解折叠域以进一步打开多肽分子结构, 即突变型 T284/F/Y/I/L/W 或 G214/F/Y/I/L/W。

[0011] 步骤 (5) 氨基酸正交识别中, 采用针对每一种带电特性的线性多肽分子测序, 设计至少包含以下六种正交识别的纳米孔道:

(a) 基于静电相互作用, 即突变型 T218K/R/H/D/E、S278K/R/H/D/E、S276K/R/H/D/E、T274K/R/H/D/E、A224Q/N/D/E/R/K/H 拟实现对第一类氨基酸的识别, 包括但不限于 H、R、K、E、D、Q、N、W;

(b) 基于氢键及亲水作用, 即突变型 T218N/Q、Q212R/K/H、D209S/T、S276Q/N、D222G/A/S 或 A224E/D, 拟实现对第二类氨基酸的识别, 包括但不限于 Q、N、Y、T、S、C、G、H; 其中, 组氨酸 His 中 R 基的 pKa 为 7, 可通过微调 pH 使其不带电, 从而基于其与关键区域的氢键相互作用使其该特异性纳米孔道中实现区分;

(c) 基于范德华相互作用, 即突变型 R220S/T/A、D222G/A、S236I/L/V、G270I/L、T232I/L/V、T274G/A/I/L 或 K238F/Y/W, 拟实现对第三类氨基酸的识别, 包括但不限于 I、L、M、V、P、A、C、G;

(d) 基于部分氨基酸侧链的大 p 键, 即突变型 D222W/H/F/Y、S276F/Y、A224K/R/W、

S272W/H 或 T274W/H/F/Y, 拟实现对第四类氨基酸的识别, 包括但不限于 W、P、F、Y、H、I、L、V;

(e) 基于小位阻效应, 即突变型 S276F/Y/I/L、S278F/Y/I/L/P、T274W/P、S236W 或 K238G/W/I/L/F/Y/P, 拟实现对第五类小体积氨基酸的识别, 包括但不限于 A、C、G、S、T、V;

(f) 基于大位阻效应, 即突变型 T218G/A、S276G/A、S278G/A、T274G/A、N226D/E、Q268S/T/G/A, 拟实现第六类大体积氨基酸的识别, 包括但不限于 W、H、I、K、R、Y。

[0012] 步骤(6) 限域微扰辅助氨基酸识别中, 针对部分结构差异较小的氨基酸及异构体氨基酸之间可能引入的识别误差, 配合特定纳米孔道进一步提高测序准确性, 引入交变电场、光学微扰测量体系, 同时设计扰动体系微扰放大纳米孔道, 所述特定纳米孔道如下所示:

(a) 突变型 S236D/E/K/H/R、A260D/E/K/H/R、K238H/R/D/E、T240D/E、S256H/R/W 结合交变电场微扰体系;

(b) 突变型 S236W/H、K238I/L、S256Y/F/W、P249W、V250I/L/F/Y/W 结合光学微扰体系。

[0013] 步骤(7) 单分子蛋白质序列测定的方法为:

(a) 使所述蛋白或多肽与所述孔接触, 使得所述蛋白或多肽相对于所述孔移动;

(b) 在所述蛋白或多肽相对于所述孔移动时, 测量穿过所述孔的电流, 其中所述电流指示所述蛋白或多肽的一或多个特征, 包含电流信号的形状、幅值、持续时间, 根据数学变换解析出该电流信号的特征, 建立多肽数据库进行数据的相互校正, 从而表征所述蛋白或多肽。

[0014] 优选的, 上述采用 Aerolysin 纳米孔道蛋白质/多肽测序的方法, 具体步骤如下所示:

(1) 样本前处理: 采用升高温度至 60-100℃、降低溶液 pH 至 0-5 的方法破坏蛋白质内部氢键, 同时采用三(2-羧乙基)膦(TCEP)或二硫苏糖醇(DTT)还原剂破坏蛋白质其 S-S 键, 释放并线性化单个蛋白质中的多肽链;

(2) 通过在多肽链的 N 端特异性修饰肽核酸 PNA、寡聚核苷酸、多肽链或有机功能团使其在进入纳米孔道的起始或终止产生特异性的离子流阻断台阶信号或荧光信号,

从而确定单个多肽分子纳米孔道测序的起点，并为多个正交识别纳米孔道并行测序信号的互相校正提供起始时间标签；

(3) 采用变性剂以及设计构建“三级结构解折叠纳米孔道”实现多肽三级结构的解折叠，其中，对于“三级结构解折叠纳米孔道”的设计为在 Aerolysin 纳米孔道入口处仿生构建蛋白酶体 19S 域的中心氨基酸环境，用于增强纳米孔道入口处与多肽的特异性非共价相互作用，并借助电驱动力，逐步破坏多肽分子内部的弱相互作用，驱动其进入限域孔道内部并实现线性解折叠，从而克服多肽三级结构对纳米孔道多肽测序的一大挑战；

(4) 设计可驱动不同带电性质多肽的功能化 Aerolysin 纳米孔道，初步筛选多肽的电荷性质以匹配选择下一步正交测序纳米孔道；

(5) 基于静电作用、氢键及亲水作用、范德华相互作用、氨基酸大 π 键相互作用、大位阻效应和小位阻效应，针对由于每一种带电特性的多肽构建 6 类正交识别纳米孔道用以专一性识别氨基酸序列；

(6) 在每一个正交识别纳米孔道入口区域引入易形成氢键的氨基酸，并调节该区域限域孔道结构，从而设计构建多肽二级结构标签域，在该区域中，孔道内部氨基酸残基将与不同二级结构的多肽形成特定规则的氢键相互作用，从而诱导特异性的离子流阻断及特定离子迁移率的变化，形成二级结构标签离子流特征，用以数据处理时对单个蛋白质测序离子流电信号的校准降噪；

(7) 针对氨基酸正交识别可能存在的氨基酸识别误差，采用离子流限域微扰技术，结合特定设计的纳米孔道放大温度扰动、交变电场扰动及光学扰动对孔道内离子迁移率的影响，进一步识别离子迁移频率特征，从而提高纳米孔道测量界面的氨基酸识别能力，精准获得蛋白质单分子序列信息；

(8) 在所述蛋白质或多肽相对于所述 Aerolysin 纳米孔道移动时，进行一或多个测量，测量和分析穿过所述孔的电流，包括电流幅值、频率、形状和持续时间等特征，从而测定所述分析物是否存在或其一或多个特征。根据数学变换解析出该电流信号的特征，建立多肽数据库进行数据的相互校正，从而表征所述蛋白或多肽。

[0015]

与现有技术相比，本发明的有益效果在于：

(1) 由于多条多肽链通过氢键或 S-S 键盘曲折叠而成的蛋白质高级结构，其体积较大，

难以进入最窄处仅为 1 nm 的 Aerolysin 纳米孔道。设计蛋白质高级结构解折叠模块，将具有高级结构的蛋白质分子解折叠成线性多肽分子。基于此，本发明采用升高温度、降低溶液 pH 等方法破坏蛋白质内部氢键，同时采用三(2-羧乙基)膦 (TCEP)、二硫苏糖醇 (DTT) 等还原剂破坏其 S-S 键，释放并线性化单个蛋白质中多条多肽链。

