



(12)发明专利



(10)授权公告号 CN 105026002 B

(45)授权公告日 2017.05.24

(21)申请号 201480008025.4

(22)申请日 2014.02.10

(65)同一申请的已公布的文献号
申请公布号 CN 105026002 A

(43)申请公布日 2015.11.04

(30)优先权数据
61/762,671 2013.02.08 US

(85)PCT国际申请进入国家阶段日
2015.08.07

(86)PCT国际申请的申请数据
PCT/US2014/015574 2014.02.10

(87)PCT国际申请的公布数据
W02014/124365 EN 2014.08.14

(73)专利权人 康奈尔大学
地址 美国纽约州
专利权人 路易斯安那州州立大学及农业机械学院管理委员会
北卡罗来纳大学查佩尔山分校
美国东北大学

(72)发明人 S·A·索普 F·巴拉尼
G·格日尔斯 R·L·麦卡利
C·J·米奇尼 D·摩尔多万
M·C·莫菲 D·E·尼科图浦洛斯
S·帕克 E·J·波德拉哈-莫菲

(74)专利代理机构 北京万慧达知识产权代理有限公司 11111

代理人 白华胜 代峰

(51)Int.Cl.
C12M 1/40(2006.01)
C12M 1/34(2006.01)
C12M 1/00(2006.01)
G01N 33/68(2006.01)
C12Q 1/68(2006.01)

(56)对比文件
EP 1047107 A2,2000.10.25,全文.
US 2007/0190546 A1,2007.08.16,全文.
US 2012/0129716 A1,2012.05.24,说明书第8,14,16,21,28,32,35-37,62,81,91,120,168,220,298段.
W0 2013/012440 A2,2013.01.24,全文.
Uba et al.Nanogap electrical detection of science molecules translocating through a nanochannel with transverse nanoelectrodes and funnels populated with an array of nanopillars.《15th International Conference on Miniaturized Systems for Chemistry and Life Sciences》.2011,第401页第2-3段,第402页第2,5段,第403页及图2C,3B.

审查员 王斌

权利要求书8页 说明书59页
序列表6页 附图48页

(54)发明名称
生物分子处理平台及其用途

(57)摘要

本发明涉及一种包括生物分子处理器的装置。每个生物分子处理器具有一个或多个由固体基底界定的生物反应器室;在每个生物反应器内的支撑结构;裂解酶,所述裂解酶被固定于所述支撑结构并且操作性地定位于所述生物反应器室内以裂解所述裂解酶操作性地接合的生物聚

合物分子的单体或多聚体单元;以及在所述固体基底中形成的一个或多个飞行时间通道,所述飞行时间通道流体地联接到所述一个或多个生物反应器室。所述飞行时间通道各自具有两个或更多个传感器,这些传感器包括至少(i)邻近于所述飞行时间通道的输入端接触所述通道的第一传感器,及(ii)邻近于所述飞行时间通道的输出端接触所述通道的第二传感器。本发明还涉及使用所述装置对生物分子进行测序和鉴别的方法。

1. 一种装置,其包括:

生物分子处理器,每个生物分子处理器包括:

一个或多个由固体基底界定的生物反应器室;

在每个生物反应器室内并附接到所述固体基底的支撑结构;

裂解酶,所述裂解酶被固定于所述支撑结构并且操作性地定位于所述生物反应器室内以裂解所述裂解酶操作性地接合的生物聚合物分子的单体或多聚体单元;及

在所述固体基底中形成并且流体地联接到所述一个或多个生物反应器室的一个或多个飞行时间通道,所述一个或多个飞行时间通道各自具有输入端和输出端,其中所述一个或多个飞行时间通道各自包括两个或更多个传感器,所述传感器包括至少(i)邻近于所述一个或多个飞行时间通道的输入端接触所述一个或多个飞行时间通道的第一传感器,及(ii)邻近于所述一个或多个飞行时间通道的输出端接触所述一个或多个飞行时间通道的第二传感器。

2. 如权利要求1所述的装置,其还包括:

电场发生器,所述电场发生器操作性地定位以在所述一个或多个生物反应器室中并且沿所述一个或多个飞行时间通道的长度产生电场。

3. 如权利要求1所述的装置,其中所述一个或多个飞行时间通道是小于或等于50nm宽并且小于或等于50nm深。

4. 如权利要求1所述的装置,其中所述一个或多个飞行时间通道是5 μ m-250 μ m长。

5. 如权利要求1所述的装置,其中所述一个或多个生物反应器室是100-1000nm宽。

6. 如权利要求1所述的装置,其中在所述一个或多个生物反应器室内的所述支撑结构是50-900nm宽。

7. 如权利要求1所述的装置,其中所述固体基底是由聚合材料制成。

8. 如权利要求1所述的装置,其中所述裂解酶是核酸外切酶或核糖核酸外切酶。

9. 如权利要求8所述的装置,其中所述裂解酶是进行性核酸外切酶或核糖核酸外切酶。

10. 如权利要求1所述的装置,其中所述裂解酶是蛋白酶。

11. 如权利要求1所述的装置,其中所述两个或更多个传感器各自包括一对感测电极,所述感测电极由贵金属纳米线、分节段纳米线或纳米粒子组成。

12. 如权利要求11所述的装置,其中每一对感测电极都具有供其接触所述飞行时间通道的窄截面。

13. 如权利要求1所述的装置,其中所述一个或多个飞行时间通道包括三个或更多个传感器。

14. 如权利要求1所述的装置,其中所述生物分子处理器还包括:

由所述固体基底的壁界定的进料通道,所述进料通道的长度从邻近于所述固体基底的表面的输入端延伸到邻近于所述一个或多个生物反应器室的输出端。

15. 如权利要求14所述的装置,其还包括:

邻近于所述进料通道的输入端接触所述进料通道的传感器。

16. 如权利要求15所述的装置,其中所述进料通道传感器包括一对感测电极。

17. 如权利要求14所述的装置,其中所述进料通道在所述输入端与所述输出端之间包括一个或多个熵陷阱。

18. 如权利要求14所述的装置,其中所述进料通道是小于或等于100nm宽并且小于或等于100nm深。

19. 如权利要求1所述的装置,其中一个或多个所述飞行时间通道包括电中性亲水性壁表面、带电的亲水性壁表面或电中性疏水性壁表面。

20. 如权利要求1所述的装置,其中所述装置包括10,000个或更多个生物分子处理器。

21. 如权利要求1所述的装置,其中所述装置包括100,000个或更多个生物分子处理器。

22. 一种用于核酸分子测序的方法,其包括:

提供如权利要求1所述的装置;

将包含一个或多个核酸分子的样品在一定条件下进料至所述装置的所述生物分子处理器中,所述条件有效使所述生物反应器室内固定的裂解酶接合所述样品中的所述一个或多个核酸分子并将所述一个或多个核酸分子裂解成单体核苷酸以个别地进入所述一个或多个飞行时间通道的输入端;

跨所述一个或多个生物反应器室并且沿所述一个或多个飞行时间通道的长度施加电场以将所述裂解的单体核苷酸输送通过所述一个或多个飞行时间通道;

基于所述施加,在所述裂解的单体核苷酸穿过所述一个或多个飞行时间通道中的至少第一传感器和第二传感器时,对其进行检测;

基于所述检测,测量每个裂解的单体核苷酸穿过所述一个或多个飞行时间通道中的至少第一传感器和第二传感器所花费的时间;以及

基于所述测量,鉴别所述单体核苷酸。

23. 如权利要求22所述的方法,其还包括:

基于所述鉴别,获得所述样品中所述一个或多个核酸分子的至少一部分核苷酸序列。

24. 如权利要求22所述的方法,其还包括:

重复所述施加、所述检测、所述测量、所述鉴别,以获得所述样品中所述一个或多个核酸分子中超过一个的至少一部分核苷酸序列。

25. 如权利要求24所述的方法,其中进行所述重复以获得至少10,000个核酸分子的核苷酸序列或其部分。

26. 如权利要求24所述的方法,其中进行所述重复以获得至少100,000个核酸分子的核苷酸序列或其部分。

27. 如权利要求22所述的方法,其中所述测量还包括:

测量每个裂解的单体核苷酸穿过所述一个或多个飞行时间通道中第一传感器或第二传感器中的至少一个时的电力峰值振幅。

28. 如权利要求22所述的方法,其中所述一个或多个核酸分子选自由以下各物组成的群组:双链脱氧核糖核酸分子(DNA)、单链DNA分子、包含一个或多个甲基化核苷酸碱基的DNA分子、包含一个或多个经过修饰或受损的核苷酸碱基的DNA分子、核糖核酸(RNA)分子、包含一个或多个经过修饰或受损的核苷酸碱基的RNA分子,以及RNA/DNA杂交分子。

29. 如权利要求28所述的方法,其中所述一个或多个核酸分子包含RNA分子,所述RNA分子选自由以下各物组成的群组:长链非编码RNA(lncRNA)、核糖体RNA(rRNA)、核内小RNA(snoRNA)、微RNA(miRNA)、转运RNA(tRNA)以及小干扰RNA(siRNA)。

30. 如权利要求28所述的方法,其中所述一个或多个核酸分子包含双链DNA分子并且所

述双链DNA分子的两条链中的至少一部分单体核苷酸得到鉴别。

31. 如权利要求28所述的方法,其中所述一个或多个核酸分子包括含一个或多个核苷酸碱基插入、缺失、易位、突变和/或受损核苷酸碱基的DNA分子,并且所述一个或多个核苷酸碱基插入、缺失、易位、突变和/或受损核苷酸碱基是基于所述测量进行鉴别。

32. 如权利要求31所述的方法,其中具有一个或多个核苷酸碱基插入、缺失、易位、突变和/或受损核苷酸碱基的低丰度核酸分子不同于所述样品中与所述低丰度核酸分子具有类似核苷酸序列但不含所述一个或多个核苷酸碱基插入、缺失、易位、突变和/或受损碱基的100到10,000倍过量的核酸分子。

33. 如权利要求28所述的方法,其中所述一个或多个核酸分子包括含一个或多个甲基化或修饰的核苷酸碱基的DNA或RNA分子,并且所述一个或多个甲基化或修饰的核苷酸碱基是基于所述测量进行鉴别。

34. 如权利要求33所述的方法,其中具有一个或多个甲基化或修饰的核苷酸碱基的低丰度核酸分子不同于所述样品中与所述低丰度核酸分子具有类似核苷酸序列且具有未甲基化或未修饰的核苷酸碱基的100到10,000倍过量的核酸分子。

35. 如权利要求32所述的方法,其中具有一个或多个未甲基化或未修饰的核苷酸碱基的低丰度核酸分子不同于所述样品中与所述低丰度核酸分子具有类似核苷酸序列且具有甲基化或修饰的核苷酸碱基的100到10,000倍过量的核酸分子。

36. 如权利要求22所述的方法,其还包括:

在所述进料之前,从组织、细胞、血清、血浆、血液或外来体分离所述一个或多个核酸分子。

37. 如权利要求22所述的方法,其中在进行所述方法时,鉴别出所述一个或多个核酸分子中的200到50,000个单体核苷酸。

38. 如权利要求22所述的方法,其还包括:

基于所述鉴别,诊断或预测疾病状态并区分基因型或发展阶段。

39. 如权利要求22所述的方法,其中所述裂解酶是核酸外切酶或核糖核酸外切酶。

40. 如权利要求39所述的方法,其中所述裂解酶是进行性核酸外切酶或进行性核糖核酸外切酶。

41. 如权利要求39所述的方法,其中所述裂解酶选自 λ 核酸外切酶、核酸外切酶I、核酸外切酶III、T7核酸外切酶、XRN-1、RNA酶II及外来体复合物,或是所述酶的组分。

42. 如权利要求22所述的方法,其中所述生物分子处理器还包括:

由所述固体基底的壁界定的进料通道,所述进料通道的长度从邻近于所述固体基底的表面的输入端延伸到邻近于所述一个或多个生物反应器室的输出端,所述方法还包括:

将包含所述一个或多个核酸分子的样品引入所述进料通道中,以及

沿所述进料通道的长度施加电场以在所述进料通道内延长所述一个或多个核酸分子。

43. 如权利要求42所述的方法,其中所述进料通道在所述输入端与所述输出端之间还包括一个或多个熵陷阱以对所述一个或多个核酸分子定向。

44. 如权利要求42所述的方法,其中所述装置还包括邻近于所述进料通道的输入端接触所述进料通道的传感器,所述方法还包括:

在所述一个或多个核酸分子横过所述进料通道传感器时,对其进行检测,及

基于所述检测的持续时间,测定所述一个或多个核酸分子的长度。

45. 如权利要求42所述的方法,其中所述进料通道传感器包括一对感测电极。

46. 一种用于鉴别一种或多种蛋白质或多肽的方法,其包括:

提供如权利要求1所述的装置;

将包含一种或多种蛋白质或多肽的样品在一定条件下进料至所述装置的所述生物分子处理器中,所述条件有效使所述生物反应器室内固定的裂解酶接合所述样品中的所述一种或多种蛋白质或多肽并将所述一种或多种蛋白质或多肽裂解成较小的肽片段以进入所述一个或多个飞行时间通道的输入端;

跨所述一个或多个生物反应器室并且沿所述一个或多个飞行时间通道的长度施加电场以将所述裂解的肽片段输送通过所述一个或多个飞行时间通道;

基于所述施加,在所述肽片段穿过所述一个或多个飞行时间通道中的至少第一传感器和第二传感器时,对其进行检测;

基于所述检测,测量(i)每个肽片段穿过所述一个或多个飞行时间通道的至少第一传感器和第二传感器所花费的时间,和/或(ii)在每个肽片段穿过所述一个或多个飞行时间通道中第一传感器或第二传感器中的至少一个时的电力峰值振幅;及

基于对所述肽片段的所述测量,鉴别所述样品中所述一种或多种蛋白质或多肽的至少一部分。

47. 如权利要求46所述的方法,其还包括:

将从所述样品中的所述一种或多种蛋白质或多肽裂解的肽片段的测量值(i)和(ii)与通过已知蛋白质或多肽裂解产生的肽片段的相应测量值相比较。

48. 如权利要求46所述的方法,其还包括:

在所述进料之前,富集或分离来自生物样品的所述一种或多种蛋白质或多肽。

49. 如权利要求48所述的方法,其中所述富集或分离是使用亲和纯化、配体结合或基于抗体的富集进行的。

50. 如权利要求46所述的方法,其中所述一种或多种蛋白质或多肽选自由以下各物组成的群组:G蛋白偶联受体、核受体、电压门控离子通道、配体门控离子通道、受体酪氨酸激酶、生长因子、蛋白酶、序列特异性蛋白酶、磷酸酶、蛋白激酶、生物活性脂质、细胞因子、趋化因子、泛素连接酶、病毒调控剂、细胞分裂蛋白、支架蛋白、DNA修复蛋白、细菌核糖体、组蛋白脱乙酰酶、细胞凋亡调控剂、伴侣蛋白、丝氨酸/苏氨酸蛋白质激酶、周期蛋白依赖性激酶、生长因子受体、蛋白酶体、信号传导蛋白复合物、蛋白质/核酸转运蛋白及病毒衣壳。

51. 如权利要求46所述的方法,其还包括:

在所述进料之前,分离或富集来自一个或多个亚细胞区室的所述一种或多种蛋白质或多肽。

52. 如权利要求46所述的方法,其中所述一种或多种蛋白质或多肽包含一个或多个修饰的氨基酸残基,并且所述测量鉴别出所述一个或多个氨基酸修饰。

53. 如权利要求52所述的方法,其中所述氨基酸修饰选自甲基化、乙酰化及磷酸化。

54. 如权利要求46所述的方法,其中所述裂解酶是蛋白酶。

55. 如权利要求54所述的方法,其中所述裂解酶选自由以下各物组成的群组:蛋白酶K、金黄色葡萄球菌V-8蛋白酶、胰蛋白酶、赖氨酸蛋白酶、天冬氨酸蛋白酶、谷氨酸蛋白酶及胰

凝乳蛋白酶。

56. 一种用于鉴别样品中多个核酸分子中的一个或多个的方法, 其包括:

提供含有一个或多个核酸分子的样品, 所述核酸分子各自与其对应的互补核酸分子杂交, 其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分和适于形成发夹的单链3' 适配体部分;

使所述样品经受有效使所述一个或多个核酸分子和其互补序列的所述3' 适配体部分形成发夹的条件;

提供聚合酶;

将所述一个或多个3' 发夹状核酸分子及其互补序列与所述聚合酶混合以形成延伸混合物;

使所述一个或多个核酸分子及其互补序列从其对应的3' 发夹状区域延伸以形成一个或多个全长发夹状核酸分子;

将所述一个或多个全长发夹状核酸分子引入一个或多个生物分子处理器中, 所述生物分子处理器包括:

一个或多个由固体基底界定的生物反应器室;

在每个生物反应器室内并附接到所述固体基底的支撑结构;

裂解酶, 所述裂解酶被固定于所述支撑结构并且操作性地定位于所述生物反应器室内以裂解所述裂解酶操作性地接合的生物聚合物分子的单体或多聚体单元; 及

在所述固体基底中形成并且流体地联接到所述一个或多个生物反应器室的一个或多个飞行时间通道, 所述一个或多个飞行时间通道各自具有输入端和输出端, 其中所述一个或多个飞行时间通道各自包括两个或更多个传感器, 所述传感器包括至少 (i) 邻近于所述一个或多个飞行时间通道的输入端接触所述一个或多个飞行时间通道的第一传感器, 及 (ii) 邻近于所述一个或多个飞行时间通道的输出端接触所述一个或多个飞行时间通道的第二传感器; 及

对所述一个或多个全长发夹状核酸分子的至少一部分进行测序以鉴别所述样品中的一个或多个核酸分子。

57. 如权利要求56所述的方法, 其中在所述延伸期间, 所述聚合酶将所述一个或多个核酸分子从其对应的互补核酸分子解离。

58. 如权利要求57所述的方法, 其中所述聚合酶包含5' 到3' 核酸外切酶活性或链置换活性。

59. 如权利要求56所述的方法, 其还包括:

在所述延伸之前, 使所述一个或多个3' 发夹状核酸分子与其互补序列解链。

60. 如权利要求56所述的方法, 其中所述一个或多个全长发夹状核酸分子的5' 端被磷酸化。

61. 如权利要求56所述的方法, 其中所述提供样品包括:

提供包含一个或多个核酸分子及其互补序列的样品, 及

将所述适配体部分附接到所述一个或多个核酸分子及其互补序列的所述靶特异性部分上。

62. 如权利要求61所述的方法, 其中所述适配体部分, 或其互补序列是使用一种或多种

酶附接到所述一个或多个核酸分子或其互补序列的所述靶特异性部分上,所述酶选自连接酶、聚合酶、重组酶、末端转移酶、核酸内切酶、DNA修复酶及反转录酶组成的群组。

63. 如权利要求56所述的方法,其中所述适配体部分另外包含一个或多个独特的核苷酸序列标签。

64. 如权利要求56所述的方法,其还包括:

在所述测序之前,将迁移率调节物附接到所述一个或多个全长发夹状核酸分子的发夹状部分上。

65. 一种用于鉴别样品中的多个核酸分子中的一个或多个的方法,其包括:

提供含有一个或多个核酸分子的样品,所述核酸分子各自与其对应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分、5' 适配体部分及3' 适配体部分,其中所述5' 适配体部分和所述3' 适配体部分各自适于形成发夹;

使所述一个或多个核酸分子与其互补序列解链以形成一个或多个单链核酸分子及其互补序列的集合,其中所述一个或多个单链核酸分子及其互补序列各自的5' 适配体部分及3' 适配体部分的发夹序列形成发夹;