[0016] (2) 一条多肽分子可以从 N 端或 C 端进入纳米孔道。如实现根据时序纳米孔道特征信号顺序读出氨基酸序列，亟需确定单个多肽分子进入纳米孔道的端位，即确定测序的起始方向。针对这一目标，本发明通过在多肽的 N 端特异性修饰肽核酸 PNA (如包含多个腺嘌呤的 PNA 序列)、寡聚核苷酸、多肽链或有机功能团 (如 FAM)，使其在进入纳米孔道的起始或终止产生特异性的离子流阻断台阶信号或荧光信号等特殊信号，从而确定单个多肽分子纳米孔道测序的起点，并为多个正交识别纳米孔道并行测序信号的互相校正提供起始时间标签。

[0017] (3) 由于在溶液中多肽二级结构可能进一步盘曲折叠成三级结构，使其具有较大的三维尺寸，因此难以进入纳米孔道。针对这一情况，本发明采用变性剂 (如盐酸胍 GdHCl) 以及设计构建“三级结构解折叠纳米孔道”等方法实现多肽三级结构的解折叠。其中，对于“三级结构解折叠纳米孔道”的设计，拟在 Aerolysin 纳米孔道入口处仿生构建蛋白酶体 19S 域的中心氨基酸环境 (如突变型 T210Y&S213W 等)，增强纳米孔道入口处与多肽的特异性非共价相互作用，并借助电泳力、电渗流、介电泳力等多种电驱动力，逐步破坏多肽分子内部的弱相互作用，驱动其进入限域孔道内部并，实现线性解折叠，从而克服多肽三级结构对纳米孔道多肽测序的一大挑战。

[0018] (4) 本发明设计可驱动不同带电性质多肽的功能化 Aerolysin 纳米孔道，初步筛选多肽的电荷性质以匹配选择下一步正交测序纳米孔道。本发明设计至少 4 种“电性初筛纳米孔道”分别针对性捕获 4 类带电性质多肽即负电荷多肽、正电荷多肽、电中性且正负电荷相互屏蔽多肽以及电中性且正负电荷分离多肽。4 种“电性初筛纳米孔道”设计如下：

(i) 特异性识别带负电多肽的“电性初筛纳米孔道”。通过调节孔道内部关键区域直径或转移孔道内电荷至直径较宽区域 (如突变型 T274N/Q/I/L、T232D/E、K238H/D/R/F/A/C/G/Q/E/K/L/M/N/S/Y/T/I/W/P/V、S280T/N/Q/H/I/L 等)，控制其孔道内部由孔道最窄处电荷决定的电渗流为零，并减小介电泳力，使单个带负电多肽由电泳力驱动进入纳米孔道。

[0019] (ii) 特异性识别带正电多肽的“电性初筛纳米孔道”。通过引入或增大孔道内负电荷的分布（如突变型 T274D/E、T218D/E、S276D/E、S278D/E、K238A/N/D/E/Q、R282D/E/S/T/N/Q/A、R220D/E/S/T/N/Q/A 等），从而调节由孔道内部阳离子决定的电渗流。在实验中，施加反向电压，从而实现对单个正电荷多肽的高效捕获。

[0020] (iii) 特异性识别电中性且正负电荷相互屏蔽多肽的“电性初筛纳米孔道”。对于电中性且其中正负电荷互相屏蔽（即正负电荷距离较近）的多肽，通过在孔道内直径较小区域引入正电荷氨基酸（如突变型 T218K/R/H/N/Q、S276K/R/H、S278K/R/H/N/Q、S274K/R/H、N226K/R/H、S272K/R/H、G270K/R/H、S228K/R/H、Q268K/R/H、T230K/R/H、A266K/R/H、T232K/R/H、S264K/R/H、G234K/R/H、N262K/R/H、S236K/R/H、A260K/R/H、S280N/Q 等），增强孔道内部由阴离子决定的电渗流，从而增强纳米孔道对于电中性且正负电荷相互屏蔽多肽的捕获效率，获得特异性离子流响应。

[0021] (iiii) 特异性识别电中性且正负电荷分离多肽的“电性初筛纳米孔道”。对于电中性且其中正负电荷分离的多肽，通过增强孔道入口处的电势梯度（如突变型 S280Q/N/A、T284Q/N/A、G214Q/N/A），调控于非线性的电场强度，从而利用介电泳力驱动单个电中性且正负电荷分离多肽进入孔道。

[0022] (5) 基于静电作用、氢键及亲水作用、范德华相互作用、氨基酸大 π 键相互作用、大位阻效应和小位阻效应，本发明针对由于每一种带电特性的多肽构建至少 6 类正交识别纳米孔道用以专一性识别氨基酸序列，如下所示：

(I) 基于静电相互作用，在孔道内现有电流传感区域中引入带电荷氨基酸，如突变型 T218K/R/H/D/E、S278K/R/H/D/E、S276K/R/H/D/E、T274K/R/H/D/E、A224Q/N/D/E/R/K/H 等。同时，带电荷氨基酸的引入会增强孔道与待测氨基酸之间的氢键、盐键和阳离子- π 相互作用，从而实现对第一类氨基酸的识别，包括但不限于 H、R、K、E、D、Q、N、W。

[0023] (II) 基于氢键及亲水作用，调控孔道内电流传感区域的电势梯度，如突变型 T218N/Q、Q212R/K/H、D209S/T、S276Q/N、D222G/A/S、A224E/D 等，加快带电荷氨基酸通过该区域的速度并延长极性不带电氨基酸在此区域内的停留时间，从而实现对第二类氨基酸的识别，包括但不限于 Q、N、Y、T、S、C、G、H。其中，组氨酸 H 中 R 基的 pK_a 约为 7，可通过微调 pH 使其不带电，从而基于其与纳米孔道关键传感区域的氢键相互作用，使其在该特异性纳米孔道中实现特征区分。

[0024] (III) 基于范德华相互作用, 调控孔道整体的电势分布和孔道立体结构分布, 引入疏水氨基酸域, 如突变型 R220S/T/A、D222G/A、S236I/L/V、G270I/L、T232I/L/V、T274G/A/I/L、K238F/Y/W 等, 使得电流传感区域从野生型孔道的静电敏感域转移到突变型孔道的疏水域中, 从而通过相互作用延长特定氨基酸在该区域的停留时间, 获得特征离子流信号, 从而拟实现对第三类氨基酸的识别, 包括但不限于 I、L、M、V、P、A、C、G。

[0025] (IV) 基于部分氨基酸大 π 键相互作用, 在调控孔道立体结构和电势分布的基础上, 重构 Aerolysin 纳米孔道内的电流传感区域组成, 构建以正电荷氨基酸及输水氨基酸为主导的灵敏区域, 如突变型 D222W/H/F/Y、S276F/Y、A224K/R/W、S272W/H、T274W/H/F/Y 等, 增强该灵敏区域与特定氨基酸的 π - π 相互作用、阳离子- π 键相互作用、 π - π 相互作用等, 从而实现对第四类氨基酸的识别, 包括但不限于 W、P、F、Y、H、I、L、V。

[0026] (V) 基于大位阻效应, 进一步减小孔道内电流传感区域的限域空间, 增大该区域的位阻, 如突变型 S276F/Y/I/L、S278F/Y/I/L/P、T274W/P、S236W、K238G/W/I/L/F/Y/P 等, 延长所有氨基酸通过该区域的时间, 增强小体积氨基酸的离子流信号电流幅值, 并使得大体积氨基酸造成近乎全阻断离子流台阶, 从而特异性区分氨基酸体积, 实现对第五类小体积氨基酸的识别, 包括但不限于 A、C、G、S、T、V。

[0027] (VI) 基于小位阻效应, 调控孔道内的立体结构, 增大电流关键区域的尺寸, 如突变型 T218G/A、S276G/A、S278G/A、T274G/A、N226D/E、Q268S/T/G/A 等, 并基于多肽整体的带电荷类型减小纳米孔道内的电渗流, 从而进一步减小小体积氨基酸的电流响应, 增强大体积氨基酸的电流差异, 实现第六类大体积氨基酸的识别, 包括但不限于 W、H、I、K、R、Y。

[0028] (6) 本发明在每一个正交识别纳米孔道靠近入口区域引入易形成氢键的氨基酸, 并调控该区域限域孔道结构, 从而设计构建多肽二级结构标签域。在该区域中, 孔道内部氨基酸残基将与不同二级结构的多肽形成特定规则的氢键相互作用, 从而诱导特异性的离子流阻断及特定离子迁移率的变化, 形成二级结构标签离子流特征, 用以数据处理时对单个蛋白质测序离子流电信号的校准降噪。

[0029] (7) 针对氨基酸正交识别可能存在的氨基酸识别误差, 本发明基于申请人团队前期对孔道内离子流迁移轨迹的研究, 采用离子流限域微扰技术, 结合特定设计的纳米

孔道放大温度扰动、交变电场扰动及光学扰动对孔道内离子迁移率的影响，进一步识别离子迁移频率特征，从而提高纳米孔道测量界面的氨基酸识别能力，精准获得蛋白质单分子序列信息。

[0030] (8) 在所述蛋白质或多肽相对于所述孔移动时，进行一或多个测量，测量和分析穿过所述孔的电流，包括电流幅值、频率、形状和持续时间等特征，从而测定所述分析物是否存在或其一或多个特征。根据数学变换解析出该电流信号的特征，建立多肽数据库进行数据的相互校正，从而表征所述蛋白或多肽。

[0031] 其中，同种多肽分子在步骤(6)所示的6类正交识别纳米孔道中的测量时间具有差异，且不同氨基酸在每一种孔道内的传感灵敏区域停留时间不同，导致了各个孔道之间对氨基酸的测序时间轴存在差异。因此，本发明在六大类氨基酸识别中分别引入了标签氨基酸，如第(I)、(II)、(IV)、(VI)类正交识别纳米孔道中的组氨酸H，第(III)、(IV)、(VI)类正交识别纳米孔道中的异亮氨酸(I)，第(II)、(III)、(V)类正交识别纳米孔道”中的半胱氨酸C，第(II)、(IV)、(VI)类正交识别纳米孔道中的酪氨酸Y等。大多数待测氨基酸均有至少两类正交识别纳米孔道进行专一性识别，从而从多角度修正、统一不同孔道内的测序时间轴，实现对六种孔道测得离子流电信号的交叉、校正和精准整合。

[0032] 此外，本项目对氨基酸的翻译后修饰可以在氨基酸序列测定的同时被检测。例如，丝氨酸S、苏氨酸T、组氨酸H的磷酸化修饰可在步骤(6)所示的第(I)、(II)类正交识别纳米孔道中识别。天门冬氨酸D、谷氨酸E的甲基化修饰可在步骤(6)所示的第(I)类正交识别纳米孔道”中识别。天门冬酰胺N、苏氨酸T、丝氨酸S的糖基化修饰可在步骤(6)所示的第(I)、(II)、(V)类正交识别纳米孔道中识别。因此，本项目可在识别多肽氨基酸序列的同时，有望同时实现对特定氨基酸翻译后修饰种类、数量、位置的精准测定。

[0033] 其中，步骤(7)中在温度扰动方面是利用温度变化来显著改变分子的随机振动以及分子间的相互作用等，本发明通过在0-40℃范围内精准调控实验体系温度，实现最大化刺激由相互作用产生的离子迁移率变化程度，从而提升单个氨基酸频率特征的信噪比。在频率为0.1-1 MHz的交变电场扰动方面拟设计功能化扰动体系微扰放大纳米孔道，通过定点突变调整孔道的直径和电荷分布，并在孔道内引入离子结合氨基酸，如突变型S236D/E/K/H/R、A260D/E/K/H/R、K238H/R/D/E、T240D/E、S256H/R/W等，如Mg²⁺、

Ni⁺等，从而使得在特定交变电场频率下诱导诱导放大阳离子相互作用引起的离子迁移率变化，从而增强识别具有较小差异的氨基酸，如天门冬酰胺 N 和谷氨酰胺 Q、异亮氨酸 I 和缬氨酸 V 等。

[0034] 在步骤（7）中光学扰动方面，本发明设计高度限域的纳米孔道，使其在电荷敏感区域相互作用类型多样，如突变型 S236W/H、K238I/L、S256Y/F/W、P249W、V250I/L/F/Y/W 等，从而利用特定频率的红外（10000-25000nm）或紫外光（180-400nm），扰动待测多肽与孔道内部特定氨基酸的弱相互作用，如氢键、p-p 相互作用、p-p 相互作用等，提升具有相似弱相互作用以及同分异构体氨基酸的识别，如丝氨酸 S 和苏氨酸 T、亮氨酸 L 和异亮氨酸 I 等。

[0035] 在步骤（8）中测量蛋白质或多肽的一个、两个、三个、四个或五个或更多个特征。一或多个特征优选地是选自：(i)蛋白质或多肽的序列；(ii)蛋白质或多肽是否被修饰以及被修饰的氨基酸种类，位置，个数；(iii)蛋白质或多肽的长度；(iv)蛋白质或多肽的一致性；(v)蛋白质或多肽的构象；(vi)蛋白质或多肽的二级结构。

[0036]

附图说明

[0037] 图 1：蛋白质/多肽测序方法流程图。

[0038] 图 2：N226Q 电性初筛 Aerolysin 纳米孔道分别检测 Glu-Gly-Cys 和 Glu-Cys-Gly 两种三肽分子的原始电流轨迹。

[0039] 图 3：T232K 电性初筛 Aerolysin 纳米孔道检测 Glu-Gly-Cys 和 Glu-Cys-Gly 两种三肽分子混合物的原始电流轨迹。

[0040] 图 4：T232K/K238Q 多肽测序 Aerolysin 纳米孔道分别检测 Glu-Gly-Cys 和 Glu-Cys-Gly 两种三肽分子的原始电流轨迹。

[0041] 图 5：野生型电性初筛 Aerolysin 纳米孔道分别检测模板多肽分子和两种磷酸化多肽分子的原始电流轨迹。

[0042] 图 6：T232K/K238Q 磷酸化检测 Aerolysin 纳米孔道分别检测模板多肽分子和两种磷酸化多肽分子的原始电流轨迹。

[0043]

实施例 1

一种利用半胱氨酸专一型 Aerolysin 纳米孔道多肽分子测序的方法，多肽以 Glu 为引导

链，两种多肽分子的氨基酸序列分别为 Glu-Gly-Cys 和 Glu-Cys-Gly。具体步骤如下：

(1) 设计了 N226Q 和 T232K 两个电性初筛孔道，利用定点突变技术表达并纯化了相应的突变型 Proerolysin 蛋白用于孔道构建，具体步骤如参考专利 CN202010131704.8。

[0044] (2) 将 1mg/mL 的 Proerolysin 蛋白与胰蛋白酶 10: 1 混合并在室温下孵育 6 小时，得到具有成孔活性的 Aerolysin 单体蛋白。