提供包含多个固定的捕获寡核苷酸探针的固体支撑物,每个捕获寡核苷酸探针包含与所述一个或多个核酸分子或其互补序列的靶特异性部分的区域互补的核苷酸序列;

使一个或多个所述解链的单链核酸分子及其互补序列与所述固体支撑物上固定的其互补捕获寡核苷酸杂交;

提供聚合酶;

使所述聚合酶与含有一个或多个所述杂交的核酸分子及其互补序列的所述固体支撑物接触;

使所述一个或多个核酸分子或其互补序列从其对应的3' 发夹状区域延伸以形成一个或多个全长发夹状核酸分子,由此使所述一个或多个核酸分子从所述捕获寡核苷酸和从所述固体支撑物释放;

将所述一个或多个全长发夹状核酸分子引入一个或多个生物分子处理器中,所述生物分子处理器包括:

一个或多个由固体基底界定的生物反应器室;

在每个生物反应器室内并附接到所述固体基底的支撑结构;

裂解酶,所述裂解酶被固定于所述支撑结构并且操作性地定位于所述生物反应器室内以裂解所述裂解酶操作性地接合的生物聚合物分子的单体或多聚体单元;及

在所述固体基底中形成并且流体地联接到所述一个或多个生物反应器室的一个或多个飞行时间通道,所述一个或多个飞行时间通道各自具有输入端和输出端,其中所述一个或多个飞行时间通道各自包括两个或更多个传感器,所述传感器包括至少(i)邻近于所述一个或多个飞行时间通道的输入端接触所述一个或多个飞行时间通道的第一传感器,及(ii)邻近于所述一个或多个飞行时间通道的输出端接触所述一个或多个飞行时间通道的第二传感器;及

对所述一个或多个全长发夹状核酸分子的至少一部分进行测序以鉴别所述样品中的一个或多个核酸分子。

66. 如权利要求65所述的方法,其中所述提供样品包括:

提供包含一个或多个核酸分子及其互补序列的样品,及

将所述适配体部分附接到所述一个或多个核酸分子及其互补序列的所述靶特异性部分上。

67. 如权利要求66所述的方法,其中所述适配体部分,或其互补序列是使用一种或多种酶附接到所述一个或多个核酸分子或其互补序列的所述靶特异性部分上,所述酶选自自由连接酶、聚合酶、重组酶、末端转移酶、核酸内切酶、DNA修复酶及反转录酶组成的群组。

68. 如权利要求65所述的方法,其中所述适配体部分还包含一个或多个独特的核苷酸序列标签。

69. 如权利要求65所述的方法,其中所述全长发夹状靶核酸分子的5'端被磷酸化。

70. 如权利要求65所述的方法,其还包括:

在所述测序之前,将迁移率调节物附接到所述全长发夹状核酸分子的发夹状部分上。

71. 一种用于鉴别样品中的多个靶核酸分子中的一个或多个的方法,其包括:

提供含有一个或多个核酸分子的样品,所述核酸分子各自与其对应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分、5'适配体部分及包含单链均聚物重复序列的3'适配体部分;

使所述一个或多个核酸分子与其互补序列解链以形成一个或多个单链核酸分子及其互补序列的集合;

提供与所述一个或多个单链核酸分子及其互补序列的3'端互补的多个寡核苷酸引物;

提供聚合酶;

将所述一个或多个单链核酸分子及其互补序列的集合与所述聚合酶和所述引物混合以形成延伸混合物;

使所述寡核苷酸引物与其互补的所述一个或多个单链核酸分子及其互补序列的3'端杂交;

使所述杂交的引物延伸以形成一个或多个全长双链核酸分子;

将所述一个或多个全长发夹状核酸分子引入一个或多个生物分子处理器中,所述生物分子处理器包括:

一个或多个由固体基底界定的生物反应器室;

在每个生物反应器室内并附接到所述固体基底的支撑结构;

裂解酶,所述裂解酶被固定于所述支撑结构并且操作性地定位于所述生物反应器室内以裂解所述裂解酶操作性地接合的生物聚合物分子的单体或多聚体单元;及

在所述固体基底中形成并且流体地联接到所述一个或多个生物反应器室的一个或多个飞行时间通道,所述一个或多个飞行时间通道各自具有输入端和输出端,其中所述一个或多个飞行时间通道各自包括两个或更多个传感器,所述传感器包括至少(i)邻近于所述一个或多个飞行时间通道的输入端接触所述一个或多个飞行时间通道的第一传感器,及(ii)邻近于所述一个或多个飞行时间通道的输出端接触所述一个或多个飞行时间通道的第二传感器;及

对所述一个或多个全长双链核酸分子的至少一部分进行测序以鉴别所述样品中的一个或多个核酸分子。

72. 如权利要求71所述的方法,其中所述提供样品包括:

提供包含所述一个或多个核酸分子及其互补序列的样品；

将所述3' 适配体部分和所述5' 适配体部分附接到所述一个或多个核酸分子及其互补序列的所述靶特异性部分上；及

使所述一个或多个核酸分子及其互补序列中具有单链均聚物重复序列的所述3' 适配体部分延伸。

73. 如权利要求72所述的方法，其中所述附接和延伸是使用一种或多种酶进行的，所述酶选自自由连接酶、聚合酶、重组酶、末端转移酶、核酸内切酶、DNA修复酶及反转录酶组成的群组。

74. 如权利要求71所述的方法，其中所述适配体部分包含一个或多个独特的核苷酸序列标签。

75. 如权利要求71所述的方法，其中所述寡核苷酸引物包含3' 适配体特异性部分和3' 均聚物重复序列特异性部分。

76. 如权利要求71所述的方法，其中所述5' 适配体部分包含5' 磷酸酯部分。

77. 如权利要求71所述的方法，其还包括：

提供包含多个固定的捕获寡核苷酸探针的固体支撑物，每个捕获寡核苷酸探针包含与所述一个或多个核酸分子或其互补序列的靶特异性部分的区域互补的核苷酸序列，及

在所述混合之前，使所述一个或多个解链的单链核酸分子及其互补序列与所述固体支撑物上的固定的其互补捕获寡核苷酸杂交以实现靶特异性富集，其中在所述延伸之前或期间，所述一个或多个核酸分子及其互补序列从所述固体支撑物的所述捕获寡核苷酸解离。

78. 如权利要求77所述的方法，其中在所述延伸期间，所述聚合酶将所述一个或多个核酸分子及其互补序列从所述固体支撑物的所述捕获寡核苷酸解离。

79. 如权利要求78所述的方法，其中所述聚合酶具有5' 到3' 核酸外切酶活性或链置换活性。

80. 如权利要求77所述的方法，其还包括：

在所述杂交所述寡核苷酸引物之后并且在所述延伸之前，使所述一个或多个核酸分子及其互补序列与所述固体支撑物的所述捕获寡核苷酸解链。

81. 如权利要求71所述的方法，其还包括：

在所述测序之前，将迁移率调节物附接到所述一个或多个延伸的全长双链靶核酸分子的3' 端上。

生物分子处理平台及其用途

[0001] 本申请要求2013年2月8日提交的美国临时专利申请序列号61/762,671的权益,其在此以引用的方式整体并入。

[0002] 本发明是在国立卫生研究院授予的基金号NIH-R21-HG006278-01、国家科学基金会授予的NSF-EPS-0346411及NSF-EPS-0701491下由政府支持进行。政府在本发明中享有一定权利。

发明领域

[0003] 本发明涉及一种适用于核酸测序及蛋白质/多肽鉴别的装置和方法。

[0004] 发明背景

[0005] \$1,000基因组的益处已在文献中充分地记载(Kaiser,J.,“DNA Sequencing-A Plan to Capture Human Diversity in 1000 Genomes,”*Science* 319:395-395(2008); Kuehn,B.M.,“1000 Genomes Project Promises Closer Look at Variation in Human Genome,”*JAMA*300:2715-2715(2008);Mardis,E.,“Anticipating the\$1,000 Genome,”*Genome Biol.*7:112(2006);Metzker,M.L.,“Emerging Technologies in DNA Sequencing,”*Genome Res.*15:1767-1776(2005);Schloss,J.,“How to Get Genomes at One Ten-Thousandth the Cost,”*Nature Biotechnol.*26:1113-1115(2008))。一些重要影响包括;(i)个体化药物,其可参与更有效的疾病预防、改善诊断和预后以便经由序列变异的全基因组评价使适当的疗法与特定患者相匹配;(ii)理解全基因组复杂性;(iii)设计新的治疗剂;以及(iv)开发一种不管序列变化型式的体外诊断(IVD)的实际上的标准。

[0006] 存在用作多种疾病的有效生物标志物的很多不同的遗传变异,如自然突变、遗传突变、单核苷酸多态性(SNP)、甲基化模式(表观遗传)、基因表达、拷贝数变化、微卫星不稳定性等等。令人遗憾地,所有这些结构修饰需要独特的测定形式且因而由于进行每个分子测定所要求的专门设备和专门技术而难以在临床上实施(Thomas等人,“Biomedical Microelectromechanical Systems(BioMEMS)Using Electrophoresis for the Analysis of Genetic Mutations,”*Molecular Review Diagnostics* 2:429-447(2002))。可使用几乎不需要操作者专门技术的单一仪器揭露所有序列变异的存在/不存在的“标准”测定形式将扩展IVD的全部效用。在很多情况下,选出的外显子在基因组中的广泛再测序可提供具有所需的灵敏度而不管序列变异类型的必需的临床信息。

[0007] DNA测序的进展有希望标准化并开发非侵入性分子诊断以改善产前护理、移植功效、癌症及其它疾病检测以及个体化治疗。目前,具有易感性或早期疾病的患者不能被鉴别,并且患有疾病的那些患者不被给予最佳治疗—全都是因为诊断水平的不足。因此,急需开发可在临床实验室中使用的自动超快测序平台。需要所述低成本的台式机器来促进分子标志物的发现、批准及临床使用。

[0008] 例如,在癌症领域中,存在开发以下技术的需要:早期检测、指导疗法以及监测复发—全部都来自血液样品。这包括开发以下的需要:(i)启动子超甲基化和甲基化不足的高灵敏度检测(当以1%至0.01%的无细胞DNA存在时),(ii)常见和不常见的突变在已知基因

中的高灵敏度检测(当以1%至0.01%的无细胞DNA存在时), (iii) 肿瘤特异性mRNA和从源自肿瘤的外来体或血液中的RISC复合体分离的miRNA的精确量化, (iv) 在从循环肿瘤细胞中分离的DNA中的肿瘤特异性拷贝变化的精确量化, (v) 在从循环肿瘤细胞中分离的DNA中的突变、启动子超甲基化和甲基化不足的精确量化。上述所有情况(在从循环肿瘤细胞中分离的DNA中的肿瘤特异性拷贝变化的量化除外) 需要集中在对基因组的靶基因或靶区域的测序上。此外, 来自原始片段的两条链的序列信息或甲基化状态的确定提供极度需要的罕见事件证实。

[0009] 在产前护理领域中, 急需开发用于以下的非侵入性测定: 常见的非整倍体, 如21三体、18三体或13三体; 小缺失, 如由杜兴氏肌营养不良 (DMD) 基因中的缺失所造成的那些; 其它小的拷贝数异常, 如引起孤独症的那些; 决定潜在的临床表现的平衡易位; 甲基化变化, 其可导致与印记 (imprinting) 有关的疾病如安格曼综合征或普-威综合征; 三联体重复变化, 其引起如亨廷顿病的疾病; 点突变, 如引起囊性纤维化的CFTR基因中的那些。

[0010] 单分子测序 (SMS) 提供基于整体的策略不可用的一些独特属性, 如就实现由\$1, 000基因组计划所提出的雄心勃勃的授权而言, 基于PCR的策略。例如, SMS (i) 使样品处理管线流线化以降低完成的碱基测序成本 (Bayley, H., "Sequencing Single Molecules of DNA," *Curr. Opin. Chem. Biol.* 10:628-637 (2006)); (ii) 消除对扩增及其相关偏移以及试剂的需要以及对设计适合于基因组的困难区域 (即, 高度重复的区域) 的引物的需要; (iii) 提供以高鉴别效率直接寻找甲基化位点、罕见突变及其它序列变异的能力; (iv) 提供用于监测拷贝数变化的高灵敏度; 以及 (v) 生成可减少装配需求的长的读取, 及经过高重复区的序列。

[0011] 一种类型的SMS方法涉及合成产生聚合酶的互补DNA链, 其是由荧光标记的脱氧核苷酸构成 (Davis等人, "Rapid DNA Sequencing Based Upon Single Molecule Detection," *Genetic Analysis-Biomolecular Engineering* 8:1-7 (1991); Goodwin等人, "Application of Single Molecule Detection to DNA Sequencing," *Nucleosides & Nucleotides* 16:543-550 (1997); Goodwin等人, "DNA Sequencing by Single Molecule Detection," *Prog. Biophys. Mol. Biol.* 65:SMI02-SMI02 (1996))。使用抗生蛋白链菌素: 生物素连接物将互补DNA链固定在微珠上。光学捕获用于将珠粒: DNA复合体悬浮在用高度进行性核酸外切酶填充的油流中, 所述高度进行性核酸外切酶依次地剪切末端单核苷酸 (dNMP), 末端单核苷酸将它们经由激活激光束来发送, 所述激光束产生荧光光子爆发, 其中颜色提供用于鉴别的机制。虽然概念上很简单且不管单色测序的证明 (Werner等人, "Progress Towards Single-Molecule DNA Sequencing: A One Color Demonstration," *J. Biotechnol.* 102:1-14 (2003)), 已遇到此方法的几个挑战, 包括不能使用完全是染料修饰的dNTP来构建互补序列、由按比例缩放问题引起的扩散错序、以及在单分子检测期间减小信噪比的杂质荧光 (Demas等人, "Fluorescence Detection in Hydrodynamically Focused Sample Streams: Reduction of Diffusional Defocusing by Association of Analyte With High-Molecular Weight Species," *Appl. Spectroscopy* 52:755-762 (1998) 及 Goodwin等人, "DNA Sequencing by Single Molecule Detection," *Prog. Biophys. Mol. Biol.* 65:SMI02-SMI02 (1996))。

[0012] 最近, 已提出替代的基于荧光的SMS策略, 其在由聚合酶并入荧光标记的dNTP事件

之后并且使用监测用磷酸连接到dNTP上的光谱上不同的染料标记的dNTP的零模式波导(Eid等人,“Real-Time DNA Sequencing From Single Polymerase Molecules,”*Science*323:133-138(2009))。另一种方法使用排列在固体支撑物上的单DNA分子,该固体支撑物具有生成光子的荧光爆发的每一掺入事件(Braslavsky等人,“Sequence Information Can be Obtained From Single DNA Molecules,”*Proc.Nat’l.Acad.Sci., U.S.A.*100:3660-3964(2003))。虽然这些是直接从单分子获得序列信息的优良实例,但其确实提供一些常见的挑战,如对荧光底物的需要、分子系统之间生成串扰或交叉激发的大量光谱重叠以及对广泛的光学硬件的需要以读取所得签名。

[0013] 为了防止对从SMS形式进行基于荧光读取的需要,已提出了纳米孔技术来允许直接从包含靶DNA的单核苷酸的电子签名直接读取DNA序列,从而避免了对荧光的需要(Akeson等人,“Microsecond Time-Scale Discrimination Among Polycytidylic Acid, Polyadenylic Acid, and Polyuridylic Acid as Homopolymers or as Segments Within Single RNA Molecules,”*Biophys.J.*77:3227-3233(1999);Deamer&Branton,“Characterization of Nucleic Acids by Nanopore Analysis,”*Acc.Chem.Res.*35:817-825(2002);Meller&Branton,“Single Molecule Measurements of DNA Transport Through a Nanopore,”*Electrophoresis*23:2583-2591(2002);Meller等人,“Voltage-Driven DNA Translocations Through a Nanopore,”*Phys.Rev.Lett.*86:3435-3438(2001);以及Meller等人,“Rapid Nanopore Discrimination Between Single Polynucleotide Molecules,”*Proc.Nat’l.Acad.Sci.U.S.A.*97:1079-1084(2000))。在大多数研究中,纳米孔是 α -溶血素,其是由细菌金黄色葡萄球菌产生的蛋白质膜通道。从应用观点来看,使用此孔隙具有几个限制:(1)其机械和化学稳定性在很多情况下是不够的;(2)其具有允许仅选出类型的分子的转导的固定孔径;以及(3)制造高密度阵列的所述纳米孔的能力可能成问题。这些 α -溶血素限制造成合成纳米孔的使用(Rhee&Burns,“Nanopore Sequencing Technology:Research Trends and Applications,”*Trends Biotechnol.*24:580-586(2006)及Storm等人,“Fabrication of Solid-State Nanopores With Single-Nanometre Precision,”*Nat.Mater.*2:537-541(2003)),所述合成纳米孔可使用电子或离子束在聚合物或四氯化三硅膜中以1-50nm的尺寸来制造。合成纳米孔的有吸引力的特征是采用不同读出模式的能力,如使用装饰合成孔隙以监测隧道电流或电导变化的扰动的横向电极(Lagerqvist等人,“Fast DNA Sequencing Via Transverse Electronic Transport,”*Nano Lett.*6:779-782(2006);Lagerqvist等人,“Influence of the Environment and Probes on Rapid DNA Sequencing Via Transverse Electronic Transport,”*Biophys.J.*93:2384-2390(2007);Zikic等人,“Characterization of the Tunneling Conductance Across DNA Bases,”*Phys.Rev.E* 74(1Pt 1):011919(2006);以及Zwolak&Di Ventra,“Colloquium:Physical Approaches to DNA Sequencing and Detection,”*Rev.Modern Physics* 80:141-165(2008))。

[0014] 原则上,DNA的结构信息,不管是使用天然还是合成纳米孔,都通过由随着完整的DNA分子移动穿过孔隙时的阻塞电流幅值推导核苷酸的身份而获得。此DNA测序方法的优点包括:(1)为大DNA片段(≥ 50 kbp)测序的能力;(2)不需要使用扩增或亚克隆技术;(3)不需要使用荧光标记的脱氧核苷酸或双脱氧核苷酸;(4)需要小输入DNA样品尺寸,对于全基因

组测序大约为 1×10^8 个拷贝,以及;(5)可获得DNA序列信息的速度可提供近乎实时的读出。令人遗憾的是,直接从纳米孔进行DNA测序的研究示范还有待论证。

[0015] 已有几篇综述集中于纳米孔技术用于DNA测序的潜力,且同时这些综述指出,存在为实现此激动人心的新平台及其用于DNA测序的潜力的许多挑战(Branton等人,“The Potential and Challenges of Nanopore Sequencing,”*Nat.Biotechnol.*26:1146-1153 (2008)及Zwolak&Di Ventra,“Colloquium:Physical Approaches to DNA Sequencing and Detection,”*Rev.Modern Physics* 80:141-165 (2008))。首先,通过孔隙的易位时间相当高(每核苷酸1-20 μ s),需要在MHz范围内发挥作用的读出电子器件的带宽。其次,读出分辨率需要孔隙厚度等于或小于DNA分子的单碱基间距, ~ 0.34 nm。因为合成的及 α -溶血素孔隙的厚度比此间距大得多(5-15nm),所以多个碱基同时驻留在孔隙内。即使纳米孔可以此必要厚度来制造,有效的电场读取区域也将在孔隙的任一侧上延伸大约1个孔径单位(Liu等人,“The Effect of Translocating Cylindrical Particles on the Ionic Current Through a Nanopore,”*Biophys.J.*92:1164-1177 (2007))。第三,纳米孔阵列的生产必须以高产量模式可再生产地完成,且以必要的尺寸大小和低成本以适应预期应用。第四,高质量的基因组DNA必须从一系列不同的样品(血液、组织、骨髓、尿、唾液等等)中提取,接着处理以产生DNA片段(约50kbp),其被用作测序的输入。样品制备和测序步骤应该整合到单一平台中并且以基本的交钥匙模式操作以允许广泛的用户基础。