[0045] (3) 将实验温度控制在 $22 \pm 1^\circ\text{C}$ 。在两个检测池中分别加入 1mL (1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, pH=8) 缓冲溶液，用提拉法制备磷脂双分子层，具体步骤参考专利 CN201510047662.9。

[0046] (4) 在稳定的磷脂双分子层形成后，施加 200mV 电压，在 cis 检测池中加入 1 μL 的 Aerolysin 单体蛋白质。Aerolysin 单体自组装形成七聚体并插入磷脂膜中形成稳定的纳米孔道，同时，离子流发生跃升，在 100mV 电压下获得稳定的开孔电流。

[0047] (5) 在 cis 检测池中加入 4 μL 浓度为 50mM 的三肽链，施加 120mV 的外加电压。采集到的原始电流轨迹如图 2-3 所示。图 2 和图 3 为 N226Q 和 T232K 电性初筛 Aerolysin 纳米孔道分别检测 Glu-Gly-Cys 和 Glu-Cys-Gly 两种三肽分子的原始电流轨迹。首先由电性初筛孔道判断多肽的电性，两种三肽分子在 N226Q 电性初筛孔道中分别出现了电流阻断信号（图 2）。在 T232K 电性初筛孔道中没有出现电流阻断信号，图 3 中为加了两种多肽分子混合物的电流轨迹。因此，我们判断两种多肽分子都是由负电荷驱动的。

[0048] (6) 进一步设计了 T232K/K238Q 双突变型多肽测序孔道，利用定点突变技术表达并纯化了突变型 Proerolysin 蛋白用于孔道构建。

[0049] (7) 重复以上纳米孔道构建步骤，分别在 cis 检测池加入两种多肽分子，如图 4 所示，两种多肽分子在 T232K/K238Q 双突变型多肽测序孔道中产生了两种阻断信号，其中阻断时间较长的特征性双台阶的阻断信号是由引导链 Glu 牵引的 N 端进孔的多肽分子产生的阻断信号，而阻断时间很短的信号为非引导链牵引进孔的多肽分子产生的阻断信号。根据阻断信号的形状、阻断时间和程度可以区分两种多肽分子的序列。

[0050]

实施例 2

一种利用突变型 Aerolysin 纳米孔道检测磷酸化多肽的方法，以 S-K-I-G 为引导链，模板多肽序列为 S-K-I-G-S-T-E-N-L，分别在第五位的丝氨酸和第六位的苏氨酸进行磷酸化修饰的到的序列分别为 S-K-I-G-^PS-T-E-N-L 和 S-K-I-G-S-^PT-E-N-L。具体步骤如下：

(1) 设计了野生型电性初筛孔道，表达并纯化了野生型的 Proerolysin 蛋白用于孔道构建，具体步骤如参考专利 CN202010131704.8。

[0051] (2) 将 1mg/mL 的 Proerolysin 蛋白与胰蛋白酶 10: 1 混合并在室温下孵育 6 小时，得到具有成孔活性的 Aerolysin 单体蛋白。

[0052] (3) 将实验温度控制在 $22 \pm 1^\circ\text{C}$ 。在两个检测池中分别加入 1mL (1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, pH=8) 缓冲溶液，用提拉法制备磷脂双分子层，具体步骤参考专利 CN201510047662.9。

[0053] (4) 在稳定的磷脂双分子层形成后，施加 200mV 电压，在 *cis* 检测池中加入 1 μ L 的 Aerolysin 单体蛋白质。Aerolysin 单体自组装形成七聚体并插入磷脂膜中形成稳定的纳米孔道，同时，离子流发生跃升，获得稳定的开孔电流。

[0054] (5) 在 *cis* 检测池中加入 5 μ L 浓度为 1 mM 的多肽溶液，施加 100mV 的外加电压。采集到的原始电流轨迹如图 5 所示，三种多肽分子在野生型 Aerolysin 纳米孔道中产生了很少数的电流阻断信号，判断三种多肽分子由负电荷驱动。

[0055] (6) 进一步设计了 T232K/K238Q 双突变型磷酸化检测孔道，利用定点突变技术表达并纯化了突变型 Proerolysin 蛋白用于孔道构建。

[0056] (7) 重复以上纳米孔道构建步骤，分别在 *cis* 检测池加入三种多肽分子，如图 6 所示，根据多肽分子在 T232K/K238Q 孔道产生的阻断信号的形状、阻断时间和程度可以区分模板多肽分子和两种磷酸化多肽分子。

[0057] 以上显示和描述了本发明的基本原理和主要特征和本发明的优点。本行业的技术人员应该了解，本发明不受上述实施例的限制，上述实施例和说明书中描述的只是说明本发明的原理，在不脱离本发明精神和范围的前提下，本发明还会有各种变化和改进，这些变化和改进都落入要求保护的本发明范围内。本发明要求保护范围由所附的权利要求书及其等效物界定。

权 利 要 求 书

1. 一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 包括以下步骤:

(1) 蛋白质解折叠; (2) 端位测序起点标记; (3) 带电性初步筛选; (4) 多肽三级结构解折叠; (5) 氨基酸正交识别; (6) 限域微扰辅助氨基酸识别; (7) 单分子蛋白质序列测定。

2. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 所述步骤 (1) 蛋白质解折叠中, 单个蛋白质分子在纳米孔道测序前必须解开其高级结构以线性直链的方式进入单个纳米孔道, 单个蛋白质在经由温度调控和 pH 调控的方法将去折叠。

3. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 步骤 (2) 端位测序起点标记利用特定序列的肽核酸、寡聚核苷酸、多肽链或有机功能团标记去折叠多肽链的 N 端或 C 端为测序起点, 从而获得离子流起点标签信号。

4. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 步骤 (3) 带电性初步筛选中, 设计电性的初步筛选纳米孔道。

5. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 步骤 (4) 多肽三级结构解折叠中, 为辅助电荷初筛, 进一步设计三级结构解折叠纳米孔道, 即在生物纳米孔道入口处构建解折叠域以进一步打开多肽分子结构, 即突变型 T284/F/Y/I/L/W 或 G214/F/Y/I/L/W。

6. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 步骤 (5) 氨基酸正交识别中, 采用针对每一种带电特性的线性多肽分子测序, 设计至少包含以下六种正交识别的纳米孔道:

(a) 基于静电相互作用, 即突变型 T218K/R/H/D/E、S278K/R/H/D/E、S276K/R/H/D/E、T274K/R/H/D/E、A224Q/N/D/E/R/K/H 拟实现对第一类氨基酸的识别, 包括但不限于 H、R、K、E、D、Q、N、W;

(b) 基于氢键及亲水作用, 即突变型 T218N/Q、Q212R/K/H、D209S/T、S276Q/N、D222G/A/S 或 A224E/D, 拟实现对第二类氨基酸的识别, 包括但不限于 Q、N、Y、T、S、C、G、H; 其中, 组氨酸 His 中 R 基的 pKa 为 7, 可通过微调 pH 使其不带电, 从而基于其与关键区域的氢键相互作用使其该特异性纳米孔道中实现区分;

(c) 基于范德华相互作用, 即突变型 R220S/T/A、D222G/A、S236I/L/V、G270I/L、T232I/L/V、T274G/A/I/L 或 K238F/Y/W, 拟实现对第三类氨基酸的识别, 包括但不限于

于 I、L、M、V、P、A、C、G；

(d) 基于部分氨基酸侧链的大 p 键，即突变型 D222W/H/F/Y、S276F/Y、A224K/R/W、S272W/H 或 T274W/H/F/Y，拟实现对第四类氨基酸的识别，包括但不限于 W、P、F、Y、H、I、L、V；

(e) 基于小位阻效应，即突变型 S276F/Y/I/L、S278F/Y/I/L/P、T274W/P、S236W 或 K238G/W/I/L/F/Y/P，拟实现对第五类小体积氨基酸的识别，包括但不限于 A、C、G、S、T、V；

(f) 基于大位阻效应，即突变型 T218G/A、S276G/A、S278G/A、T274G/A、N226D/E、Q268S/T/G/A，拟实现第六类大体积氨基酸的识别，包括但不限于 W、H、I、K、R、Y。

7. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法，其特征在于，步骤（6）限域微扰辅助氨基酸识别中，针对部分结构差异较小的氨基酸及异构体氨基酸之间可能引入的识别误差，配合特定纳米孔道进一步提高测序准确性，引入交变电场、光学微扰测量体系，同时设计扰动体系微扰放大纳米孔道，所述特定纳米孔道如下所示：

(a) 突变型 S236D/E/K/H/R、A260D/E/K/H/R、K238H/R/D/E、T240D/E、S256H/R/W 结合交变电场微扰体系；

(b) 突变型 S236W/H、K238I/L、S256Y/F/W、P249W、V250I/L/F/Y/W 结合光学微扰体系。

8. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法，其特征在于，步骤（7）单分子蛋白质序列测定的方法为：

(a) 使所述蛋白或多肽与所述孔接触，使得所述蛋白或多肽相对于所述孔移动；

(b) 在所述蛋白或多肽相对于所述孔移动时，测量穿过所述孔的电流，其中所述电流指示所述蛋白或多肽的一或多个特征，包含电流信号的形状、幅值、持续时间，根据数学变换解析出该电流信号的特征，建立多肽数据库进行数据的相互校正，从而表征所述蛋白或多肽。

9. 根据权利要求 1 所述采用 Aerolysin 纳米孔道蛋白质/多肽测序的方法，具体步骤如下所示：

样本前处理：采用升高温度至 60-100℃、降低溶液 pH 至 0-5 的方法破坏蛋白质内部氢键，同时采用三(2-羧乙基)膦（TCEP）或二硫苏糖醇（DTT）还原剂破坏蛋白质其 S-S

键，释放并线性化单个蛋白质中的多肽链；

通过多肽链的 N 端特异性修饰肽核酸 PNA、寡聚核苷酸、多肽链或有机功能团使其在进入纳米孔道的起始或终止产生特异性的离子流阻断台阶信号或荧光信号，从而确定单个多肽分子纳米孔道测序的起点，并为多个正交识别纳米孔道并行测序信号的互相校正提供起始时间标签；

采用变性剂以及设计构建“三级结构解折叠纳米孔道”实现多肽三级结构的解折叠，其中，对于“三级结构解折叠纳米孔道”的设计为在 Aerolysin 纳米孔道入口处仿生构建蛋白酶体 19S 域的中心氨基酸环境，用于增强纳米孔道入口处与多肽的特异性非共价相互作用，并借助电驱动力，逐步破坏多肽分子内部的弱相互作用，驱动其进入限域孔道内部并实现线性解折叠，从而克服多肽三级结构对纳米孔道多肽测序的一大挑战；

设计可驱动不同带电性质多肽的功能化 Aerolysin 纳米孔道，初步筛选多肽的电荷性质以匹配选择下一步正交测序纳米孔道；

基于静电作用、氢键及亲水作用、范德华相互作用、氨基酸大 π 键相互作用、大位阻效应和小位阻效应，针对由于每一种带电特性的多肽构建 6 类正交识别纳米孔道用以专一性识别氨基酸序列；

在每一个正交识别纳米孔道入口区域引入易形成氢键的氨基酸，并调节该区域限域孔道结构，从而设计构建多肽二级结构标签域，在该区域中，孔道内部氨基酸残基将与不同二级结构的多肽形成特定规则的氢键相互作用，从而诱导特异性的离子流阻断及特定离子迁移率的变化，形成二级结构标签离子流特征，用以数据处理时对单个蛋白质测序离子流电信号的校准降噪；

针对氨基酸正交识别可能存在的氨基酸识别误差，采用离子流限域微扰技术，结合特定设计的纳米孔道放大温度扰动、交变电场扰动及光学扰动对孔道内离子迁移率的影响，进一步识别离子迁移频率特征，从而提高纳米孔道测量界面的氨基酸识别能力，精准获得蛋白质单分子序列信息；

在所述蛋白质或多肽相对于所述 Aerolysin 纳米孔道移动时，进行一或多个测量，测量和分析穿过所述孔的电流，包括电流幅值、频率、形状和持续时间等特征，从而测定所述分析物是否存在或其一或多个特征，根据数学变换解析出该电流信号的特征，建立多肽数据库进行数据的相互校正，从而表征所述蛋白或多肽。

10. 根据权利要求 1 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法，其特征在

于, 所述步骤(2)中的有机功能团为FAM、VIC、CY5、HEX、ROX。

11. 根据权利要求1所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 所述步骤(3)中采用的变性剂为盐酸胍或 GdHCl。

12. 根据权利要求1所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 所述步骤(3)中所述电驱动力为电泳力、电渗流、介电泳力。

13. 根据权利要求1所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 所述步骤(3)中设计可驱动不同带电性质多肽的功能化 Aerolysin 纳米孔道具体方法为: 采用4种“电性初筛纳米孔道”分别针对性捕获4类带电性质多肽即负电荷多肽、正电荷多肽、电中性且正负电荷相互屏蔽多肽以及电中性且正负电荷分离多肽。

14. 根据权利要求5所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法, 其特征在于, 所述步骤4种“电性初筛纳米孔道”分别为特异性识别带负电多肽的“电性初筛纳米孔道”、特异性识别带正电多肽的“电性初筛纳米孔道”、特异性识别电中性且正负电荷相互屏蔽多肽的“电性初筛纳米孔道”、特异性识别电中性且正负电荷分离多肽的“电性初筛纳米孔道”; 这4种孔道的设计分别如下:

(i) 特异性识别带负电多肽的“电性初筛纳米孔道”: 通过调节 Aerolysin 纳米孔道内部关键区域直径或转移孔道内电荷至直径较宽区域, 即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 T274N/Q/I/L、T232D/E、K238H/D/R/F/A/C/G/Q/E/K/L/M/N/S/Y/T/I/W/P/V 或 S280T/N/Q/H/I/L, 控制其纳米孔道内部由孔道最窄处电荷决定的电渗流为零, 并减小介电泳力, 使单个带负电多肽由电泳力驱动进入纳米孔道;

(ii) 特异性识别带正电多肽的“电性初筛纳米孔道”: 通过引入或增大 Aerolysin 纳米孔道内负电荷的分布, 即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 T274D/E、T218D/E、S276D/E、S278D/E、K238A/N/D/E/Q、R282D/E/S/T/N/Q/A 或 R220D/E/S/T/N/Q/A, 从而调节由孔道内部阳离子决定的电渗流; 在实验中, 施加反向电压, 从而实现对单个正电荷多肽的高效捕获;

(iii) 特异性识别电中性且正负电荷相互屏蔽多肽的“电性初筛纳米孔道”: 对于电中性且其中正负电荷互相屏蔽的多肽, 通过在 Aerolysin 纳米孔道内直径较小区域引入正电荷氨基酸, 即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 T218K/R/H/N/Q、S276K/R/H、S278K/R/H/N/Q、S274K/R/H、N226K/R/H、S272K/R/H、G270K/R/H、