[0016] 如果核苷酸在物理上彼此分离同时在例如通过使用核酸外切酶从DNA上剪切之后维持其原始顺序的话,便可减少读出分辨率限制(Davis等人,“Rapid DNA Sequencing Based on Single Molecule Detection,”*Los Alamos Science* (1992))。已使用高度进行性核酸外切酶证明其是可行的,所述核酸外切酶从完整的DNA片段上依次剪切个别核苷酸且使这些碱基直接穿过配备有环糊精领的 α -溶血素纳米孔(Wu等人,“Protein Nanopores With Covalently Attached Molecular Adapters,”*J.Am.Chem.Soc.*129:16142-16148 (2007)及Clarke等人,“Continuous Base Identification for Single-Molecule Nanopore DNA Sequencing,”*Nature Nanotechnol.*4:265-270 (2009))。令人遗憾的是,使用阻塞电流的单碱基鉴别效率是93-98%(Astier等人,“Toward Single Molecule Sequencing:Direct Identification of Ribonucleoside and Deoxyribonucleoside 5'-Monophosphates by Using an Engineered Protein Nanopore Equipped With a Molecular Adaptor,”*J.Am.Chem.Soc.*128:1705-1710 (2006)),且因此,单独使用阻塞电流的测序误差不会生成鉴别例如突变位点所要求的必需的测序准确性。同样,最佳核酸外切酶活性所需的盐条件不能与高准确度碱基识别所需的条件相匹配且因此鉴别效率在90%至99%范围内。因此,必须考虑另外的碱基鉴别策略。

[0017] 本发明克服了本领域中的这些及其它缺点。

[0018] 发明概述

[0019] 本发明的第一方面涉及一种包括生物分子处理器的装置。每个生物分子处理器包括一个或多个由固体基底界定的生物反应器室、和附接于固体基底上的在每个生物反应器室内的支撑结构。生物分子处理器还具有被固定于支撑结构上并操作性地定位在所述生物反应器室内的裂解酶以裂解被裂解酶操作性地接合的生物聚合物分子的单体或多聚体单元。生物分子处理器还具有一个或多个飞行时间通道,其在固体基底中形成并流体地联接

到所述一个或多个生物反应器室。所述一个或多个飞行时间通道各自具有输入端和输出端,其中所述一个或多个飞行时间通道各自具有两个或更多个传感器,包括至少(i)邻近所述一个或多个飞行时间通道的输入端接触所述一个或多个飞行时间通道的第一传感器及(ii)邻近所述一个或多个飞行时间通道的输出端接触所述一个或多个飞行时间通道的第二传感器。

[0020] 本发明的其它方面涉及使用本发明的装置进行核酸测序和蛋白质或多肽鉴别的方法。

[0021] 本发明的另一方面涉及一种用于鉴别样品中的多个核酸分子中的一个或多个的方法。此方法涉及提供含有一个或多个核酸分子的样品,所述核酸分子各自与其相应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分和适用于发夹形成的单链3' 适配体部分。样品经受有效使所述一个或多个核酸分子及其互补序列的3' 适配体部分形成发夹的条件。所述方法进一步涉及提供聚合酶并将所述一个或多个3' 发夹核酸分子及其互补序列与聚合酶混合以形成延伸混合物。所述一个或多个核酸分子及其互补序列是从其相应的3' 发夹区域延伸以形成一个或多个全长发夹核酸分子。所述方法进一步涉及为所述一个或多个全长发夹核酸分子的至少一部分测序以鉴别样品中的一个或多个核酸分子。

[0022] 本发明的另一方面涉及一种用于鉴别样品中的多个核酸分子中的一个或多个的方法,所述方法包括在固体支撑物上的固定以实现一个或多个所需核酸分子的靶特异性富集。此方法涉及提供含有一个或多个核酸分子的样品,所述核酸分子各自与其相应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分、5' 适配体部分及3' 适配体部分,其中5' 和3' 适配体部分各自适用于发夹形成。所述方法还涉及使所述一个或多个核酸分子与其互补序列解链以形成一个或多个单链核酸分子及其互补序列的集合,其中一个或多个单链核酸分子及其互补序列各自的5' 和3' 适配体部分的发夹序列形成发夹。提供包括多个固定的捕获寡核苷酸探针的固体支撑物,所述捕获寡核苷酸探针包含与所述一个或多个核酸分子或其互补序列的靶特异性部分区域互补的核苷酸序列。所述一个或多个解链的单链核酸分子及其互补序列与被固定在固体支撑物上的其互补捕获寡核苷酸杂交。此方法还涉及提供聚合酶并将所述聚合酶与含有所述一个或多个杂交核酸分子及其互补序列的固体支撑物接触。所述一个或多个核酸分子及其互补序列从其相应的3' 发夹区域延伸以形成一个或多个全长发夹状靶核酸分子,从而使所述一个或多个核酸分子从捕获寡核苷酸且从固体支撑物上释放。对所述一个或多个核酸分子的至少一部分进行测序以鉴别样品中的所述一个或多个核酸分子。

[0023] 本发明的另一方面涉及一种用于鉴别样品中的多个核酸分子中的一个或多个的方法。此方法涉及提供含有一个或多个核酸分子的样品,所述核酸分子各自与其相应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分、5' 适配体部分及包含单链均聚物重复序列的3' 适配体部分。所述方法还涉及使所述一个或多个核酸分子与其互补序列解链以形成单链核酸分子及其互补序列的集合。提供多个与一个或多个核酸分子或其互补序列的3' 端互补的寡核苷酸引物和聚合酶,并且将所述一个或多个核酸分子及其互补序列与聚合酶和引物混合以形成延伸混合物。所述方法还涉及使所述寡核苷酸引物与其互补的所述一个或多个核酸分子及其互补序列的3' 端杂交,并且

使所述杂交的引物延伸以形成全长双链核酸分子。对所述一个或多个全长双链核酸分子的至少一部分进行测序以鉴别样品中的所述一个或多个核酸分子。

[0024] 本发明涉及一种生物聚合物测序平台,在本文中被称为生物分子处理器,其使用全自动平台可显著减少与获取测序信息有关的成本、劳动力及时间。所述平台使用纳米尺度传感器,其基于它们的特征性的通过2维(2D)飞行时间通道的飞行时间来鉴别生物聚合物分子的单体和多聚体单元,所述飞行时间通道是使用低成本的微米和纳米复制技术在可模压的固体支撑物中制造。所述单体和多聚体单元是从完整的生物聚合物,例如核酸分子或蛋白质,使用裂解酶而生成,所述裂解酶共价地固定于生物分子处理器的生物反应器室内所容纳的支撑结构。裂解的单体或多聚体单元被进料至飞行时间通道中,其中单体或多聚体的身份是从通过飞行时间通道的分子依赖性飞行时间而推导。飞行时间使用保持在通道的各端处的至少两对传感器,用由被单体或多聚体单元诱发的电响应中的扰动产生的信号来转换。

[0025] 本文中提出的方法直接解决与用于DNA序列测定的基于纳米孔的技术有关的主要挑战,如由Branton等人,“The Potential and Challenges of Nanopore Sequencing,”*Nat. Biotechnol.* 26:1146-1153 (2008) 所指示,该文献在此以引用的方式整体并入。首先,与需要易位速率减小的纳米孔测序有关的带宽限制被否定,这是由于以下的事实:个别碱基是以由酶剪切速率和施加的电场以及单体的电泳迁移率所决定的长度间隔(Δt)。例如,使用1,000个核苷酸 s^{-1} 的 λ -核酸外切酶剪切速率(Matsuura等人,“Real Time Observation of a Single DNA Digestion by Lambda Exonuclease Under a Fluorescence Microscope Field,”*Nucleic Acids Res.* 29:e79 (2001),其在此以引用的方式整体并入),和0.1mm s^{-1} 的线速度,核苷酸间距将是100nm。核苷酸在飞行管纳米通道内的多个占用不是问题,因为开始和停止时间是由保持在飞行管内的至少两个电极对来测量。其次,<0.34nm厚的纳米孔的单碱基分辨率要求不是必需的,这是由于dNMP的核酸外切酶施加的间距和阻塞电流不被用于鉴别核苷酸碱基的事实。实际上,核苷酸的身份是基于这些实体的非常确实的色谱或电泳性质(Czarnecka等人,“Application of Solid Phase Extraction and High-Performance Liquid Chromatography to Qualitative and Quantitative Analysis of Nucleotides and Nucleosides in Human Cerebrospinal Fluid,”*J. Chromatogr. B* 822:85-90 (2005),其在此以引用的方式整体并入)。另外,使用跨导或隧道电流以及关于对转换的信号分子取向作用的潜在问题并不存在于本发明中(Zikic等人,“Characterization of the Tunneling Conductance Across DNA Bases,”*Phys. Rev. E* 74(1Pt 1):011919 (2006),其在此以引用的方式整体并入)。因为在可模压的材料中使用既定复制技术生产微流体和纳米流体元件,所以它们可以低成本和高生产率来生产。另外,纳米尺度电极将使用在多孔模板中的模板化方法或本文所述的其它方法生成以允许感测元件的大规模生产。第四,本发明的生物分子处理器可与如本文所述的微米尺度的模块相接以产生在提交至序列分析之前用于处理样品的自治系统。例如,DNA可选自靶细胞,在以全自动方式测序之前被纯化和剪切。另外,由大量的可模压材料以及所采用的模块化方法提供的材料性质灵活性将允许选择对应用最优的基底材料来建造这些系统。最终,这些模块的低成本生产和它们的直接前向装配以及装配系统的高功能性将提供低成本平台并且它们的自动化能力将产生用于直接在临床环境或基于小发现的实验室中再测序

或从头测序的平台。

[0026] 附图简述

[0027] 图1是本发明的生物分子处理器的示意性俯视图。

[0028] 图2A是本发明的生物分子处理器的透视图。图2B显示了被固定于生物反应器支撑结构16并且与核酸分子15操作性接合的裂解酶36。

[0029] 图3A-3D是本发明的生物分子处理器的各个部分的扫描电子显微镜检 (SEM) 图像。图3A是含有微米柱和/或纳米柱阵列的v形进料通道进口28的SEM图像。图3B是含有熵陷阱30并且联接到生物反应器室14的进料通道20的SEM图像。图3C是含固定有裂解酶36的固体支撑物16的生物反应器室14的SEM图像。生物反应器室流体地联接到进料通道20和飞行时间通道24。图3D显示了飞行时间通道24,在通道24内具有第一对纳米感测电极32和第二对纳米感测电极34。如所示,这些纳米感测电极各自连接到微接触垫40,所述微接触垫将纳米感测电极联接到装置的外部电子部件。

[0030] 图4A-4B显示了进料通道样品进口的各种设计。图4A显示了各种进口的SEM图像,包括槽形进口、V形微米/亚微米槽形进口、具有微米柱/纳米柱阵列的柱形进口,及漏斗形进口。图4B是显示相较于较陡的微流体到纳米通道界面(“原始”),使用图4A的各种进口设计引起的DNA捕获增加的图。

[0031] 图5A-5B是显示用于将一种或更多种裂解酶固定于本发明装置的生物反应器室中的固体支撑物上的替代性方法的示意图。

[0032] 图6A-6B显示了具有含三个纳米感测电极的飞行时间通道的本发明的生物分子处理器。图6A是生物分子处理器10的俯视图,所述生物分子处理器包括进料通道20、生物反应器室14,以及含有第一对纳米感测电极32、第二对纳米感测电极34和第三对纳米感测电极33的飞行时间通道24。图6B是显示了单核苷酸46进入和排出含三对纳米感测电极32、33、34的飞行时间通道24的透视图。

[0033] 图7是本发明的生物分子处理器的俯视图,其中飞行时间通道含有多个飞行管,每个飞行管具有两对由具有独特表面化学性质的一部分飞行时间通道分隔的纳米感测电极。每个飞行管的第一传感器和第二传感器检测生物聚合物单元的飞行时间随特定飞行时间管表面化学性质的变化。

[0034] 图8A-8B是显示制备双链基因组DNA的两条链以根据本发明方法进行序列分析的替代性方法的示意图。

[0035] 图9是用于富集靶基因组DNA样品并制备所富集的双链DNA的两条链以根据本发明方法进行序列分析的方法的示意性综述。

[0036] 图10是显示制备双链基因组DNA的两条链以根据本发明方法进行序列分析的方法的示意图。

[0037] 图11是用于富集靶基因组DNA样品并制备所富集的双链DNA的两条链以根据本发明方法进行序列分析的方法的示意性综述。

[0038] 图12是适于制造本发明的生物分子处理器的第一方法的示意性综述。

[0039] 图13是适于制造本发明的生物分子处理器的第二方法的示意性综述。

[0040] 图14是适于制造本发明的生物分子处理器的第三方法的示意性综述。

[0041] 图15是形成用于制造本发明的生物分子处理器的硅母模的示意性综述。

[0042] 图16A-16B描绘了由硅母模形成树脂模(图16A),随后使用树脂模制造微流体和纳米流体网络以及与本发明的生物分子处理器相关的其它结构(图16B)所使用的纳米压印光刻工艺。

[0043] 图17A-17B显示了形成适用于本发明的生物分子处理器中的单相纳米线的工艺(图17A)以及在这一工艺中制造的纳米线的扫描电子显微镜检(SEM)图像(图17B)。

[0044] 图18显示了形成适用于本发明的生物分子处理器中的分节段纳米线的工艺。

[0045] 图19是显示了用于形成适用于本发明的生物分子处理器中的超细纳米线的工艺的示意图。

[0046] 图20显示了使用选择性蚀刻使纳米线变细的工艺。

[0047] 图21是示出了使用化学贴片将纳米线定位于本发明的生物分子处理器上的感测电极位置处的示意图。

[0048] 图22是显示了使用剥离法分子组装图案化(molecular assembly patterning by lift-off, MAPL)技术在生物分子处理器的基底上形成化学贴片的示意性综述。

[0049] 图23是显示了使用纳米模版掩模在生物分子处理器的基底上形成化学贴片的示意性综述。

[0050] 图24是显示了使用纳米模版掩模将纳米线定位于生物分子处理器上的电极传感器位置处的示意性综述。

[0051] 图25是显示了在不施加磁场的情况下,将纳米线经磁力引导到其电极传感器位置的工艺的示意性综述。

[0052] 图26是显示了使用外加磁场将纳米线经磁力引导到其电极传感器位置的工艺的示意性综述。

[0053] 图27描绘了使用纳米压印光刻法同时形成纳米间隙和纳米通道的工艺。

[0054] 图28是当在基底上预先形成纳米通道和其它流体网络时,使用纳米压印光刻法形成纳米间隙的工艺。

[0055] 图29A-29B示意性示出了用于形成微接触垫的替代性工艺,所述微接触垫将飞行时间传感器的纳米感测电极联接到外部电子部件。

[0056] 图30A-30C显示了用于将盖板热组装到生物分子处理器上,同时使结构变形减到最少的工艺。图30A是组装工艺的示意图。图30B是用于组装工艺中的压力加热器的示意图。图30C是粘接的盖板的图像。

[0057] 图31描绘了利用机载信号处理电子装置在微接触垫与印刷电路板(PCB)之间建立各种电连接的工艺。

[0058] 图32显示了本发明的生物分子处理器的电子装置。

[0059] 图33A-33D显示了含有四个核碱基的dNMP结构:图33A显示了质子化的3'dAMP;图33B显示了非质子化的3'dGMP;图33C显示了质子化的5'dCMP;并且图33D显示了非质子化的5'dTMP。图33B结构中划圈的原子与图33A结构中相应原子上局部电荷的差异使图33C结构得到图33D结构中划圈的原子上的局部电荷。图33B与图33D结构中划圈的原子之间的唯一结构差异是在C3'和C5'原子处。在模拟时使用了如图33D一般的具有四个不同核碱基的结构。核碱基的身份不影响磷酸酯基上的局部电荷。核碱基上标为轴1的轴被用于分析狭缝壁上dNMP的吸附和解吸附,并论述于本文中。

[0060] 图34A-34B显示了用于平衡和不平衡模拟的系统及DNA的化学结构。图34A显示了在纳米狭缝中含有水、氯化钠及dNMP的模拟系统。灰色球体表示壁原子,绿色表示氯化物,茶色表示磷,黄色表示钠,红色表示非水氧,白色表示非水氢,青色表示非壁碳,蓝色表示氮并且紫色表示水分子。为清楚起见,一些壁原子未示出。在x和y方向上存在周期性边界条件。图34B显示了两种dNMP(仅仅示出极性氢)。单环嘧啶(C、T)和双环嘌呤(A、G)核碱基被附接到糖-磷酸酯骨架。 λ -核酸外切酶可以消化双链DNA的一条链,从5'端开始,一次消化一个核苷酸,并且在每个释放的核苷酸的5'端上留下磷酸酯基。

[0061] 图35描绘了在算法中测定吸附和解吸附时间的步骤。为了计算在吸附期间dNMP与壁之间的能量,关于吸附使用了一种不同的定义。如果能量低于 -2.0kcal/mol ,那么认为dNMP被吸附。

[0062] 图36显示了四种dNMP的自由能随距疏水性壁的距离 d_w 的变化。这些曲线是所有不同组的壁配置的平均值。

[0063] 图37显示了在z方向(垂直于壁表面)上dNMP质量中心的典型50ns轨线。

[0064] 图38显示了对于吸附,在平衡情况(灰色)和不平衡情况(黑色)下轴1与表面平面的角度随 d_w 的变化。这些结构上的箭头指示轴指向的方向。负数表示从狭缝的中心平面指向外($d_w=1.5\text{nm}$)。

[0065] 图39显示了对于解吸附,在平衡情况(灰色)和不平衡情况(黑色)下轴1与表面平面的角度随 d_w 的变化。这些结构上的箭头指示轴指向的方向。负数表示从狭缝的中心平面指向外($d_w=1.5\text{nm}$)。

[0066] 图40A-40D描绘了轴2的定义,该轴处于核碱基的平面内。图40A描绘了dAMP。图40B描绘了dGMP。图40C描绘了dCMP。图40D描绘了dTMP。

[0067] 图41显示了对于吸附,在平衡情况(灰色)和不平衡情况(黑色)下轴2与表面平面的角度随 d_w 的变化。这些结构上的箭头指示轴指向的方向。负数表示从狭缝的中心平面指向外($d_w=1.5\text{nm}$)。

[0068] 图42显示了对于解吸附,在平衡情况(灰色)和不平衡情况(黑色)下轴2与表面平面的角度随 d_w 的变化。这些结构上的箭头指示轴指向的方向。负数表示从裂隙的中心平面指向外($d_w=1.5\text{nm}$)。

[0069] 图43是显示在 $5.9\mu\text{m}$ 内的飞行时间分布的图。对于每个dNMP,总计使用了10,000个飞行时间,其中每个数据点是通过由 0.5nm 内的飞行时间分布计算 $N_{\text{dAMP,dGMP}}=11887$ 随机样品的飞行时间总和来产生。每一dNMP的最短分析时间($\Delta t_{\text{分析}}$)是约 $10\mu\text{s}$ 。