S228K/R/H、Q268K/R/H、T230K/R/H、A266K/R/H、T232K/R/H、S264K/R/H、G234K/R/H、N262K/R/H、S236K/R/H、A260K/R/H 或 S280N/Q，用以增强孔道内部由阴离子决定的电渗流，从而增强纳米孔道对于电中性且正负电荷相互屏蔽多肽的捕获效率，获得特异性离子流响应；

(iiii) 特异性识别电中性且正负电荷分离多肽的“电性初筛纳米孔道”；

对于电中性且其中正负电荷分离的多肽，通过增强 Aerolysin 纳米孔道入口处的电势梯度，即将将 Aerolysin 纳米孔道中的 Aerolysin 蛋白变为突变型 S280Q/N/A、T284Q/N/A 或 G214Q/N/A，调控于非线性的电场强度，从而利用介电泳力驱动单个电中性且正负电荷分离多肽进入孔道。

15. 根据权利要求 5 所述一种采用 Aerolysin 纳米孔道蛋白质/多肽测序方法，其特征在于，所述步骤（5）中 6 类正交识别纳米孔道构建方法如下所示：

（I）基于静电相互作用，在孔道内现有电流传感区域中引入带电荷氨基酸，即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 T218K/R/H/D/E、S278K/R/H/D/E、S276K/R/H/D/E、T274K/R/H/D/E 或 A224Q/N/D/E/R/K/H；带电荷氨基酸的引入会增强孔道与待测氨基酸之间的氢键、盐键和阳离子- π 相互作用，从而实现对第一类氨基酸的识别，包括但不限于 H、R、K、E、D、Q、N、W；

（II）基于氢键及亲水作用，调控孔道内电流传感区域的电势梯度，即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 T218N/Q、Q212R/K/H、D209S/T、S276Q/N、D222G/A/S 或 A224E/D，用以加快带电荷氨基酸通过该区域的速度并延长极性不带电氨基酸在此区域内的停留时间，从而实现对第二类氨基酸的识别，包括但不限于 Q、N、Y、T、S、C、G、H；其中，组氨酸 H 中 R 基的 pKa 约 7，可通过微调 pH 使其不带电，从而基于其与纳米孔道关键传感区域的氢键相互作用，使其在该特异性纳米孔道中实现特征区分；

（III）基于范德华相互作用，调控孔道整体的电势分布和孔道立体结构分布，引入疏水氨基酸域，即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 R220S/T/A、D222G/A、S236I/L/V、G270I/L、T232I/L/V、T274G/A/I/L 或 K238F/Y/W，使得电流传感区域从野生型孔道的静电敏感域转移到突变型孔道的疏水域中，从而通过相互作用延长特定氨基酸在该区域的停留时间，获得特征离子流信号，从而拟实现对第三类氨基酸的识别，包括但不限于 I、L、M、V、P、A、C、G；

(IV) 基于部分氨基酸大 π 键相互作用, 在调控孔道立体结构和电势分布的基础上, 重构 Aerolysin 纳米孔道内的电流传感区域组成, 构建以正电荷氨基酸及输水氨基酸为主导的灵敏区域, 即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 D222W/H/F/Y、S276F/Y、A224K/R/W、S272W/H 或 T274W/H/F/Y, 用以增强该灵敏区域与特定氨基酸的 π - π 相互作用、阳离子- π 键相互作用、 π - π 相互作用, 从而实现对第四类氨基酸的识别, 包括但不限于 W、P、F、Y、H、I、L、V;

(V) 基于大位阻效应, 进一步减小孔道内电流传感区域的限域空间, 增大该区域的位阻, 即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 S276F/Y/I/L、S278F/Y/I/L/P、T274W/P、S236W 或 K238G/W/I/L/F/Y/P, 从而延长所有氨基酸通过该区域的时间, 增强小体积氨基酸的离子流信号电流幅值, 并使得大体积氨基酸造成近乎全阻断离子流台阶, 从而特异性区分氨基酸体积, 实现对第五类小体积氨基酸的识别, 包括但不限于 A、C、G、S、T、V;

(VI) 基于小位阻效应, 调控孔道内的立体结构, 增大电流关键区域的尺寸, 即将 Aerolysin 纳米孔道中的 Aerolysin 蛋白改为突变型 T218G/A、S276G/A、S278G/A、T274G/A、N226D/E 或 Q268S/T/G/A, 并基于多肽整体的带电荷类型减小纳米孔道内的电渗流, 从而进一步减小小体积氨基酸的电流响应, 增强大体积氨基酸的电流差异, 实现第六类大体积氨基酸的识别, 包括但不限于 W、H、I、K、R、Y。