[0070] 图44是显示了在 0.048cm/s 的中心线速度(v_{max})和 2.47mm 的通道长度下所估算的飞行时间分布的图。

[0071] 图45是显示了将 λ -核酸外切酶(λ -Exo)固定于聚(甲基丙烯酸甲酯)(PMMA)表面的工艺的示意图。

[0072] 图46A-46C描述了在PMMA表面上进行的 λ -Exo固定。图46A和46B是仅暴露于 λ -Exo(图46A)或暴露于 λ -Exo和3-(3-二甲基氨基丙基)碳化二亚胺(EDC)/N-羟基琥珀酰亚胺(NHS)偶联剂(图46B)的PMMA表面的原子力显微镜检(AFM)扫描。图46C是显示固定的 λ -Exo酶在整个聚合物表面上的尺寸分布的柱状图。

[0073] 图47显示了在自由溶液 λ -Exo消化和固相反应器 λ -Exo消化之后得到的 λ -DNA的荧

光谱。同时显示出未消化的 λ -DNA的荧光谱作为对照。

[0074] 图48是显示实时荧光监测 λ -Exo消化活性的图。将荧光标记的 λ -DNA与固定的 λ -Exo一起在MgCl₂辅因子的存在(■)和不存在(●)下培育,并且在反应时间过程中,监测相对荧光强度的降低,此指示 λ -Exo消化活性。

[0075] 图49是显示通过毛细管电泳(CE)基线分离天然脱氧核苷酸单磷酸酯(dNMP)的图。

[0076] 图50A-50B显示了热塑性(PMMA)表面处理对电渗流(electroosmotic flow, EOF)的影响。图50A显示了在天然未经过改性的PMMA纳米通道(◆)以及用O₂等离子体处理的PMMA纳米通道(▲)中的EOF。图50B显示了在天然未经过改性的PMMA纳米通道(◆)以及经过等离子体氧化并在EDC/NHS偶联剂存在下用乙二胺处理的PMMA纳米通道(■)中的EOF。

[0077] 图51A-51C描述了通过PMMA纳米通道的单一实体的电泳迁移率。图51A是显示在4000V/cm电场强度下通过PMMA纳米通道的单一实体的电泳驱动的易位时间的柱状图。这一数据与在较低场强度下得到的数据的比较指示,由于在较高电场下分散效应减弱,易位时间的分配与场强度呈反相关。图51B是在不存在电场情况下放在PMMA纳米通道中的单一银纳米粒子(Ag-NP)图像(3D图)。图51C是显示在PMMA纳米通道中电驱动的单一体易位的电泳迁移率和板数目(图51A中所示柱状图的宽度的测量值)的图。

[0078] 图52A-52I描绘了通过PMMA纳米通道的T4DNA易位事件的光学和电学测量值。图52A和52B显示了由短DNA分子(约<5 μ m)易位产生的离子电流尖峰。图52C是显示短DNA分子易位穿过纳米流体装置的纳米柱阵列和纳米通道的示意图。图52D-52F和52G-52I分别显示了由较长DNA(并非全长T4DNA)和全长T4DNA分子易位所产生的离子电流尖峰。如本文中所描述,DNA分子的长度与所产生的电流瞬态长度相关。

[0079] 图53A-53B是含有通过电化学蚀刻产生的纳米间隙的分节段纳米线的图像。

[0080] 发明详述

[0081] 本发明的第一方面涉及一种包括生物分子处理器的装置。每个生物分子处理器包括一个或多个由固体基底界定的生物反应器室、和附着于固体基底上的在每个生物反应器室内的支撑结构。生物分子处理器还具有被固定于支撑结构上并操作性地定位在所述生物反应器室内的裂解酶以裂解被裂解酶操作性地接合的生物聚合物分子的单体或多聚体单元。生物分子处理器还具有一个或多个飞行时间通道,其在固体基底中形成并流体地联接到所述一个或多个生物反应器室。所述一个或多个飞行时间通道各自具有输入端和输出端,其中所述一个或多个飞行时间通道各自具有两个或更多个传感器,包括至少(i)邻近所述一个或多个飞行时间通道的输入端接触所述一个或多个飞行时间通道的第一传感器及(ii)邻近所述一个或多个飞行时间通道的输出端接触所述一个或多个飞行时间通道的第二传感器。

[0082] 所述装置的生物分子处理器可进一步含有同样由所述固体基底壁界定的进料通道。进料通道具有从邻近固体基底表面的输入端延伸至邻近一个或多个生物反应器室的输出端的长度。

[0083] 在一个实施方案中,本发明的装置包括100-1,000个生物分子处理器、1,000-10,000个生物分子处理器或10,000-100,000个生物分子处理器。在另一实施方案中,本发明的装置包括超过100,000个生物分子处理器。

[0084] 图1是本发明装置的生物分子处理器10的示意性俯视图。生物分子处理器包括各

种多尺度流体网络及微米和纳米结构,它们共同发挥作用以对存在于样品中的生物聚合物分子进行鉴别和/或测序。具体说来,生物分子处理器具有位于进料通道20的输入端处的样品进口28,其能够接收含有生物聚合物分子如核酸分子或蛋白质的样品。进料通道20在其输出端处流体地联接到生物反应器室14,所述生物反应器室14含有在生物反应器室14内固定于固体支撑结构16的裂解酶。裂解酶当生物聚合物从进料通道20进入生物反应器室时接合并将其裂解为单体或多聚体单元。裂解的单体或多聚体单元随后进入流体联接的飞行时间通道24。如本文中更详细地描述,飞行时间通道含有至少两个传感器32、34,所述传感器当裂解的单体或多聚体单元被动电地转运穿过通道时检测这些单元。裂解的单体或多聚体单元在飞行时间通道中穿过第一和第二传感器所花费的时间是单体或多聚体单元的“飞行时间”。裂解单元的飞行时间被用于表征和/或鉴别单体或多聚体单元和/或生物聚合物分子。

[0085] 图2是本发明装置的生物分子处理器的透视图。此图中未示出完全包围装置的微流体部件的生物分子处理器的盖板。

[0086] 生物分子处理器10的固体基底12可以是聚合材料或其它可模压材料。适合的聚合材料包括但不限于聚(甲基丙烯酸甲酯)(PMMA)、聚碳酸酯(PC)、环氧基树脂、共聚物、聚砜、弹性体、环状烯烃共聚物(COC)及聚合有机硅。或者,生物分子处理器10的固体基底12可以是玻璃、二氧化硅或熔融二氧化硅基质材料。

[0087] 参看图2,生物分子处理器10的进料通道20是由固体基底12内的壁界定。进料通道20在输入端具有样品进口28,且在其输出端18流体地联接到生物反应器室14。进料通道接收并对生物聚合物分子,例如核酸分子15进行定向,之后将分子进料至生物反应器室中用于酶促处理。

[0088] 含有一个或多个生物聚合物分子的样品经由进料通道进口28进入生物分子处理器。样品进口可在结构上以多种方式修改以便促进生物聚合物分子从装置的另一模块或部件捕获和负载到生物分子处理器中。例如,样品进口可以是简单的微米/亚微米槽形进口、V形微米/亚微米槽形进口、具有微米/纳米柱阵列的柱形进口或如图3A和4A所示的漏斗进口。图4B是显示与较陡的微流体网络到纳米通道界面(在图4B中标记为“原始”)相比,修改的进口结构明显提高在各种激励电压下DNA捕获和负载到进料通道中的速度。

[0089] 在本发明的一个实施方案中,进料通道的尺寸在宽度上小于或等于100nm且在深度上小于或等于100nm。在本发明的另一实施方案中,进料通道的尺寸在宽度上小于或等于75nm且在深度上小于或等于75nm。在本发明的另一实施方案中,进料通道的尺寸在宽度上小于或等于50nm且在深度上小于或等于50nm。进料通道的长度优选地在约1 μ m至约100 μ m或更长的范围内,不过较长和较短的进料通道长度也适用于本发明的生物分子处理器。进料通道可具有任何所需的几何截面,即,圆形、三角形、正方形、矩形、五边形、六边形、七边形或八边形。

[0090] 参看图2,生物分子处理器的进料通道可进一步包括一个或多个熵陷阱30,所述熵陷阱被用于在进料通道20内对生物聚合物分子如核酸分子15进行成型、拉伸和定位。适合的熵核酸陷阱在本领域中是已知的,参见,例如Han和Craighead,“Separation of Long DNA Molecules in Microfabricated Entropic Trap Array,”*Science* 288:1026-1029 (2000),O'Donnell等人,“Pressure-driven DNA Transport Across an Artificial

Nanotopography,”*New Journal of Physics* 11:075032 (2009), 及 Reisner 等人, “Direct Self-Organization of Single DNA Molecules in a Nanoslit via Embedded Nanopt Arrays,”*Proc. Natl. Acad. Sci. USA* 106:79–84 (2009), 其在此以引用的方式整体并入。

[0091] 进料通道可进一步含有一个或多个预处理室或陷阱, 用于在生物反应器室中下游酶促裂解之前对生物聚合物分子进行定向或处理。例如, 如果生物聚合物是 mRNA 分子, 那么所述分子可需要去封端或去腺苷酸化以产生适用于酶促裂解的底物。适合的加工酶可被连接或联接到进料通道的壁或位于生物反应器室上游的进料通道中的生物反应器预处理室的壁。

[0092] 熵陷阱、进料通道及任何其它生物反应器预处理室还可含有固定的捕获寡核苷酸。在通道、陷阱或室内固定于壁或支撑结构的这些捕获寡核苷酸可包括与经过通道进料的生物聚合物分子例如核酸的末端互补的均聚物重复序列。固定的捕获寡核苷酸与二聚物分子的互补区域之间的杂交有助于当分子横穿进料通道时对其进行定向。

[0093] 进料通道 20 还可包括一个或多个与进料通道相交的传感器 26。如图 1 和 2 中描绘的实施方案所示, 进料通道 20 包括传感器 26, 例如一对感测电极, 其邻近进料通道 20 的输入端并能测量垂直于进料通道 20 的电信号。每个感测电极 26 都联接到微接触垫 40, 所述微接触垫提供与生物分子处理器的外部电子部件的连接。当生物聚合物分子 15 经过感测电极 26 时, 检测到电信号中的扰动, 并且此扰动的持续时间可提供分子长度的近似值 (Liang 等人, “Nanogap Detector Inside Nanofluidic Channel for Fast Real-Time Label Free DNA Analysis,”*Nano Lett.* 8 (5):1472–1476 (2008), 其在此以引用的方式整体并入)。

[0094] 生物分子处理器 10 的进料通道 20 在其输出端 18 处流体地联接到生物反应器室 14。图 3B 是显示联接到生物反应器室 14 的含有熵陷阱 30 的进料通道 20 的一部分的扫描电子显微照片 (SEM)。生物反应器室 14 还由生物分子处理器的固体基底的壁界定并且在宽度上为约 100 至约 1000 nm。如图 3C 的 SEM 所示, 生物反应器室 14 含有支撑结构 16, 如柱, 其上共价地附接有一种或多种裂解酶 36, 例如核酸外切酶、核糖核酸外切酶或蛋白酶。生物反应器室的支撑结构可包含与生物分子处理器的固体基底相同或不同的材料。在一个实施方案中, 支撑结构是聚合物支撑结构 (例如 PMMA、PC 或 COC 聚合物)。生物反应器室的支撑结构在宽度上为约 50 nm 至约 900 nm 且在高度上为约 10 nm 至约 100 nm。支撑结构的表面积为约 1500 至约 285,000 nm²。固定在支撑结构上的裂解酶接合并裂解从进料通道 20 的输出端 18 进入生物反应器室 14 的生物聚合物分子。在一些实施方案中, 裂解酶可需要活化。将含有活化剂例如 Mg²⁺ 的缓冲剂或电泳缓冲剂保存在流体地联接到生物反应器室 14 的储存室 42 中 (图 2)。

[0095] 图 2B 显示被固定在生物反应器支撑结构 16 上且与核酸分子 15 接合的裂解酶 36。可使用本领域中已知的标准偶联化学方法将裂解酶连接于生物反应器室的支撑结构。在一个实施方案中, 经由偶发固定将酶连接于支撑结构 (图 5A)。可使用常规地用于多种抗体的非常确实的碳二亚胺 (1-乙基-3-(3-二甲基氨基丙基) 碳二亚胺, EDC) 化学物质来固定商购获得的裂解酶如核酸外切酶 (New England Biolabs) 或其它酶 (McCarley 等人, “Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,”*J. Am. Chem. Soc.* 127:842–843 (2005); Chen 等人, “Functional Template-Derived Poly (methyl methacrylate) Nanopillars for Solid-Phase Biological Reactions,”*Chem. Mater.* 19:3855–3857 (2007); 及 Wang 等人, “Microarrays Assembled in

Microfluidic Chips Fabricated From Poly(methyl methacrylate) for the Detection of Low-Abundant DNA Mutations,”*Anal.Chem.*75:1130-1140 (2003), 其在此以引用的方式整体并入)。

[0096] 在本发明的另一实施方案中,经由固定适合的捕获部分将裂解酶连接于支撑结构,其中所述酶含有或被工程化以含有捕获部分结合配偶体。例如,在一个实施方案中,支撑结构的表面含有多个结合至Ni(II)的捕获配体,其允许使用六组氨酸标记工程化的酶进行酶固定(图5B)(Dapprich,J.,“Single-Molecule DNA Digestion by Lambda-Exonuclease,”*Cytometry* 36:163-168 (1999), 其在此以引用的方式整体并入)。Ni(II)表面可通过将Ni(II)配位至用1-乙基-4-苄基-三氮杂环壬烷修饰的CT-PMMA(Acbztacn)(Johnson&Martin,“Controlling Protein Orientation at Interfaces Using Histidine Tags:An Alternative to Ni/NTA,”*J.Am.Chem.Soc.*127:2018-2019 (2005); Warden等人,“Synthesis of Novel Derivatives of 1,4,7-Triazacyclononane,”*Organic Lett.*3:2855-2858 (2001), 其在此以引用的方式整体并入)或次氨基三乙酸(NTA)而产生。Acbztacn-PMMA表面可通过将CT-PMMA表面暴露于EDC/Acbztacn以经由三氮杂环壬烷的仲胺形成Acbztacn与PMMA的酰胺键而形成(参见图5B)(Johnson&Martin,“Controlling Protein Orientation at Interfaces Using Histidine Tags:An Alternative to Ni/NTA,”*J.Am.Chem.Soc.*127:2018-2019 (2005), 其在此以引用的方式整体并入)。

[0097] 可用于将一种或多种裂解酶连接于支撑结构的替代的捕获和结合配偶体包括但不限于生物素和抗生物素蛋白链菌素、麦芽糖和麦芽糖结合蛋白、壳多糖和壳多糖结合蛋白、淀粉酶和MBP、谷胱甘肽转移酶和谷胱甘肽-S-转移酶、整联蛋白和整联蛋白结合肽、核酸寡核苷酸或核酸类似物寡核苷酸及其互补寡核苷酸。

[0098] 在另一实施方案中,使用固定抗体将酶连接于支撑结构上。例如,已被工程化以含有六组氨酸标签的裂解酶可经由抗His-标签抗体固定于支撑结构上(Perkins等人,“Sequence-Dependent Pausing of Single Lambda Exonuclease Molecules,”*Science* 301:1914-1918 (2003), 其在此以引用的方式整体并入)。商购获得的抗组氨酸-标签抗体(Roche、Qiagen或Novagen)或其它抗体的固定可使用常规地用于多种抗体的非常确实的碳二亚胺(1-乙基-3-(3-二甲基氨基丙基)碳二亚胺,EDC)化学物质来实现(McCarley等人,“Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,”*J.Am.Chem.Soc.*127:842-843 (2005); Chen等人,“Functional Template-Derived Poly(methyl methacrylate) Nanopillars for Solid-Phase Biological Reactions,”*Chem.Mater.*19:3855-3857 (2007); 及Wang等人,“Microarrays Assembled in Microfluidic Chips Fabricated From Poly(methyl methacrylate) for the Detection of Low-Abundant DNA Mutations,”*Anal.Chem.*75:1130-1140 (2003), 其在此以引用的方式整体并入)。含有工程化的肽捕获序列的酶的抗体捕获的另外实施方案包括但不限于:FLAG表位和抗FLAG抗体;及Myc标签表位和抗Myc标签抗体。

[0099] 裂解酶对支撑结构的表面覆盖率可通过评价固定之前和之后固定溶液中的蛋白质含量(溶液差分法)来评定(Smith等人,“Measurement of Protein Using Bicinchoninic Acid,”*Anal.Biochem.*150:76-85 (1985); Stoscheck,C.M.,“Quantitation of Protein,”*Methods in Enzymol.*182:50-68 (1990), 其在此以引用的方式整体并入),且

其活性将使用标准动力学方法来测定 (Chen等人, “Functional Template-Derived Poly (methyl methacrylate) Nanopillars for Solid-Phase Biological Reactions,” Chem.Mater.19:3855-3857 (2007), 其在此以引用的方式整体并入)。

[0100] 在本发明的一个实施方案中,生物反应器室的固定的裂解酶是核酸外切酶。如本文所用,“核酸外切酶”包括能催化单个核苷酸从DNA或RNA分子末端上水解的任何酶。在一个实施方案中,核酸外切酶是进行性酶,即,其催化模板的一系列连续裂解事件,而不释放模板。所述核酸外切酶可以是单体酶、多聚体酶或由多个亚基组成的酶复合体。适合的核酸外切酶包括但不限于 λ 核酸外切酶,其在5'-3'方向上裂解双链和单链DNA底物;核酸外切酶I,其在3'-5'方向上裂解单链DNA底物;核酸外切酶III,其在3'-5'方向上裂解双链DNA底物;T7核酸外切酶,其在5'-3'方向上裂解双链DNA底物、或RNA/DNA杂交体底物;XRN-1,其在5'-3'方向上裂解单链RNA底物 (Geisler和Coller, “XRN1: A Major 5' to 3' Exoribonuclease in Eukaryotic Cells,” Enzymes 31:97-110 (2012), Jinek等人, “Coupled 5' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay,” Mol.Cell 41:600-608 (2011), 及Jones等人, “The 5' $\rightarrow$ 3' Exoribonuclease XRN1/ Pacman and its Functions in Cellular Processes and Development,” WIREs RNA 3: 455-468 (2012), 其在此以引用的方式整体并入); RNase II, 其在3'-5'方向上裂解单链RNA底物 (Zuo等人, “Structural Basis for Processivity and Single-Strand Specificity of RNaseII,” Mol.Cell 24:149-156 (2006), 其在此以引用的方式整体并入); 以及外来体复合体, 其在3'-5'方向上裂解单链RNA底物 (Lee等人, “Elastic Coupling Between RNA Degradation and Unwinding by an Exoribonuclease, Science 336:1726-1729 (2012), 其在此以引用的方式整体并入)。生物反应器室的支撑结构可包含上述核酸外切酶中的任何一种或多种。在一些实施方案中, 酶的组合, 例如像外来体复合体与XRN-1的组合确保在RNA分子进入生物反应器室时对其进行裂解, 而不管RNA方向如何 (即5' -3' 或3' -5')。在一些实施方案中, 可能需要前处理步骤, 例如, 对信使RNA进行去封端以产生用于XRN-1的5'磷酸化底物, 或对信使RNA进行去腺苷酸化以产生用于外来体复合体的适合的底物。这种额外的酶可存在于预反应室、进料通道中, 或附着于生物反应器室的支撑结构上。

[0101] 在本发明的另一实施方案中,生物反应器室的固定的裂解酶是蛋白酶或肽酶。如本文所用,“蛋白酶”和“肽酶”可互换地用于指示能够通过肽键的水解进行蛋白水解的任何酶。适合的蛋白酶包括但不限于丝氨酸蛋白酶、苏氨酸蛋白酶、半胱氨酸蛋白酶、天冬氨酸盐蛋白酶 (在天冬氨酸盐残基的氨基侧裂解)、天冬酰胺蛋白酶、赖氨酸蛋白酶 (在赖氨酸残基的羧基侧裂解)、金属蛋白酶及谷氨酸蛋白酶 (在谷氨酸盐或天冬氨酸盐的羧基侧裂解)。适用于本发明中的特定蛋白酶包括但不限于蛋白酶K, 其在脂族、芳族及疏水性残基的羧基侧裂解; 金黄色葡萄球菌V-8蛋白酶, 其在天冬氨酸盐和谷氨酸盐残基的羧基侧裂解; 胰蛋白酶, 其在精氨酸和赖氨酸残基的羧基侧裂解; 以及胰凝乳蛋白酶, 其在酪氨酸、苯丙氨酸、色氨酸及亮氨酸的羧基侧裂解。

[0102] 如图2所示,生物反应器室14流体地联接到在进料通道20对侧运行的输出通道44。输出通道44可用来收集非裂解生物聚合物组分, 例如, 当生物聚合物分子是双链DNA分子且仅一条DNA链是由裂解酶裂解时, 输出通道收集DNA的非消化链。或者, 输出通道可用作第二进料通道, 其收集非消化的生物聚合物单元并且将其转运到一个或多个其它生物反应器室

的下游用于裂解。

[0103] 生物反应器室14还流体地联接到由生物分子处理器10的固体基底12的壁所界定的飞行时间通道24。飞行时间通道的尺寸在宽度上小于或等于50nm且在深度上小于或等于50nm。在本发明的另一实施方案中,飞行时间通道的尺寸在宽度上小于或等于25nm且在深度上小于或等于25nm。在本发明的另一实施方案中,飞行时间通道的尺寸在宽度上小于或等于15nm且在深度上小于或等于15nm。在本发明的另一实施方案中,飞行时间通道的尺寸在宽度上小于或等于10nm且在深度上小于或等于10nm。在本发明的另一实施方案中,飞行时间通道的尺寸在宽度上小于或等于5nm且在深度上小于或等于5nm。飞行时间通道在长度上可以是1 μ m至>250 μ m或5 μ m至250 μ m,并且可具有任何所需几何截面,即半球形、三角形、正方形、矩形、五边形、六边形、七边形或八边形。

[0104] 如图2和图3D的SEM所示,飞行时间通道24具有至少一个第一传感器32,例如,邻近飞行时间通道24的输入端22与飞行时间通道24相交的一对纳米感测电极;和第二传感器34,例如,邻近飞行时间通道24的输出端38与飞行时间通道24相交的一对纳米感测电极。纳米感测电极各自联接到微接触垫40,所述微接触垫提供与生物分子处理器的外部电子和检测器部件的连接。飞行时间通道可含有超过两个传感器。如图6A所示,飞行时间通道24可包含至少一个第一32、第二34和第三33传感器,例如,三对纳米感测电极。如图7所示,飞行时间通道可包含超过三个传感器。

[0105] 飞行时间通道的纳米感测电极能够测量垂直于通道的电子签名。在本发明的一个实施方案中,每对电极能够检测当分子实体驻留或穿过电极之间时由于溶液电导或阻抗的变化而造成的在电极之间运行的电流的变化。

[0106] 本发明的装置还具有电场发生器,其被操作性地定位以在一个或多个生物反应器室中且沿着一个或多个飞行时间通道的长度产生电场。电场施加于生物反应器室14和飞行时间通道24,以便将生物聚合物分子的裂解单元动电地转运出生物反应器室14并穿过飞行时间通道24。当裂解的单体或多聚体单元移动穿过飞行时间通道24时,这些单元由至少第一32和第二34传感器以及飞行时间通道中的任何另外的传感器来检测。从核酸分子上裂解的单核苷酸横穿含有三个传感器的飞行时间通道的三维描绘显示于图6B中。如本文所述,将两个或更多个传感器之间(例如,第一对与第二对纳米电极之间和/或第二对与第三对纳米电极之间)的每个生物聚合物单元的飞行时间用于鉴别和/或表征裂解的生物聚合物单元的身份。另外,如由第一和/或第二或更多传感器检测的裂解单元的电力峰值振幅也被用于帮助鉴别和/或表征裂解的生物聚合物单元。

[0107] 生物聚合物单元穿过飞行时间通道的行程时间是由施加的电场、飞行时间通道的长度及生物聚合物单元的数量和通道壁相互作用或个别生物聚合物单元的电泳迁移率的差异决定。如果壁相互作用或电泳迁移率是分子依赖性的(例如,单核苷酸碱基),那么飞行时间将提供生物聚合物单元身份的直接指示。

[0108] 生物聚合物单元与飞行时间通道的一个或多个壁之间的分子特异性相互作用可通过飞行时间通道壁的组成和官能化来控制。在一个实施方案中,飞行时间通道的壁包含与基底相同的组成,有或没有修改。或者,飞行通道的一个或多个壁可包含与基底不同的组成,有或没有修改。在另一实施方案中,分子依赖性飞行时间是由每个裂解的生物聚合物单元的独特电泳迁移率决定且因此不需要壁相互作用。

[0109] 在本发明的一个实施方案中,飞行时间通道包括聚合材料,例如PMMA、PC、环氧树脂、共聚物、聚砜、弹性体及聚合有机硅、或这些材料的任何组合。聚合材料可呈其天然状态,或替代地,被表面修饰以增强生物聚合物单元辨别。例如,聚合飞行时间通道壁可包括具有不同链有序度的中性、疏水性、碳氢化合物表面。在另一实例中,飞行时间通道壁表面可包括电中性亲水性表面。在又一实例中,飞行时间通道壁表面可包括带电的亲水性表面。

[0110] 包括具有不同链有序度的中性、疏水性、碳氢化合物表面的飞行时间通道壁表面可以是由构建在聚合物纳米通道表面上的具有各种长度的甲基封端的烷烃链单层形成(Henry等人,“Surface Modification of Poly(methyl methacrylate) Used in the Fabrication of Microanalytical Devices,”*Anal.Chem.*72:5331-5337 (2000),其在此以引用的方式整体并入)。单层可通过将氨基-烷烃附接于羧酸封端的表面而形成(McCarley等人,“Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,”*J.Am.Chem.Soc.*127:842-843 (2005); Wei等人,“Photochemically Patterned Poly(methyl methacrylate) Surfaces Used in the Fabrication of Microanalytical Devices,”*J.Phys.Chem.B* 109:16988-16996 (2005),其在此以引用的方式整体并入)。或者,单层可由在胺官能团上的尿素连接的烷烃层形成,所述胺官能团经由酰胺键附着于聚合物(Henry, A.C., “Surface Modification and Characterization of PMMA Used in the Construction of Microelectromechanical Systems,” In *Chemistry*, 第147页, Louisiana State University, Baton Rouge (2001); Henry等人, “Surface Modification of Poly(methyl methacrylate) Used in the Fabrication of Microanalytical Devices,” *Anal.Chem.*72:5331-5337 (2000),其在此以引用的方式整体并入)。例如,充分有序的十八烷基单层可通过异氰酸正十八酯与胺封端的PMMA表面反应而在PMMA表面上形成(Henry&McCarley, “Selective Deposition of Metals on Plastics Used in the Construction of Microanalytical Devices: Photo-Directed Formation of Metal Features on PMMA,” *J.Phys.Chem.B* 105:8755-8761 (2001),其在此以引用的方式整体并入),并且这些C₁₈-PMMA表面对于模压通道中的色谱分离是优良的(Galloway等人, “Contact Conductivity Detection in Poly(methyl methacrylate)-Based Microfluidic Devices for Analysis of Mono- and Polyanionic Molecules,” *Anal.Chem.*74:2407-2415 (2002),其在此以引用的方式整体并入)。因此,各种链长的正烷基异氰酸酯可用于制造具有不同有序度的疏水性聚合物表面,这将影响单体如dNMP的飞行时间。关于非零电渗流(EOF)的问题可通过为未反应的官能团封端来解决(Henry, A.C., “Surface Modification and Characterization of PMMA Used in the Construction of Microelectromechanical Systems,” In *Chemistry*. Louisiana State University, Baton Rouge (2001); Wei等人, “Photochemically Patterned Poly(methyl methacrylate) Surfaces Used in the Fabrication of Microanalytical Devices,” *J.Phys.Chem.B* 109:16988-16996 (2005),其在此以引用的方式整体并入)。

[0111] 后两个目标是通过将具有(a)乙二醇和(b)有机酸或胺端基的材料附接到羧基或胺封端的聚合物表面来完成。用于形成亲水性电中性表面的一种方法涉及将适当活化的羧酸封端聚合物表面与乙醇胺或氨基-三(乙二醇)反应(Wei, S., “Multianalyte Detection of Breast Cancer by Fabrication of Hybrid microarrays on Polymer-based

Microanalytical Devices,”In Chemistry.Louisiana State University,Baton Rouge (2005),其在此以引用的方式整体并入)。作为一个替代方案,胺封端的PMMA和PC表面可用乙二醇修饰,使得表面生成羧基,如羟基乙酸或羧基-三(乙二醇)。阳离子表面(在用于核酸外切酶的pH条件下)可使用用于产生胺封端的聚合物的非常确实的方法形成(Henry&McCarley,“Selective Deposition of Metals on Plastics Used in the Construction of Microanalytical Devices:Photo-Directed Formation of Metal Features on PMMA,”J.Phys.Chem.B 105:8755-8761 (2001);Henry等人,“Surface Modification of Poly(methyl methacrylate)Used in the Fabrication of Microanalytical Devices,”Anal.Chem.72:5331-5337 (2000);McCarley等人,“Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,”J.Am.Chem.Soc.127:842-843 (2005);Wei等人,“Photochemically Patterned Poly(methyl methacrylate)Surfaces Used in the Fabrication of Microanalytical Devices.J.Phys.Chem.B109:16988-16996 (2005),其在此以引用的方式整体并入)。阴离子表面将由生成羧酸封端的表面(McCarley等人,“Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,”J.Am.Chem.Soc.127:842-843 (2005);Vaidya等人,“Surface Modification and Characterization of Microfabricated Poly(carbonate) Devices:Manipulation of Electroosmotic Flow,”Analyst 127:1289-1292 (2002),其在此以引用的方式整体并入)或带有磺酸的表面(后者具有几乎非pH依赖性表面电荷)(Henry,A.C.,“Surface Modification and Characterization of PMMA Used in the Construction of Microelectromechanical Systems,”In Chemistry,第147页,Louisiana State University,Baton Rouge (2001),其在此以引用的方式整体并入)的途径产生。

[0112] 大多数修饰化学方法是基于生成支架,例如羧基,其包含可区域特别图案化的官能团,因为仅某些部位通过掩盖不意在被活化的区域且UV暴露此组件而在基底上被活化(McCarley等人,“Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,”J.Am.Chem.Soc.127:842-843 (2005);Wei等人,“Photochemically Patterned Poly(methyl methacrylate)Surfaces Used in the Fabrication of Microanalytical Devices.J.Phys.Chem.B109:16988-16996 (2005),其在此以引用的方式整体并入)。单层元件可经由掩盖仅选择性地固定于纳米通道域,留下其余的固体支撑物基底及其中的结构呈其天然形式。尽管纳米通道尺寸小于激活光的衍射极限(254nm),但通道的长度远高于衍射极限且这是临界尺寸。

[0113] 如图7所示,本发明装置的生物分子处理器可包括联接在一起的多个飞行管。每个飞行管具有由飞行时间通道的一部分分隔开的传感器。每个飞行管的特征为与裂解的生物聚合物单元区别地相互作用的独特或不同的化学特性。此设计形式允许多维生物聚合物单元分隔以增强个别的生物聚合物单元的鉴别和表征。此外,这样一种装置可用于形成由输入样品中的一种或多种蛋白质的蛋白水解所生成的肽片段的指纹。利用具有两个或更多个传感器和/或两个或更多个飞行管的飞行时间通道,所述飞行时间通道具有不同的壁表面化学特性以便每通道生成一个或多个飞行时间测量结果,从而形成十分复杂的肽指纹,例如2-维(使用4个传感器)、3-维(使用6个传感器)或4-维指纹(使用8个传感器)以将基本上

所有可能的蛋白质及其修饰形式与给定部分区分开(参见图7)。

[0114] 如上所述,飞行时间通道中的传感器能够当裂解的生物聚合物单元传送穿过通道时对其进行检测。在本发明的一个实施方案中,每个传感器包括一对感测电极,所述感测电极是由一对贵金属纳米线、分节段纳米线或以穿过通道的纳米间隙分隔的纳米颗粒形成。在本发明的另一实施方案中,每对感测电极具有其中所述电极接触飞行时间通道的狭窄横截面。纳米线具有小于50nm、更优选地小于约30nm、更优选地小于约10nm的直径。在本发明的另一实施方案中,纳米线具有在约5nm与约20nm之间的直径。感测电极之间的距离,即纳米间隙,在1nm与10nm之间,且在一些实施方案中,在1nm与5nm之间。感测电极和间隙尺寸控制多聚体或单体单元的信噪比,其中较小的纳米线和间隙尺寸得到更高的信噪比(Liang等人,“Nanogap Detector Inside Nanofluidic Channel for Fast Real-Time Label-free DNA Analysis,”*Nano Letters* 8:1472-1476 (2008);Tsutsui等人,“Identifying Single Nucleotides by Tunneling Current,”*Nature Nanotechnology* 5:286-290 (2010); Ohshiro等人,“Single-Molecule Electrical Random Resequencing of DNA and RNA,”*Scientific Reports* 2:1-7 (2012);其在此以引用的方式整体并入)。

[0115] 如本文中更详细地描述,纳米线可以是由金属组分的组合形成的分节段纳米线。示例性组合包括但不限于金(Au)和钴(Co) (Au/Co)以及金(Au)和铂(Pt)。

[0116] 生物分子处理器是本发明装置的一个模块或部件,且如上所述,该装置可含有100-100,000个生物分子处理器或超过100,000个生物分子处理器。生物分子处理器可与装置的一个或多个前端处理模块相连接以生成用于生物聚合物分析的高度集成系统。模块化设计方法允许在系统中混合和匹配不同模块的机会以适应应用需要。另外,因为模块是自治的,所以在系统级集成之前可就模块的性能度量对其进行优化。装置的前端处理模块包括例如且不限于:用于从细胞溶解产物固相萃取核酸(DNA和RNA)的模块(Witek等人,“96-Well Polycarbonate-Based Microfluidic Titer Plate for High-Throughput Purification of DNA and RNA,”*Analytical Chemistry* 80:3483-3491 (2008);Park等人,“A Titer Plate-Based Polymer Microfluidic Platform for High Throughput Nucleic Acid Purification,”*Biomedical Microdevices* 10:21-33 (2008),其在此以引用的方式整体并入)或用于蛋白质/多肽分离和富集的模块、用于将核酸剪切成适当尺寸以进入生物分子处理器中的模块、热扩增模块(Hashimoto等人,“Rapid PCR in a Continuous Flow Device,”*Lab On A Chip* 4:638-645 (2004);Hashimoto等人,“Ligase Detection Reaction/Hybridization Assays Using Three-Dimensional Microfluidic Networks for the Detection of Low-Abundant DNA Point Mutations,”*Analytical Chemistry* 77:3243-3255 (2005),其在此以引用的方式整体并入)、罕见细胞选择模块(Adams等人,“Highly Efficient Circulating Tumor Cell Isolation From Whole Blood and Label-Free Enumeration Using Polymer-Based Microfluidics With an Integrated Conductivity Sensor,”*J. Am. Chem. Soc.* 130:8633-8641 (2008);Dharmasiri等人,“Capture and Enumeration of LNCaP Prostate Cancer Cells Using Aptamers Immobilized to a PMMA Microsampling Unit,”*Electrophoresis* 30:3289-3300 (2009),其在此以引用的方式整体并入)以及DNA阵列(Xu等人,“Polymer Microfluidic Chips with Integrated Waveguides for Reading Microarrays,”*Analytical Chemistry* 79:

9007-9013 (2007), 其在此以引用的方式整体并入)。

[0117] 本发明的另一方面涉及一种使用含有生物分子处理器的本发明装置进行核酸分子测序的方法。所述方法涉及在有效使固定的裂解酶在样品中接合一个或多个核酸分子并将一个或多个核酸分子裂解成单体核苷酸(所述单体核苷酸个别地进入一个或多个飞行时间通道的输入端)的条件下将含有一个或多个核酸分子的样品进料至装置的生物分子处理器中。所述方法还涉及跨一个或多个生物反应器室并沿一个或多个飞行时间通道的长度施加电场以便将裂解的单体核苷酸转运穿过一个或多个飞行时间通道。当裂解的单体核苷酸在一个或多个飞行时间通道中穿过至少第一和第二传感器时,基于所述施加来检测裂解的单体核苷酸。测量每个裂解的单体核苷酸在一个或多个飞行时间通道中穿过至少第一和第二传感器所花费的时间,并且基于所述测量鉴别裂解的单体核苷酸。

[0118] 如上所述,装置的生物分子处理器可进一步含有进料通道,所述进料通道流体地联接到生物反应器室。进料通道可含有一个或多个熵陷阱和预处理室,所述熵陷阱和预处理室在核酸分子在生物反应器室中裂解之前对其进行定向和延长。因此,在本发明的一些实施方案中,包含一个或多个核酸分子的样品被引入进料通道的通道样品进口中。沿进料通道的长度施加电场以在核酸分子进入生物反应器室之前穿过进料通道延长所述核酸分子。

[0119] 可使用本发明的方法和装置测序的核酸分子包括但不限于脱氧核糖核酸(DNA)分子和核糖核酸(RNA)分子。具体地说,双链DNA、单链DNA分子、DNA和RNA分子包含一个或多个甲基化的核苷酸碱基,DNA和RNA分子包含一个或多个修饰或损伤的核苷酸碱基。具体说来,本发明的方法可容易地鉴别包含一个或多个核苷酸碱基插入、缺失、易位及突变的DNA和RNA分子,特别是当以低丰度存在于样品中时。当一个或多个核酸分子是双链DNA分子时,本发明的方法提供鉴别来自双链DNA分子的两条链的单体核苷酸的至少一部分的能力。可使用本发明方法测序的RNA分子包括例如长的非编码RNA(lncRNA)、核糖体RNA(rRNA)、小核RNA(snoRNA)、微小RNA(miRNA)、转移RNA(tRNA)及小干扰RNA(siRNA)、以及RNA/DNA杂交分子。一个或多个核酸分子可由任何生物来源,包括但不限于组织、细胞、血清、血浆、血液或外来体分离。

[0120] 对于核酸分子测序,生物反应器室中的裂解酶是核酸外切酶或核糖核酸外切酶。适合的核酸外切酶和核糖核酸外切酶如上所述。对于核酸分子测序,裂解酶优选地是进行性核酸外切酶或核糖核酸外切酶,其在核酸分子进入生物反应器室时接合所述核酸分子并且将核酸分子依次地裂解成单体单元,即个别的核苷酸碱基。裂解的单体单元被动电地转运出生物反应器室并进入和穿过飞行时间通道,在其中所述裂解的单体单元由传感器来检测并且基于其穿过通道的飞行时间来鉴别。

[0121] 如本文实施例中更详细地描述,在溶液中的每个核苷酸碱基,例如脱氧核苷酸5'-单磷酸碱基移动穿过飞行时间通道,以依赖于其个别分子身份、尺寸及通道表面的化学身份(例如,如上所述,具有不同的链有序度的中性、疏水性、碳氢化合物表面比对电中性亲水性表面比对带电的亲水性表面)的方式与通道壁相互作用。因此,每个单核苷酸碱基是基于其穿过飞行时间通道的时间来区分,所述时间是通过通道的至少两个在空间上分开的传感器来检测和测量。由于每个单核苷酸的电力峰值振幅可不同(Tsutsui等人,“Identifying Single Nucleotides by Tunneling Current,”Nat.Nanotech.5:286-290(2010),其在此