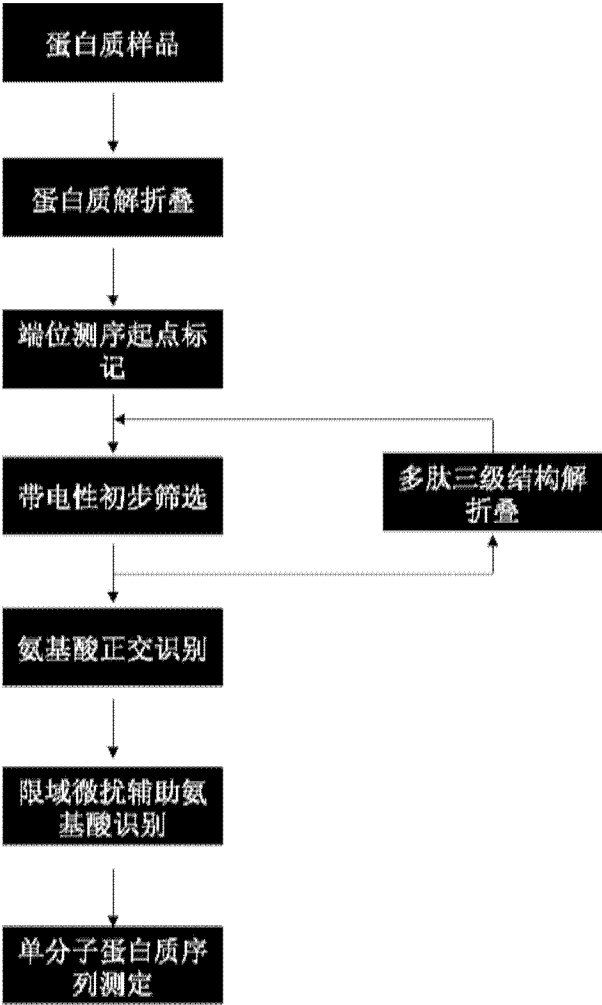


图 1

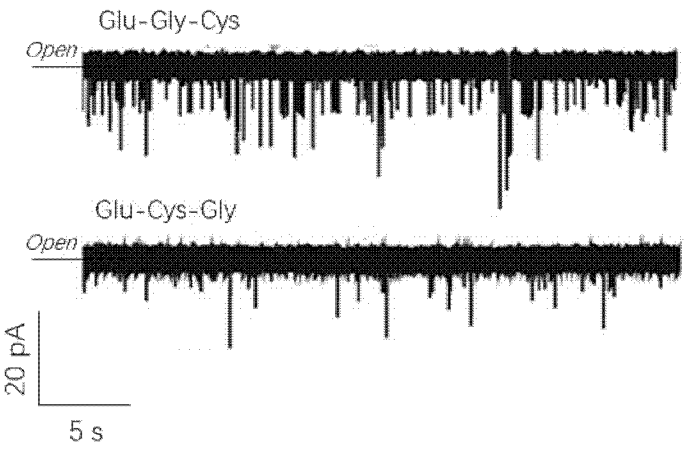


图 2



图 3

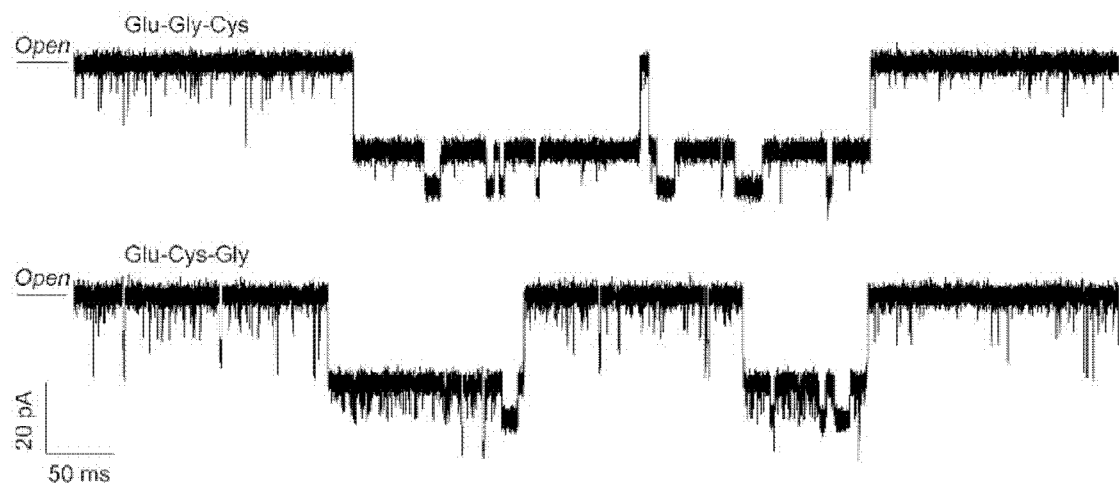


图 4

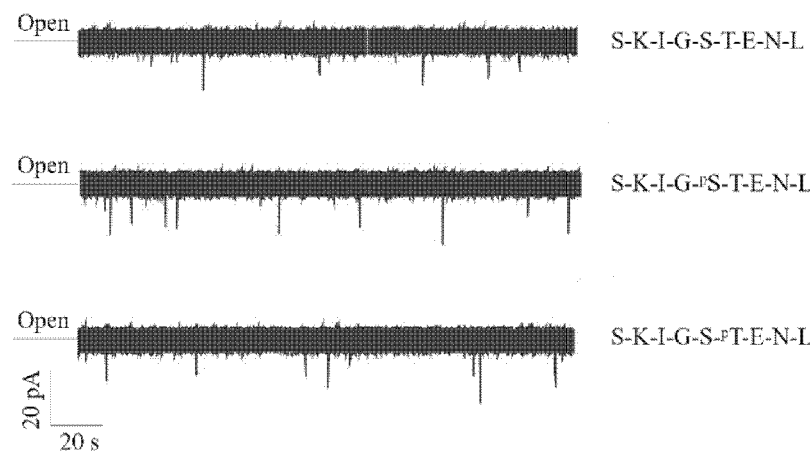


图 5

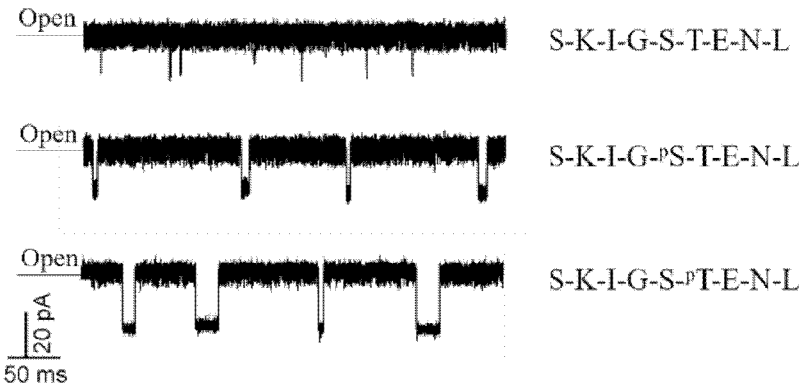


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/094480

A. CLASSIFICATION OF SUBJECT MATTER

C07K 1/113(2006.01)i; C07K 1/12(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, SIPOABS, DWPI, CNTXT, WOTXT, EPTXT, USTXT, CNKI, 百度搜索, WEB OF SCIENCE, PubMed: 纳米孔 (道), 蛋白质, 多肽, 测序, 解折叠, 端位, 起点标记, 带电性初步筛选, 正交识别, 单分子, 限域微扰, Aerolysin, nanopores, protein, polypeptide, sequencing, unfolding

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 112480204 A (NANJING UNIVERSITY) 12 March 2021 (2021-03-12) see claims 1-15	1-15
Y	WO 2020055246 A1 (UNIV. GRONINGEN) 19 March 2020 (2020-03-19) see claims 1-20, description page 1 lines 8-10, page 11 lines 20-22, page 12 lines 17-20	1-4
A	CN 109890980 A (UNIVERSITY OF GRONINGEN) 14 June 2019 (2019-06-14) see entire document	1-15
A	CN 103820311 A (TSINGHUA UNIVERSITY) 28 May 2014 (2014-05-28) see entire document	1-15
A	CN 104685361 A (TECHNISCHE UNIVERSITEIT DELFT) 03 June 2015 (2015-06-03) see entire document	1-15
A	CN 109295187 A (NANJING UNIVERSITY) 01 February 2019 (2019-02-01) see entire document	1-15
A	CN 108449941 A (F. HOFFMANN-LA ROCHE AG.) 24 August 2018 (2018-08-24) see entire document	1-15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

03 August 2021

Date of mailing of the international search report

01 September 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/094480

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Hadjer Ouldali et al. "Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore" <i>Nat Biotechnol.</i> , Vol. 38, No. 2, 28 February 2020 (2020-02-28), pages 176-181, see abstract, overview	1-4
A	Fabien Piguet et al. "Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore" <i>Nat Commun.</i> , Vol. 9, No. 1, 06 March 2018 (2018-03-06), pp. 1-13	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/094480

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	112480204	A	12 March 2021	None			
WO	2020055246	A1	19 March 2020	CA	3111488	A1	19 March 2020
				AU	2019337340	A1	25 March 2021
CN	109890980	A	14 June 2019	EP	3485029	B1	11 November 2020
				US	2019292235	A1	26 September 2019
				BR	112019000680	A2	24 April 2019
				JP	2019533637	A	21 November 2019
				CA	3030704	A1	18 January 2018
				WO	2018012963	A1	18 January 2018
				EP	3485029	A1	22 May 2019
				AU	2017295442	A1	07 February 2019
				AU	2017295442	A2	21 February 2019
CN	103820311	A	28 May 2014	CN	103820311	B	24 February 2016
CN	104685361	A	03 June 2015	EP	2872898	B1	07 December 2016
				CN	104685361	B	10 May 2017
				US	2015185199	A1	02 July 2015
				WO	2014014347	A1	23 January 2014
				JP	2015524561	A	24 August 2015
				CA	2879288	A1	23 January 2014
				ES	2613745	T3	25 May 2017
				NL	2009191	C	20 January 2014
				NL	2009191	C2	20 January 2014
				EP	2872898	A1	20 May 2015
CN	109295187	A	01 February 2019	CN	109295187	B	09 February 2021
CN	108449941	A	24 August 2018	WO	2017050722	A1	30 March 2017
				US	2018362594	A1	20 December 2018
				JP	2018531002	A	25 October 2018
				AU	2016326867	B2	13 September 2018
				CA	2998970	A1	30 March 2017
				EP	3353194	A1	01 August 2018
				AU	2016326867	A1	08 March 2018
				US	10752658	B2	25 August 2020
				US	2020392191	A1	17 December 2020