以引用的方式整体并入),所以还可以测量每个裂解的单核苷酸当穿过至少一个传感器时的电力峰值振幅。此测量可结合单核苷酸飞行时间使用以鉴别当横穿飞行通道时的碱基。

[0122] 在另一方式中,单一的单核苷酸可不与飞行管相互作用,但仍显示分子依赖性的飞行时间。因为单核苷酸被电驱动穿过飞行管,所以其可显示在电泳迁移率上的差异,从而显示依赖于特定的单核苷酸而不需要壁相互作用的独特的飞行时间。在醚形式(ether format)中,飞行时间是从感测电极对之间的行程时间推导出。

[0123] 在本发明的一个实施方案中,飞行时间通道具有两个传感器,例如,两对纳米感测电极,其测量个别核苷酸单体横穿飞行时间通道时的飞行时间。仅举例来说,当两个电极之间的间距是10 μ m(微米)且所施加的电压是10,000V/cm时,四个单核苷酸将具有以下平均飞行时间值(基于传统的毛细管电泳结果):

[0124] GMP 242 μ s(微秒)

[0125] AMP 246 μ s

[0126] CMP 254 μ s

[0127] TMP 266 μ s

[0128] 对于每次运行,给定的单核苷酸将具有限定的飞行时间值。因此,对于生物聚合物的个别单体单元,在飞行时间值中将存在高斯分布。

[0129] 在另一实施方案中,飞行时间通道含有三个或更多个传感器,例如,三对或更多对纳米感测电极,沿着纳米通道近似相等地间隔开(参见图6A-6B)。当裂解的生物聚合物单元进入飞行时间通道时,其经过第一传感器,所述传感器产生第一电子签名,可测量所述第一电子签名以给出第一信号。第一信号在某一时间出现,并且还具有给定的形状和幅度。当生物聚合物单元继续穿过飞行时间通道时,其经过第二传感器,并且这生成第二电子签名,可测量所述第二电子签名以给出第二信号。第二信号在某一时间出现,并且还具有给定的形状和幅度。此过程重复至第三个,并且直到第n个传感器。根据此实施方案,获得三个信号并且可计算每个传感器(1、2或1、3或2、3)之间的飞行时间。此提供三个飞行时间数值以及三个独立的信号签名。

[0130] 三个飞行时间数值彼此依赖,因为传感器1与3之间的飞行时间表示传感器1与2和传感器2与3之间的飞行时间总和。尽管如此,如果飞行时间除以行进的距离,那么获得两个独立的速度,以及对于每个生物聚合物单元的平均速度。对于n个电极,获得n-1个独立的速度,并且通过控制电极数量来控制每个平均值的统计完整性。这允许明显地改善进入飞行时间纳米通道的给定的生物聚合物单元的识别(call)准确性(减小误差率至0.25%)。此外,幅度和形状的另外特征(信号的上升和下降)提供生物聚合物单元识别的潜在的进一步验证的机会。例如,如果AMP和CMP的飞行时间重叠2%的时间,那么三个信号签名将有助于进一步辨别。

[0131] 如果生物聚合物单元吸附到壁上或显示在其电泳迁移率上的较大差异,且当另一生物聚合物单元在飞行时间通道中时离开,这可能导致错序错误。然而,使用超过两对感测电极将立即标记此错误。使用多个感测站将明显地改善碱基识别的准确性。不能校正可能在进入飞行时间通道之前出现的错误,如碱基损失或错序错误。所述错误可通过如本文所述对双链靶DNA的两条链进行测序而减少。

[0132] 飞行时间通道可进一步包括一个以上联接在一起的飞行管,每个飞行管由至少两

个传感器组成,所述传感器被具有独特的壁表面化学特性的飞行时间通道的一部分分隔。联接的飞行时间管在飞行时间通道内的不同构造提供多维分隔(参见图7)。获得生物聚合物单元穿过包含不同表面化学特性的通道的飞行时间测量将提高极类似的单核苷酸(例如甲基化或羟甲基化比对非甲基化碱基)之间的辨别。

[0133] 对每个依次裂解的单体核苷酸碱基重复此过程以获得样品中的一个或多个核酸分子的至少一部分的核苷酸序列的至少一部分。在一些实施方案中,获得样品中的一个或多个核酸分子的至少一部分的整个核苷酸序列。重复进行所述方法的施加、检测、测量及鉴别步骤以获得样品中的一个以上核酸分子的至少部分核苷酸序列。在一些实施方案中,所述方法能够获得样品中的一个以上核酸分子的整个或基本上整个核苷酸序列。

[0134] 困扰着一些当前可用的测序技术的一个问题是不能精确地检测罕见的突变或甲基化事件,例如,启动子超甲基化或甲基化不足或以1%至0.01%总DNA存在的突变。举例来说,考虑到循环肿瘤DNA在具有p53基因中的突变的血浆中的存在-或一百个基因中的任何一个可突变,或一百个启动子区域在那个类型的肿瘤中可甲基化。所述样品将含有大部分从正常细胞中产生的无细胞DNA,其中肿瘤DNA可仅包含0.01%的总的无细胞DNA。因此,如果试图通过总测序发现所述突变DNA的存在,那么将需要对100,000个基因组进行测序以鉴别具有突变的10个基因组。这将需要对300,000GB的DNA进行测序,一个在当前测序技术能力范围之外的任务,更不必说庞大的数据管理问题。为了解决此问题,许多小组试图捕获特定靶区域或PCR扩增所讨论的区域。序列捕获遭受失灵,使得大概90-95%的所需序列被捕获,但所需片段丢失。或者,PCR扩增提供引入不能与真正的突变区别的罕见错误的风险。此外,PCR丢失甲基化和羟甲基化信息。虽然亚硫酸氢盐处理在传统上用于确定启动子甲基化的存在,但其还破坏DNA样品并且缺乏鉴别无细胞的DNA中的多个甲基化变化的能力。

[0135] 虽然单分子测序可摒除一些上述测序限制,但传统的单分子测序的主要限制是与大多数这些技术有关的高误差率。当使用基于核酸外切酶的测序时,破坏一条链,这免除了为了检验突变或甲基化状态而对双链DNA的两条链的测序。这对于基于 λ 核酸外切酶的测序尤其是个挑战,所述测序需要5' 磷酸盐以及双链DNA底物以便确保仅单核苷酸通过核酸酶活性以进行性方式生成。如下所述的本发明方法提供若干途径,以允许对核酸分子的两条链上的靶区域进行直接捕获、富集及测序,这提供精确的校正和对极低丰度的突变或序列差异的证实,从而克服了当前测序技术的上述限制。

[0136] 因此,本发明的另一方面涉及一种用于鉴别样品中的多个核酸分子中的一个或多个的方法。具体说来,所述方法允许鉴别样品中的多个核酸分子中的一个或多个序列差异。此方法涉及提供含有一个或多个核酸分子的样品,所述核酸分子各自与其相应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分和适用于发夹形成的单链3' 适配体部分。样品经受有效使所述一个或多个核酸分子及其互补序列的3' 适配体部分形成发夹的条件。所述方法进一步涉及提供聚合酶并将所述一个或多个3' 发夹核酸分子及其互补序列与聚合酶混合以形成延伸混合物。所述一个或多个核酸分子及其互补序列是从其相应的3' 发夹区域延伸以形成一个或多个全长发夹核酸分子。所述方法还涉及为所述一个或多个全长发夹状核酸分子的至少一部分测序以鉴别样品中的一个或多个核酸分子中的一个或多个。

[0137] 根据本发明的此方面,所提供的样品含有一个或多个核酸分子及其互补序列,其

中含有一个或多个有待鉴别的潜在序列差异。样品中的核酸分子可被随机片段化并处理以便将适配体(即,适用于发夹形成的3' 适配体部分)附加到片段化的核酸分子的两端。例如,DNA分子的末端,不管是平端抑或变得平齐,都可使用多种酶如T4聚合酶或大肠杆菌聚合酶和T4激酶来磷酸化。不具有3' 至5' 校正活性的聚合酶(如Klenow(外))被用于将额外的“A”添加至3' 末端,形成单碱基3' 突出端。使用本领域中已知的多种酶促反应中的任一种将适配体部分附加到核酸分子及其互补序列。适合的酶包括不限于连接酶(例如,大肠杆菌连接酶或T4DNA连接酶)、聚合酶(例如,Taq聚合酶、T4聚合酶或大肠杆菌聚合酶)、重组酶、末端转移酶、核酸内切酶、DNA修复酶以及逆转录酶。

[0138] 根据本发明的此方面,3' 适配体部分含有适用于发夹形成的序列。另外,适配体部分可含有一个或多个独特的核苷酸序列标签,例如,患者鉴别序列标签和/或链鉴别序列标签。另一个任选特性是设计散布有随机序列的适配体,以使得一旦适配体延伸并变成单链,它们由于在各末端上的适配体彼此杂交而不会形成锅柄结构。这还减少如果使用具有5' - 3' 核酸外切酶活性的聚合酶,在3' 发夹状末端扩增期间偶然地破坏片段一侧上的适配体标识符区域的机会。

[0139] 仅举例来说,适用于本发明的此方面中的适配体部分描绘如下。在此实例中,任选的患者鉴别序列是6个碱基,即CGTGAC(允许4,096个不同的标识符标签)且以下划线标注。如所示,患者鉴别序列含有散布的随机核苷酸(即在患者鉴别区域(CGNTNGNANC;SEQ ID NO:1的核苷酸1-10)中的“N”碱基)。4个“N”的使用提供256倍多样性,从而明显地减少在片段的两端上具有相同适配体序列的机会。片段鉴别序列是10个碱基(允许1,048,576个不同的组合)并且由一串加框的N表示。发夹序列以粗体表示。较短的寡核苷酸(即SEQ ID NO:2)在3' 端上含有额外的“T”,以使得所述适配体适合于连接至A尾片段。



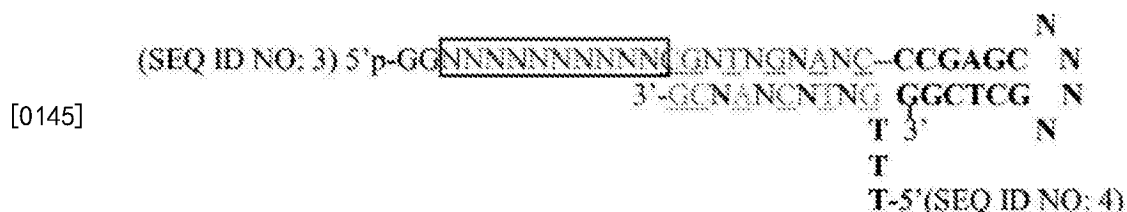
[0141] 寡核苷酸适配体是用针对每组所限定的患者标识符序列和发夹序列来合成,但以所有四个碱基的混合物形式添加其余的碱基以提供向每个片段给出独特的标识符组合所需的多样性。

[0142] 图8A是本发明的此方面的示意图。在此实例中,含有患者标识符序列和发夹状区域的适配体部分在步骤2中被附加到每个DNA片段(即,一个或多个核酸分子及其互补序列的每一个)的末端上。每个片段的3' 发夹状末端可使用具有5' 至3' 核酸外切酶活性或链置换活性的聚合酶来延伸(步骤3),从而将一个或多个核酸分子从其相应的互补核酸分子上离解。或者,3' 发夹状核酸分子可在由聚合酶延伸之前从其杂交的互补序列上解链。在此实施方案中,可利用缺乏核酸外切酶活性或链置换活性的聚合酶。在3' 发夹状末端的延伸之后,所生成的双链发夹状DNA分子理想地适于使用本发明装置进行测序。

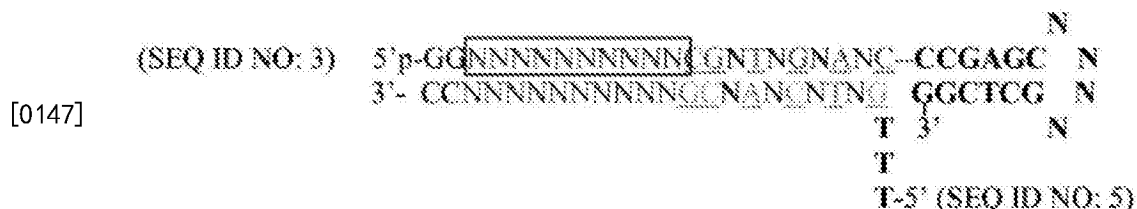
[0143] 将独特的序列标识符附加到核酸片段的两端上的一个替代手段最近已针对“双重测序(Duplex Sequencing)”而开发(Schmitt等人,“Detection of Ultra-Rare Mutations by Next-Generation Sequencing,”Proc Natl Acad Sci U S A 109(36):14508-13(2012),其在此以引用的方式整体并入)。此方法是基于使用含有12个碱基随机化标签的双

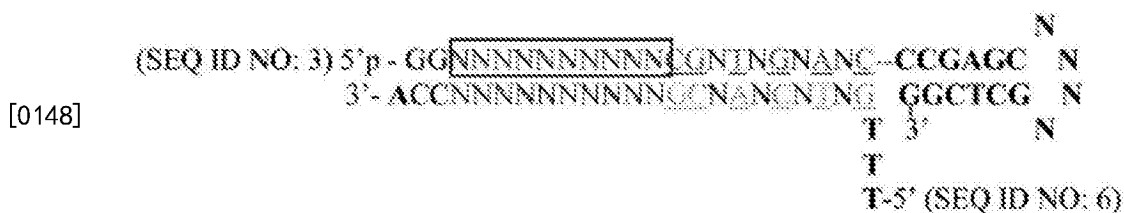
链体接头。通过扩增输入靶DNA的上链和下链,给定的片段获得独特的标识符(在各端上由12个碱基组成),以使得可经由测序对所述给定片段进行追踪。将共享一组独特的标签的序列读取(sequence read)集成成配对的家族,所述家族拥有在上链和下链方向上具有链标识符的成员。每对家族反映一个双链DNA片段的扩增。仅存在于一个或几个家族成员中的突变代表在扩增晚期出现的测序错误或PCR引入的错误。在一个成对家族的许多或所有成员中出现的突变是起因于第一轮扩增期间的PCR错误,如可能当跨过诱变的DNA损伤位点进行拷贝时出现的。另一方面,存在于DNA片段双链上的真正的突变出现在家族对的所有成员中。尽管人为突变可共同存在于具有真正突变的家族对中,但除在第一轮PCR扩增期间出现的那些以外的所有突变都可当产生错误纠正的单链共有序列时独立地被鉴别和忽视(discounted)。随后可将从个别DNA双链体的两条链的每一条处获得的序列相比较以获得双链体共有序列,所述共有序列消除了在第一轮PCR期间出现的剩余错误。此方法的缺点在于许多片段需要测序以得到家族对中每条链的至少五个成员(即,每个原始片段最少10个短读序列,但由于波动而经常需要多得多)。此外,所述方法不提供有关片段内的CpG位点的甲基化状态的信息。

[0144] 尽管如此,用于将独特的序列适配体附加到靶标末端的上述方法可被修改以用于本发明中。仅举例来说,适用于本发明的此方面中的一个替代的适配体部分描绘如下。在此实例中,任选的患者鉴别序列是6个碱基,即CGTGAC(允许4,096个不同的标识符标签)且以下划线标注。如所示,患者鉴别序列含有散布的随机核苷酸(即在患者鉴别区域(CGNTNGNANC;以下SEQ ID NO:3的核苷酸13-22)中的“N”碱基)。4个“N”的使用提供256倍多样性,从而明显地减少在片段的两端上具有相同适配体序列的机会。片段鉴别序列是10个碱基(允许1,048,576个不同的组合)并且由一串加框的N表示。发夹序列以粗体表示。较小的适配体寡核苷酸(即SEQ ID NO:4)含有5' OH或5' 尾部(即TTT,粗体的)以防止连接至在更长的适配体链上的发夹。



[0146] 在以下实例中,适配体的较小寡核苷酸的3'端是用聚合酶(即Klenow(外-))来延伸,所述聚合酶缺乏核酸外切酶或链置换活性。完整的适配体A-加尾是通过用聚合酶和dATP延长孵育来保证。





[0149] 寡核苷酸适配体是用针对每组所限定的患者标识符序列和发夹序列来合成,但以所有四个碱基的混合物形式添加其余的碱基以提供向每个片段给出独特的标识符组合所需的多样性。因为上述工序将A附加到寡核苷酸,所以靶DNA将在两侧由单个3'“T”碱基作为尾部。

[0150] 图8B是本发明的此方面的示意图。在此实例中,含有患者标识符序列和发夹状区域的适配体部分在步骤2中被附加到每个DNA片段(即,一个或多个核酸分子及其互补序列)的末端上。每个片段的3'发夹状末端使用具有链置换活性的聚合酶来延伸(步骤3)。所生成的双链发夹状DNA分子理想地适于使用本发明装置进行测序。此方法优于现有技术是因为(i)其不需要PCR扩增,(ii)其仅需要对每个片段的原始链进行测序,而不是10个拷贝,以及(iii)由于对原始链测序,所以其保护CpG甲基化信息。

[0151] 含有5'磷酸盐的全长发夹状双链DNA理想地适于使用本发明装置进行测序,特别是当裂解酶如 λ 核酸外切酶被连接于生物反应器室的固体支撑物时。发夹状区域为DNA片段定向提供机会以使得游离(5'磷酸化)末端最先遇到核酸外切酶。因为DNA被电泳转运穿过进料通道及朝向生物反应器室的熵陷阱,所以阻止移动中的含有发夹的末端,以使得相对侧(游离端)更快地移动,且因此首先到达生物反应器室中的核酸外切酶。另外,发夹序列可被设计以含有简单的重复序列(例如 A_n 或 CA_n),其适用于瞬时杂交至位于进料通道、熵陷阱或生物反应器室上游的其它室内的含有互补序列(即 T_n 或 TG_n)的固定的寡核苷酸。因为发夹区域将瞬时杂交至固定的互补序列并从固定的互补序列上解链,所以其在电泳场中比游离端更慢地迁移,由此将游离端与迁移的前部对齐以便其成为到达生物反应器室的第一末端。

[0152] 在一个替代实施方案中,发夹序列被设计以含有更长的已知序列,该已知序列适用于杂交至含有已知互补序列的寡核苷酸。互补寡核苷酸被联接到“迁移率阻碍物(mobility drag)”或“迁移率调节物(mobility modifier)”。可阻碍DNA迁移的迁移率调节物的实例包括不限于肽、蛋白质、PEG、其它高分子聚合物。在所述条件下,在形成含有有待鉴别的DNA的靶特异性区域的发夹状分子之后,将含有迁移率调节物的互补寡核苷酸例如经由杂交附加到发夹状分子上。负载有杂交迁移率调节物的发夹状末端在电泳场中比游离端更慢地迁移,提供进入生物反应器通道的另外的能障(除熵陷阱之外),其中游离端优先地进入。

[0153] 在本发明的一个相关方面中,如上所述用于鉴别样品中的多个核酸分子的一个或多个序列差异的方法还涉及在固体支撑物上固定核酸分子以实现一个或多个所需核酸分子的序列特异性富集。此方法涉及提供含有一个或多个核酸分子的样品,所述核酸分子各自与其相应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分、5'适配体部分及3'适配体部分,其中5'和3'适配体部分各自适用于发夹形成。所述方法还涉及使所述一个或多个核酸分子与其互补序列解链以形成一个或多个

单链核酸分子及其互补序列的集合,其中一个或多个单链核酸分子及其互补序列各自的5'和3' 适配体部分的发夹序列形成发夹。提供包括多个固定的捕获寡核苷酸探针的固体支撑物,所述捕获寡核苷酸探针包含与所述一个或多个核酸分子或其互补序列的靶特异性部分区域互补的核苷酸序列。所述一个或多个解链的单链核酸分子及其互补序列与被固定在固体支撑物上的其互补捕获寡核苷酸杂交。此方法还涉及提供聚合酶并将所述聚合酶与含有所述一个或多个杂交核酸分子及其互补序列的固体支撑物接触。所述一个或多个核酸分子及其互补序列从其相应的3' 发夹区域延伸以形成一个或多个全长发夹状靶核酸分子,从而使所述一个或多个核酸分子从捕获寡核苷酸且从固体支撑物上释放。对所述一个或多个全长发夹状核酸分子的至少一部分进行测序以鉴别样品中的所述一个或多个核酸分子。

[0154] 图9是本发明的此实施方案的示意图。在此实例中,将含有患者标识符序列标签、链标识符序列标签及发夹区域的适配体部分附加到目标双链DNA的末端上以使得发夹状序列被附加到每个片段的5' 端上(图9,步骤2)。双链DNA的3' 端含有接头区域。使用缺乏5' 至3' 活性的聚合酶将3' 接头端延伸整个发夹的长度以形成含有靶特异性部分、5' 适配体部分及3' 适配体部分的双链DNA产物,其中5' 和3' 适配体部分各自适用于发夹形成,如图9步骤3中所示。如图9步骤4所示,双链DNA被解链成单链DNA。有义和反义DNA链都是通过杂交至固定于固体支撑物上的互补捕获寡核苷酸来捕获。在微制造装置中的固体支撑物上的杂交时间是快速的,大约30-60秒,相比于常规阵列(大约3-5小时)(Wang等人,“Microarrays Assembled in Microfluidic Chips Fabricated From Poly(methyl methacrylate) for the Detection of Low-Abundant DNA Mutations,”*Anal.Chem.*75:1130-1140 (2003),其在此以引用的方式整体并入)。将互补链在空间上分开以避免它们之间的再杂交。

[0155] 固定的单链DNA的发夹区域形成发夹,且固定的DNA链的3' 发夹端使用具有5' -3' 核酸外切酶活性或链置换活性的聚合酶来延伸(步骤5,图9),其将双链发夹状DNA从固体支撑物上释放。或者,固定的核酸分子在延伸之前从其互补的捕获寡核苷酸上解链。在这种情况下,可利用缺乏核酸外切酶活性或链置换活性的聚合酶。

[0156] 本发明的另一方面涉及另一种用于鉴别样品中的多个靶核酸分子中的一个或多个或多个靶核酸分子中的序列差异的方法,其不涉及发夹形成,还允许对核酸分子的两条链上的靶区域进行直接捕获和测序。此方法涉及提供含有一个或多个核酸分子的样品,所述核酸分子各自与其相应的互补核酸分子杂交,其中所述一个或多个核酸分子及其杂交的互补序列各自包含靶特异性部分、5' 适配体部分及包含单链均聚物重复序列的3' 适配体部分。所述方法还涉及将一个或多个核酸分子从其互补序列上解链以形成单链核酸分子及其互补序列的集合。提供与一个或多个核酸分子或其互补序列的3' 端互补的多个寡核苷酸引物和聚合酶,并且将一个或多个核酸分子及其互补序列与聚合酶和引物混合以形成延伸混合物。所述方法还涉及将寡核苷酸引物杂交至其互补的一个或多个核酸分子及其互补序列的3' 端,并且延伸杂交引物以形成全长双链核酸分子。对所述一个或多个全长双链核酸分子的至少一部分测序以鉴别样品中的一个或多个核酸分子。

[0157] 图10是本发明的此方面的示意图。如所示,双链基因组DNA是可使用此方法鉴别的示例性核酸分子。基因组DNA被随机片段化并处理以便将适配体附加到分子的两端(图10,步骤2)。例如,如上所述将平端抑或变得平齐的DNA的末端磷酸化,并且聚合酶被用于将额外的“A”添加至3' 端,形成单碱基3' 突出端。将含有例如链标识符序列或标签的适配体部分

附加到DNA片段的各端上(图10,步骤2)。使用聚合酶延伸3' 适配体末端,接着用末端转移酶作为尾部(图10,步骤3)。加尾用于引入均聚物重复序列如polyA尾部到3' 适配体部分的末端中。

[0158] 双链DNA解链并且将引物杂交至加尾的适配体链的3' 端(图10,步骤4)。杂交引物是使用聚合酶来延伸以生成使用本发明装置理想地适于测序的双链DNA。测序链含有在两端上的链标识符序列,从而允许其来源的独特验证。另外,由于对原始基因组链测序,所以所有的碱基修饰得以保留。

[0159] 为了实现具体的DNA片段的靶特异性富集,可采用如上所述的方法的一个变型,其涉及在固体支撑物上的选择性靶-序列特异性捕获。此实施方案的示意图显示于图11中。适配体部分如上所述被附加到DNA片段的各端上(图11,步骤2)。在此实例中,适配体部分还含有患者标识符序列标签。用聚合酶延伸3' 适配体末端,且如上所述使用转移酶作为尾部(图11,步骤3)。将双链DNA解链成单链DNA。上靶链和下靶链是通过杂交至固定在固体支撑物上的互补寡核苷酸探针来捕获(图11,步骤4)。将互补链在空间上分开以使其或是靶链不会彼此再杂交。

[0160] 将寡核苷酸引物杂交至固定的单链DNA的3' 端,并且使用聚合酶进行延伸以生成理想地适用于测序的双链DNA分子(图11,步骤5)。在本发明的一个实施方案中,将固定到固体支撑物上的单链DNA在引物延伸之前从其互补的捕获探针上解链。或者,单链固定的DNA在引物延伸期间通过使用链置换聚合酶(留下完整的互补捕获链)或使用具有5' →3' 核酸外切酶活性的聚合酶(其消化掉互补捕获链)而释放。

[0161] 在本发明的此方面中使用的寡核苷酸引物与链标识符部分的已知序列以及适配体段的均聚物重复序列重叠。引物不具有5' 磷酸盐且相对于均聚物重复序列的3' 端凹进,因此这个末端不是当使用本发明装置对所生成的全长双链DNA测序时用于像λ核酸外切酶的裂解酶的底物。

[0162] 根据本发明的此方面,3' 单链均聚物重复序列(例如A_n)提供在本发明的生物分子处理器内对片段定向的机会以使得平(5' 磷酸化)端将首先遇到裂解酶,即核酸外切酶。因为DNA被电泳转运穿过进料通道及朝向生物反应器室的熵陷阱,所以阻止移动中的含有3' 单链均聚物重复序列的末端,以使得游离端具有更高的首先进入生物反应器室并遇到核酸外切酶的概率。互补均聚物序列(即T_n)可固定在进料通道或熵陷阱内以使得3' 单链均聚物重复序列将瞬时杂交至并从其上解链,从而造成3' 单链端在电泳场中比游离端更慢地迁移。或者,与3' 单链均聚物重复序列互补且还含有用作电场中的迁移率阻碍物的附接体的更长的寡核苷酸可如上所述用于减慢3' 端的迁移。

[0163] 含有上述序列设计的附加的适配体提供基于以下来独特地鉴别每个核酸的超额能力:(i) 5' 侧上的随机片段ID序列;(ii) 5' 侧上的患者ID序列;(iii) 靶序列的5' 侧上的接合位点;(iv) 捕获的靶序列;(v) 靶序列的3' 侧上的接合位点;(vi) 3' 侧上的患者ID序列;以及(vii) 3' 侧上的随机片段ID序列。两个随机序列(假定各自有10个碱基)单独提供超过 1×10^{12} 程度的多样性,这足以将靶DNA与100,000个基因组当量独特地区分。因此,真正的突变将通过基于同一序列的上链和下链两者鉴别相同的突变体来验证。同样地,在给定的CpG二核苷酸中真正的5-甲基-dC或5-羟甲基-dC将表现为在给定的CpG二核苷酸的上链和下链两者上存在5-甲基-dCMP或5-羟甲基-dCMP碱基识别。

[0164] 高一致性准确度可在高误差率存在下通过对DNA的同一区域反复地测序而实现。然而,高误差率使得极其难以鉴别低丰度的序列变体,例如当试图在正常的DNA存在下鉴别癌症突变时。因此,需要低误差率来检测相对低丰度的突变。

[0165] 本发明的装置和方法提供实现精确地检测低丰度突变所需的低误差率的许多方法,即,本发明方法能够区分含有一个或多个核苷酸碱基插入、缺失、易位、突变或损伤的碱基的低丰度的核酸分子(例如RNA和DNA)与在样品中具有与低丰度核酸分子类似的核苷酸序列但没有相应的一个或多个核苷酸碱基插入、缺失、易位、突变和/或损伤的碱基的100倍至10,000倍或更大过量的核酸分子。例如,在一种方法中,利用 λ 核酸外切酶作为裂解酶来消化dsDNA的5'端,其中两条链含有独特的序列标识符(如上所述的适配体),所以其可鉴别为由同一原始靶DNA产生。在此方法中,用于对两条链测序的误差率显著地低于只有一条链时,因为在一条链上的误识别(其匹配另一条链上相同位置处的互补误识别)的概率比仅仅任何误识别低4倍。误差率和准确度显示于下表1中。注意此方法保持不依赖于哪种酶被用于生成单核苷酸。例如,在将适当的适配体附加到样品的双链DNA的两端的独特序列标识符之后,DNA可被解链并变成单链,其适用于用核酸外切酶I从3'端上消化,且适当的配偶链是由其独特的序列标识符来鉴别。

[0166] 表1. 含有链标识符的dsDNA的 λ 核酸外切酶消化和两条链的分析

[0167]

误差率	准确度	误差率	准确度
SS序列		DS序列	
10.0%	90.0%	0.2500%	99.7500%
4.0%	96.0%	0.0400%	99.9600%
1.0%	99.0%	0.0025%	99.9975%
0.5%	99.5%	0.0006%	99.9994%
0.2%	99.8%	0.0001%	99.9999%

[0168] 另一种方法涉及使用 λ 核酸外切酶来消化dsDNA的5'端,接着用核酸外切酶I将所生成的单链DNA从3'端上消化。根据此实施方案,生成的单链DNA将被引入通向含有核酸外切酶I的第二生物反应器室的第二进料通道中。消化的单核苷酸将穿梭通过第二飞行时间通道以分隔和鉴别由第二消化释放的核苷酸。以下的计算假定在区分飞行时间通道中的核苷酸时的误差率是不同的,尽管是由不同的核酸外切酶生成。

[0169] 表2.dsDNA靶分子的第一和第二链的偶联的 λ 核酸外切酶/核酸外切酶I消化

[0170]

误差率	准确度	误差率	准确度
SS序列		DS序列	
10.0%	90.0%	0.2500%	99.7500%

[0171]

4.0%	96.0%	0.0400%	99.9600%
1.0%	99.0%	0.0025%	99.9975%
0.5%	99.5%	0.0006%	99.9994%
0.2%	99.8%	0.0001%	99.9999%

[0172] 另一种方法涉及与双感测(三对纳米电极以获得三个飞行时间测量结果)相比,使用单感测(两对纳米电极以获得单个飞行时间测量结果)的dsDNA的 λ 核酸外切酶消化。在这种情况下,比较是在仅裂解掉一条链的相同核苷酸的相同潜在误差率之间进行。因此,误差率稍微高于以上分析同一原始靶分子的两条链的方法。

[0173] 表3.dsDNA的 λ 核酸外切酶消化,比较单和双飞行时间测量结果

[0174]

误差率	准确度	误差率	准确度
单感测		双感测	
10.0%	90.0%	1.0000%	99.0000%
4.0%	96.0%	0.1600%	99.8400%
1.0%	99.0%	0.0100%	99.9900%
0.5%	99.5%	0.0025%	99.9975%
0.2%	99.8%	0.0004%	99.9996%

[0175] 甲基化或羟甲基化碱基与其未甲基化的对应物(即d5meCMP比对dCMP)的区分可比仅将两个天然碱基彼此区分更困难。因此,将d5meCMP与dCMP区分的分类误差率可更高。以下表4中示出区分甲基化与未甲基化C碱基的所计算的误差率。

[0176] 表4.甲基化、单感测与双感测的比较

[0177]

误差率	准确度	误差率	准确度
单感测		双感测	
20.0%	80.0%	4.0000%	96.0000%

[0178]

10.0%	90.0%	1.0000%	99.0000%
5.0%	95.0%	0.2500%	99.7500%
2.0%	98.0%	0.0400%	99.9600%
1.0%	99.0%	0.0100%	99.9900%

[0179] 生物上相关的甲基化经常存在于称为CpG岛(其具有高含量的G+C)的区域以及称为CpG岸的相邻区域中。因此,使用如CpG序列的插入物,且邻近CpG序列的甲基化状态的评分显著地改善在对5meC评分中的准确度,即使在单一飞行时间测量中将d5meCMP与dCMP区分的误差率较高也如此。这在以下表5中说明:

[0180] 表5.甲基化、单感测与双感测的比较、以及对于连续甲基化CpG测序的评分

[0181]

误差率	准确度	误差率	准确度
单感测		双感测	
20.0%	80.0%	0.1600%	99.8400%
10.0%	90.0%	0.0100%	99.9900%
5.0%	95.0%	0.0006%	99.9994%

[0182] 同样地,将双感测与确定两条链的序列组合,如上所述还明显降低误差率并改善识别各种启动子区域中的甲基化状态的准确度。因此,本发明方法能够区分具有一个或多

个未甲基化或未修饰的核苷酸碱基的低丰度核酸分子与样品中具有与低丰度核酸分子类似的核苷酸序列且具有甲基化或修饰的核苷酸碱基的100倍至10,000倍或更大过量的核酸分子。同样地,本发明方法能够区分具有一个或多个甲基化或修饰的核苷酸碱基的低丰度核酸分子与样品中具有与低丰度核酸分子类似的核苷酸序列且具有未甲基化或未修饰的核苷酸碱基的100倍至10,000倍或更大过量的核酸分子。

[0183] 本文所述的装置和方法将克服困扰着当前可用的测序技术的上述问题。本发明的装置和方法可鉴别来源于血浆中的无细胞DNA中的肿瘤DNA的100个基因中一个或多个突变的存在或100个启动子中一个或多个甲基化变化,其中含有突变或甲基化变化的肿瘤DNA代表在正常DNA的100,000个基因组中的10个基因组。考虑到平均基因代表序列的2,000个碱基,且平均启动子区域代表500个碱基,那么将合成捕获探针(平均长度50个碱基)以捕获独特序列的250,000个碱基。那个代表5,000个探针以捕获“上”链序列和5,000个探针以捕获“下”链序列。可大批地合成每组5,000个探针(或在Agilent芯片上并从芯片上裂解),接着附接于微制造装置的空间上分开的区域,例如,在用于寡核苷酸附接的芯片的那个区域中的柱形或其它高纵横比特征上。

[0184] 随着患者年龄增大,一些基因由于年龄相关的启动子甲基化而偶发地关闭。当肿瘤扩展时,其经常在特定的启动子区域展现全面的甲基化不足和特定的超甲基化。甲基化状态的这两种变化提供在血浆中的癌症特异性DNA的早期检测的机会。重要的是鉴别具有癌症特异性甲基化变化的多个基因座以避免由于年龄相关变化造成的伪造的假阳性。本文所述的装置和方法提供鉴别肿瘤组织中的甲基化变化的快速途径,以及用于鉴别在血浆或其它生物样品中的低丰度变化的高度精确的方法。

[0185] 本发明的方法和装置提供与用于核酸分子测序的先前方法和装置相比的许多优点。具体说来,本发明方法实现具有从第一个到最后一个碱基的高准确度水平的长读取长度(200至50,000个碱基),每次运行在低成本下均匀覆盖和快速处理量。所述装置可对100,000个基因组当量中的两条链(=500,000个碱基)上的250,000个碱基测序。这等于DNA序列的 $5 \times 10^{10} = 50\text{GB}$ 。在单个1.4小时运行中最大500GB DNA的容量下,本发明装置可同时评估10份血清样品以提供100个癌症基因和100个启动子区域的突变状态的综合评价,这些基因或区域的甲基化状态与癌症或结果状态有关。如果仅存在20个基因和20个启动子区域的筛选测试,那么将在单一运行中容易地评估50份血清样品。

[0186] 使用本发明装置和方法获得的测序读取不具有困扰着其它测序技术(例如,高GC;均聚物延伸段)的情境特异性错误。实现单分子分辨率的能力促进不均匀样品的表征并且使得能够鉴别其它测序技术所看不见的变化。两条链上的靶区域的直接捕获和测序提供精确的校对和极低丰度突变的证实。对核苷酸序列的长延伸段测序的能力可以高灵敏度和特异性解析SNP、单倍型以及大规模的结构重排。长读取还通过减少毗连群的数目并在相同的覆盖深度下产生与较短的读取测序技术相比更好的一致性准确度来简化和改善基因组装配。最终,在无扩增下的输入DNA的直接测序能够在没有亚硫酸氢盐转化下鉴别原始碱基修饰(例如,5-甲基-C、5-羟基-甲基-C或损伤的DNA碱基)。与现有的单分子DNA测序策略相反,本发明装置具有在单一测序运行中生成对天然和非天然碱基的高度精确识别的能力,这归功于飞行时间鉴别方案以及由每个单体单元生成的独特电子签名。

[0187] 本发明的装置和方法提供将能够进行用其它可用的测序仪器是不实际或不可能

的临床研究和诊断测试的低成本测序平台。例如,对从血液(例如,来自血浆、外来体或循环肿瘤细胞)分离的核酸分子进行测序将提供(i)通过高灵敏度检测启动子超甲基化和甲基化不足(当以1%至0.01%的无细胞DNA存在时)来精确鉴别各种遗传相关疾病(例如癌症), (ii)通过高灵敏度检测已知基因中的单碱基突变、小插入和小缺失(当以1%至0.01%的无细胞DNA存在时)来精确鉴别遗传相关疾病, (iii)通过量化从血液中分离的特定的mRNA和miRNA来精确鉴别早期疾病发作, (iv)精确鉴别DNA中的特定拷贝变化,以及(v)通过量化从具体的循环疾病相关细胞中分离的DNA中的突变、启动子超甲基化和甲基化不足来精确鉴别早期疾病。

[0188] 本发明的生物分子处理装置和方法特别适合进行多种诊断测定。具体说来,本发明的生物分子处理装置和方法最适用于使用母体血清样品进行产前诊断测定。近期研究显示血清中的胎儿DNA(作为母体DNA的百分比)在第一、第二和第三个三个月分别近似是6%、20%和26%。由于DNA的降解方式,母体DNA通常为约160个碱基且仍与H1组蛋白有关,而胎儿DNA为约140个碱基且与组蛋白不相关。取决于临床需要及知识将提供最佳护理的地方,可开发具有足够灵敏度的测试以在适当的三个月中检测胎儿DNA。

[0189] 存在近似3,500种其中基因是已知的隐性遗传病症。最常见的病症是由DNA拷贝异常引起,额外的染色体,如在21三体中,抑或是缺失基因的一部分,如在杜兴氏肌营养不良(DMD)基因中。在考虑产前筛选时,需要平衡遗传病症的概率与工序的风险。目前,对于年龄在35岁的待产妇,在第17周期间的标准护理推荐羊膜穿刺术,由于21三体或在200中的一个处的其它染色体非整倍体的风险现在匹配工序之后的自然流产风险。