国际检索报告

国际申请号

PCT/CN2021/094480

A. 主题的分类 C07K 1/113 (2006.01) i; C07K 1/12 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																										
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) C07K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS, SIPOABS, DWPI, CNTXT, WOTXT, EPTXT, USTXT, CNKI, 百度学术搜索, WEB OF SCIENCE, PubMed: 纳米孔 (道), 蛋白质, 多肽, 测序, 解折叠, 端位, 起点标记, 带电性初步筛选, 正交识别, 单分子, 限域微扰, Aerolysin, nanopores, protein, polypeptide, sequencing, unfolding																										
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 112480204 A (南京大学) 2021年 3月 12日 (2021 - 03 - 12) 参见权利要求1-15</td> <td>1-15</td> </tr> <tr> <td>Y</td> <td>WO 2020055246 A1 (UNIV GRONINGEN) 2020年 3月 19日 (2020 - 03 - 19) 参见权利要求1-20, 说明书第1页8-10行, 第11页第20-22行, 第12页第17-20行</td> <td>1-4</td> </tr> <tr> <td>A</td> <td>CN 109890980 A (格罗宁根大学) 2019年 6月 14日 (2019 - 06 - 14) 参见全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 103820311 A (清华大学) 2014年 5月 28日 (2014 - 05 - 28) 参见全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 104685361 A (代尔夫特科技大学) 2015年 6月 3日 (2015 - 06 - 03) 参见全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 109295187 A (南京大学) 2019年 2月 1日 (2019 - 02 - 01) 参见全文</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>CN 108449941 A (豪夫迈 罗氏有限公司) 2018年 8月 24日 (2018 - 08 - 24) 参见全文</td> <td>1-15</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 112480204 A (南京大学) 2021年 3月 12日 (2021 - 03 - 12) 参见权利要求1-15	1-15	Y	WO 2020055246 A1 (UNIV GRONINGEN) 2020年 3月 19日 (2020 - 03 - 19) 参见权利要求1-20, 说明书第1页8-10行, 第11页第20-22行, 第12页第17-20行	1-4	A	CN 109890980 A (格罗宁根大学) 2019年 6月 14日 (2019 - 06 - 14) 参见全文	1-15	A	CN 103820311 A (清华大学) 2014年 5月 28日 (2014 - 05 - 28) 参见全文	1-15	A	CN 104685361 A (代尔夫特科技大学) 2015年 6月 3日 (2015 - 06 - 03) 参见全文	1-15	A	CN 109295187 A (南京大学) 2019年 2月 1日 (2019 - 02 - 01) 参见全文	1-15	A	CN 108449941 A (豪夫迈 罗氏有限公司) 2018年 8月 24日 (2018 - 08 - 24) 参见全文	1-15
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
PX	CN 112480204 A (南京大学) 2021年 3月 12日 (2021 - 03 - 12) 参见权利要求1-15	1-15																								
Y	WO 2020055246 A1 (UNIV GRONINGEN) 2020年 3月 19日 (2020 - 03 - 19) 参见权利要求1-20, 说明书第1页8-10行, 第11页第20-22行, 第12页第17-20行	1-4																								
A	CN 109890980 A (格罗宁根大学) 2019年 6月 14日 (2019 - 06 - 14) 参见全文	1-15																								
A	CN 103820311 A (清华大学) 2014年 5月 28日 (2014 - 05 - 28) 参见全文	1-15																								
A	CN 104685361 A (代尔夫特科技大学) 2015年 6月 3日 (2015 - 06 - 03) 参见全文	1-15																								
A	CN 109295187 A (南京大学) 2019年 2月 1日 (2019 - 02 - 01) 参见全文	1-15																								
A	CN 108449941 A (豪夫迈 罗氏有限公司) 2018年 8月 24日 (2018 - 08 - 24) 参见全文	1-15																								
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																										
<table border="0"> <tr> <td style="vertical-align: top;"> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td style="vertical-align: top;"> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																						
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																									
国际检索实际完成的日期		国际检索报告邮寄日期																								
2021年 8月 3日		2021年 9月 1日																								
ISA/CN的名称和邮寄地址		受权官员																								
中国国家知识产权局 (ISA/CN) 中国 北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		魏春宝 电话号码 62411028																								

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	Hadjer Ouldali等. "Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore" Nat Biotechnol., 第38卷, 第2期, 2020年 2月 28日 (2020 - 02 - 28), 第176-181页 参见摘要, 概述	1-4
A	Fabien Piguet等. "Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore" Nat Commun., 第9卷, 第1期, 2018年 3月 6日 (2018 - 03 - 06), 第1-13页	1-15

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/094480

检索报告引用的专利文件			公布日 (年/月/日)	同族专利		公布日 (年/月/日)
CN	112480204	A	2021年 3月 12日	无		
WO	2020055246	A1	2020年 3月 19日	CA	3111488	A1 2020年 3月 19日
				AU	2019337340	A1 2021年 3月 25日
CN	109890980	A	2019年 6月 14日	EP	3485029	B1 2020年 11月 11日
				US	2019292235	A1 2019年 9月 26日
				BR	112019000680	A2 2019年 4月 24日
				JP	2019533637	A 2019年 11月 21日
				CA	3030704	A1 2018年 1月 18日
				WO	2018012963	A1 2018年 1月 18日
				EP	3485029	A1 2019年 5月 22日
				AU	2017295442	A1 2019年 2月 7日
				AU	2017295442	A2 2019年 2月 21日
CN	103820311	A	2014年 5月 28日	CN	103820311	B 2016年 2月 24日
CN	104685361	A	2015年 6月 3日	EP	2872898	B1 2016年 12月 7日
				CN	104685361	B 2017年 5月 10日
				US	2015185199	A1 2015年 7月 2日
				WO	2014014347	A1 2014年 1月 23日
				JP	2015524561	A 2015年 8月 24日
				CA	2879288	A1 2014年 1月 23日
				ES	2613745	T3 2017年 5月 25日
				NL	2009191	C 2014年 1月 20日
				NL	2009191	C2 2014年 1月 20日
				EP	2872898	A1 2015年 5月 20日
CN	109295187	A	2019年 2月 1日	CN	109295187	B 2021年 2月 9日
CN	108449941	A	2018年 8月 24日	WO	2017050722	A1 2017年 3月 30日
				US	2018362594	A1 2018年 12月 20日
				JP	2018531002	A 2018年 10月 25日
				AU	2016326867	B2 2018年 9月 13日
				CA	2998970	A1 2017年 3月 30日
				EP	3353194	A1 2018年 8月 1日
				AU	2016326867	A1 2018年 3月 8日
				US	10752658	B2 2020年 8月 25日
				US	2020392191	A1 2020年 12月 17日