[0190] 在考虑使用本文所述的核酸测序方法进行产前筛选时,推荐两级测试。在就21、13及18三体低成本筛选所有孕妇时,本发明的测序方法可用于快速地鉴别染色体21、13及18上差异表达的基因,例如,鉴别由于甲基化沉默而在胎儿中关闭但在成人中打开的那些基因。在三个对照染色体即染色体2、5、7上鉴别相似区域。即使当在第一个三个月从母体血清中分离DNA时,仍可通过比较对照染色体区域之中甲基化与未甲基化的DNA快速地计算源自胎儿的DNA的百分比-在本文的实施例中,这将是6%。如果在任何其它染色体处存在三体性,即21三体,那么来自那个染色体的启动子将显示约9%的甲基化,换言之,有些高出正常二体性情况50%。推荐1,000个基因组当量的评分,以使得三体性情况的90个甲基化拷贝计数轻易与正常样品的60个甲基化拷贝相区分。认为分析10个启动子区域 $\times$ 6个染色体 $\times$ 500个碱基 $\times$ 2条链 $\times$ 1,000个拷贝=60MB DNA序列/样品。在单个1.4小时运行中最大500GB DNA的容量下,本发明装置能够同时评估大约8,300份血清样品。

[0191] 为了确定胎儿是否含有与大致3,500种其它病症有关的遗传或自然突变,包括缺失、点突变或异常甲基化,将推荐更复杂的分析。序列分析容易确定隐性等位基因在父母双方中的存在。如果父母中的突变是不同的,那么有可能通过评估来自母体血清的无细胞DNA确定孩子是否是所述疾病的复合杂合体。获得来自母体血清中的胎儿DNA分析的完整答案可能需要两部分测定。首先是在围绕疾病基因的重复区域中建立母体SNP或多态性的相。这可通过从母亲的白细胞中或从父亲的唾液中分离高分子量DNA来完成。由于本发明的核酸测序方法能够对DNA的长延伸段测序,所以这可用于在与疾病基因以及染色体相(单倍型)有关的重复区域中建立相邻的SNP或多态性。如果存在缺失,例如母亲是DMD基因缺失的携带者,那么也将建立。

[0192] 第17周,将考虑羊膜穿刺术的时间,母亲早已在第二个三个月中,即,胎儿DNA以15%-20%范围存在。在第一种方法中,在单个2.8小时运行中最大1,000GB DNA的容量下,本发明装置将评估330个基因组当量,并且如果15%来源于胎儿,那么将为约50个基因组,或每个单倍型的覆盖率为约25。此方法是最全面的且能够检测由遗传变化以及胎儿中的从头变化引起的遗传疾病。此方法不仅将检测DMD缺失,而且检测其它小拷贝数异常,如引起孤独症的那些。每200个孕妇中有约1个导致平衡易位,且此方法可精确地测定接合位置,以确定潜在的临床表现。而且,此方法将能精确地定量可导致与印记有关的疾病如安格曼综合征或普-威综合征的甲基化变化。本发明测定甲基化状态且同时通过SNP检测(即,检测顺位的母系或父系鉴别SNP的上游或下游)确定缺失是否在父系或母系染色体上的能力将增强印记疾病的诊断辨别。最终,因为本文所述的方法不需要克隆或PCR扩增,所以可对引起如亨廷顿病和脆性X染色体综合征的疾病的三联体重复变化恰当地评分。本发明还进行蛋白质和多肽鉴别及如下所述的定量蛋白质组学的能力将允许其用于测量母体血清中的胎儿蛋白质,这是胎儿非整倍体和神经管缺陷的产前筛选的一部分。

[0193] 或者,在第二个方法中,疾病基因可被分成20种最常见的遗传疾病,然后被分成17组不太常见的突变序列,每个覆盖平均200个基因。每组基因将由两条链的捕获探针组覆盖,然后取决于亲本测序分析的结果,母体血液将被给出适当的患者标识符并在17个专用芯片的一个或多个上评估。如果给定的基因在长度上平均为约2,500个碱基,那么对于200个基因,两条链来说,在芯片上是1MB捕获探针。如果1,000个基因组当量有待被测序,那么一份血清样品将需要1GB序列。在单个1.4小时运行中最大500GB DNA的容量下,装置将在单次运行中评估500份血清样品。

[0194] 上述第一个方法将鉴别遗传和自然突变,并且确定胎儿是否从母体遗传带有突变的区域。此方法还能够确定针对x连锁遗传疾病的缺失、其它染色体缺失、胎儿中的异常甲基化、由三联体重复引起的疾病以及由染色体易位或其它重排引起的疾病的存在。

[0195] 第二个方法将鉴别针对所检查的基因的疾病状况。关键问题将是家庭获得正确答案的重要性如何。可以直接确定父母双方是否是携带者,且如果突变是不同的,那么确定父亲的疾病等位基因是否存在于胎儿中是相对简单的。如果其不存在,那么胎儿将无疾病或是携带者。如果其存在,那么遗传母体等位基因并患上疾病的机会是50%。如果已确定母体等位基因的单倍型,那么单倍型标志物可用于验证遗传的母体等位基因的存在或不存在。还可以谨慎地进行羊膜穿刺术并直接测试母体等位基因的存在。当前的推荐是对如上所述的基因进行测序,并且对父体疾病等位基因进行评分。如果存在,或如果父体和母体疾病特异性突变是相同的,那么医师推荐羊膜穿刺术。

[0196] 本发明的方法和装置可实现长的读取长度,从而使其理想地用于非侵入性产前诊断和染色体易位的植入前遗传诊断(PGD)。携带染色体易位的个体处于不孕、流产、死产和/或生下具有先天缺陷的儿童的风险中。植入前遗传诊断能够区分具有正确数量的遗传物质(平衡/正常)的胚胎与由于易位而失去遗传物质(失衡)的胚胎。其中一个成员是易位携带者的许多夫妇经历了流产或当获悉孕妇具有失衡的染色体组时必须面对艰难的决定。基于本发明的方法和装置的PGD将通过事先知晓所移植的胚胎具有平衡的染色体易位而降低不得不处理这些特殊情况的可能性。

[0197] 对输入双链DNA的两条链的测序允许在碱基识别中的最高水平的准确度。在本质

上是单链的RNA底物不提供分析两条链的机会。然而,在大多数情况下,完全正确的RNA序列可从基因组序列来推断——这又可通过对同一输入片段的两条链测序来确定,从而提供最高准确度。因此,RNA测序可提供极其重要的信息,如在字面上计数每个信使RNA,包括存在于样品中的所有剪接位点变体、所有长的非编码RNA(lncRNA)分子、所有微小RNA(miRNA)分子——其是来自从血清中分离的循环肿瘤细胞或癌症特异性外来体。因此,对单链RNA底物的测序不要求与双链DNA底物相同的高保真度,但是因为读出仍是基于飞行时间通道检测,所以定量RNA中的不寻常的或修饰的碱基的能力得以保留。

[0198] 本发明的装置在对所有mRNA测序中存在与其它技术相比的优点。当前的微阵列技术可鉴别不同外显子在转录产物中的表达,但缺乏精确地测定或定量所有剪接位点变体的能力。当前的新一代mRNA定量产生在3'端上的表现和cDNA的序列短片段,其然后被用于计数每个mRNA种类的拷贝数目。相比之下,本文所述的当前实施方案允许对整个原始mRNA链进行测序,这不依赖于其是否是14,000个碱基还是400个碱基。此方法允许装置确定癌细胞中的正常剪接和异常剪接,具有鉴别疗法的肿瘤特异性靶标的潜能。

[0199] 使用本发明的方法和装置区分RNA修饰(这不容易通过现有技术完成)的能力将帮助变革我们对基因表达和疾病起因的认识。近期研究显示高达20%的人mRNA被常规地甲基化。此外,最近还显示RNA甲基化的缺陷可能招致疾病。例如,脂肪质量和肥胖相关(FTO)基因编码能够逆转RNA甲基化的酶。FTO突变可引起摄食和代谢的异常,从而导致肥胖。FTO突变是肥胖和2型糖尿病的主要原因。近来有证据显示RNA修饰还可在癌症和神经病症如孤独症、阿尔茨海默氏病及精神分裂症中发挥作用。

[0200] 本发明的另一方面涉及使用本发明装置鉴别样品中的一种或多种蛋白质或多肽的方法。此方法涉及将包含一种或多种蛋白质或多肽的样品在一定条件下进料至装置的生物分子处理器中,所述条件有效使固定在生物反应器室内的裂解酶接合样品中的所述一种或多种蛋白质或多肽并将所述一种或多种蛋白质或多肽裂解成较小肽片段以进入一个或多个飞行时间通道的输入端。跨一个或多个生物反应器室并沿一个或多个飞行时间通道的长度施加电场以将裂解的肽片段转运穿过一个或多个飞行时间通道,并且基于所施加的电场,当肽片段经过一个或多个飞行时间通道中的至少第一和第二传感器时对其进行检测。所述方法还涉及基于所述检测测量(i)每个肽片段经过一个或多个飞行时间通道的至少第一和第二传感器所花费的时间,和/或(ii)当每个肽片段经过一个或多个飞行时间通道中的第一或第二传感器中的至少一个时的电力峰值振幅。基于所测量的肽片段鉴别样品中的一种或多种蛋白质或多肽的至少一部分。

[0201] 根据本发明的此方面,将针对裂解的肽片段所获得的测量结果(即,每个肽片段的飞行时间和电力峰值振幅)与含有通过裂解已知的蛋白质或多肽产生的肽片段的相应测量结果的数据库相比以鉴别或表征样品中的一种或多种未知的蛋白质或多肽。

[0202] 人基因组已经鉴别了约21,000个蛋白质编码基因,且使用替代剪接,这些产生约100,000个不同的细胞蛋白质。大多数的这些蛋白质经历翻译后修饰,如磷酸化、乙酰化及甲基化,并且所述修饰经常在蛋白质的功能和活性状态中发挥关键作用。

[0203] 在本发明此方面的一个实施方案中,分析的复杂性可通过添加初始步骤以从含有多种蛋白质的样品中分级分离或富集所需蛋白质而降低。此分级分离步骤通过将离散模块集成到使用模块集成方法的本发明装置中来容易地添加。此分级分离步骤可取决于物理性

质,如按大小、电荷或疏水性进行分离,使用例如色谱法。分级分离还可取决于对抗体、小分子或大分子的亲和力。例如,抗体富集步骤或分离所述蛋白质与给定的转录因子(也被特异性抗体拉下(pull down))有关。或者,蛋白质可通过结合至磷酸化、甲基化或乙酰化肽家族、或其它蛋白质结构域、或甲基化DNA、或特异性DNA序列、或DNA序列家族、或常用底物如ATP或GTP或其类似物中的一个或多个来富集,并且可通过在含有共价附接的底物或其类似物的固体支撑物上的亲和力纯化来分级分离。分级分离输入蛋白质的其它方法包括通过亚细胞区室如细胞核、细胞质、线粒体、内质网、高尔基体、细胞膜、核膜、溶酶体的分离,或与DNA、RNA、染色质、其它蛋白质支架联合。分级分离可被设计来富集一组蛋白质,包括但不限于:G-蛋白偶联受体、核受体、电压门控离子通道、配体门控离子通道、受体酪氨酸激酶、生长因子、蛋白酶、序列特异性蛋白酶、磷酸酶、蛋白激酶、生物活性脂质、细胞因子、趋化因子、泛素连接酶、病毒调节剂、细胞分裂蛋白、支架蛋白、DNA修复蛋白、细菌核糖体、组蛋白脱乙酰酶、细胞凋亡调节剂、伴侣蛋白、丝氨酸/苏氨酸蛋白激酶、细胞周期蛋白依赖性激酶、生长因子受体、蛋白酶体、信号蛋白复合物、蛋白/核酸转运蛋白以及病毒衣壳。

[0204] 当将个别蛋白质进料至生物反应器室中时,固定的裂解酶(即蛋白酶)消化蛋白质来完成,并且使用毛细管电泳或毛细管电色谱法将片段在一个或多个飞行时间纳米通道中分离并检测。此分离和检测形成用于给定蛋白质的指纹。如果某一残基被修饰,例如赖氨酸被乙酰化,那么那个片段的迁移率将与含有非乙酰化赖氨酸的相同片段相比被修饰。或者,乙酰化可防止一般将出现的在那个位置处的酶裂解(例如,胰蛋白酶不会在乙酰化的残基处裂解),从而改变肽指纹。使用含有从具有已知修饰的已知蛋白质上裂解的肽片段的测量结果或指纹的数据库,可以将针对样品中的未知蛋白质/多肽所获取的图案与数据库相比较以鉴别和表征样品中的蛋白质/多肽。此实施方案的独特方面在于其是在单分子水平下进行,因此,可观察到使用利用许多不同形式的质谱分析法的常规蛋白质处理方案不能显现的罕见修饰。

[0205] 利用一个或多个具有两个或更多个感测电极对和/或两个或更多个飞行管的飞行时间纳米通道,所述飞行管具有不同的壁表面化学特性以便每通道生成一个或多个飞行时间测量结果,从而形成十分复杂的指纹,例如2-维、3-维或4-维指纹以区分基本上所有可能的蛋白质及其修饰形式。用于生成具有两个或更多个具有不同壁表面化学特性的飞行管的n维指纹的装置结构的实例提供于图7中。修饰包括但不限于甲基化、乙酰化、磷酸化、糖基化以及苏氨酸或酪氨酸的一个或多个氨基酸残基的任何其它翻译后修饰。

[0206] 对于在区分个别蛋白质中有效的蛋白质指纹分析,蛋白酶需要在大多数时间产生相同的片段——即,其需要保证特定的裂解。实现此目的的最容易的方式是使用需要在识别袋中的特定类型残基的蛋白酶。适用于本发明中的特殊蛋白酶如上所述。

[0207] 使用本发明的方法和装置的蛋白质/多肽分析可用于精确地鉴别,即诊断和预测由从亚细胞组分富集的样品中蛋白质水平的变化或蛋白质的共价修饰引起的疾病,关于与细胞DNA、RNA、染色质、或其它蛋白质支架有关的蛋白质、结合底物或其类似物的蛋白质、或结合至大分子如蛋白质、碳水化合物、脂质、甲基化DNA、特定DNA序列或DNA序列家族的蛋白质。

[0208] 本发明装置在蛋白质分析中的使用将在推进对广泛多种健康相关研究领域的认知和理解中发挥决定性作用,包括以下研究:繁殖和寿命机制、病原细菌、不孕、癌症转化、

糖尿病、年龄相关疾病、克罗恩氏病、阿尔茨海默氏病、胰岛素抗性、致病真菌进展以及前列腺癌的预后。本发明还将关注猎枪定量蛋白质组学的成果以便最大化蛋白质序列覆盖率、实现在复杂样品中对于低丰度蛋白质的精确定量以及高置信度鉴别。本发明方法将提供在不同组织、细胞及细胞器内蛋白质复合物的精确定量以及蛋白质修饰的动态调节的灵敏测量,其中目标蛋白质和复合物经常具有极低丰度且在低动态范围内。

[0209] 本发明的方法和装置可用于检测血清标志物以区分良性与恶性前列腺疾病。前列腺癌(CaP)是男性中最常见的非皮肤恶性肿瘤。前列腺特异性抗原(PSA)筛选的引入具有明显降低的死亡率,但使用PSA作为诊断血清标志物远非理想的。本发明可用于检测另外的血清标志物,所述标志物预测CaP的存在和进展并且以高灵敏度和特异性区分良性与恶性疾病。具体说来,本发明可用于检测前列腺特异性生物标志物,所述标志物在容易获得的生物流体中是可检测的并且区分正常、良性前列腺肥大(BPH)、前列腺上皮内瘤样病变以及恶性前列腺组织。在良性与恶性疾病之间展现改变的表达并且具有补充或实现PSA作用的潜能的血清生物标志物包括血小板反应蛋白-1(TSP-1)、锯齿蛋白-2(EN2)、BCL-2、Ki-67、ERK5、嗜铬粒蛋白-A(CgA)、人腺激肽释放酶2(hK2)、尿激酶型纤溶酶原激活物(u-PA)、转化生长因子- β 1(TGF- β 1)、白细胞介素-6(IL-6)、前列腺膜特异性抗原、前列腺特异性细胞抗原、 α -甲基酰基-CoA消旋酶(AMACR)自身抗体、早期前列腺癌抗原(EPCA)、GSTP1超甲基化、性激素以及性激素结合球蛋白(SHBG)。例如,EPCA和hK2是可能的诊断因子且CgA、IL-6及TGF- β 是可能的预后预测因子。没有单个标志物能实现所需的灵敏度和特异性水平。标志物的组合或血清分布将改善CaP筛选的诊断准确度以及预后准确度以便改善治疗。本发明方法和装置的高灵敏度和高采样率使其成为此多因子诊断和预后方法的理想选择。本发明在检测和鉴别提供前列腺癌的准确或早期诊断的血液蛋白质标志物中的使用具有减少不必要的前列腺活组织检查数量并且明显改善前列腺癌的治疗和管理的潜能。

[0210] 本发明的方法和装置可用于鉴别在患有阿尔茨海默氏病的患者的不同组织中的N-糖基化蛋白质位点。N连接的糖基化是低聚糖附接至氮原子,通常天冬酰胺残基的N4,且出现在分泌或膜结合的蛋白质上。已通过早老素-1建立位点特异性糖基化与阿尔茨海默氏病的直接联系,早老素-1是 γ 分泌酶蛋白质复合物的一部分,其是将淀粉样蛋白前体蛋白(APP)加工成短的阿尔茨海默氏病相关肽淀粉样蛋白 β 中所涉及的一种蛋白酶。本发明的方法可用于构成以下发现:蛋白质上的N-糖基化位点可在阿尔茨海默氏病中发挥重要作用并且到目前为止理解并改善疾病的诊断和预后。

[0211] 目前,慢性肝脏病症的诊断需要肝活组织检查以评定纤维化阶段和坏死性炎症活性并检测肝硬化。在这些患者的大的亚组中,肝硬化是发展肝细胞癌的主要风险因素。因此,存在对血清标志物的需要,其可常规地评定肝纤维化的进展并可靠地检测肝硬化的阶段。总血清蛋白质N-糖组是用于检测肝硬化的优良生物标志物。本发明的方法和装置可用于以高特异性和良好的灵敏度检测肝硬化的血清标志物,如N-糖组,且因此在肝硬化患者中提供活组织检查的替代方案。

[0212] 本发明的方法和装置还可用于评价治疗性蛋白质产物。单克隆抗体是重要的治疗剂,靶向从癌症到神经变性病范围内的多种疾病。本发明方法可用于发展阶段和临床使用之前,用于表征这些分子,目的是向当前分析技术提出重大挑战。此外,使用重组DNA技术产生的治疗性蛋白质一般是复合物、非均质的,且在表达、纯化和长期保存期间经受多种酶

促或化学修饰。本发明在评价重组蛋白质序列中的使用可提供关于氨基酸修饰和序列变化的信息,它们具有影响治疗性蛋白质产物的可靠性和活性的潜能。所述装置可用于评价由翻译后修饰引起的重组蛋白质异质性、由蛋白水解或转录/翻译错误产生的序列变化以及在加工或最终产物储存期间形成的降解产物。本发明的方法和装置可用于分析单克隆抗体糖基化、脱酰胺、以及对于广泛多种非均质治疗性蛋白质产物的二硫键定位。

[0213] 本发明的方法和装置还可用于检测食物传染性疾病。细菌单核细胞增多性李氏杆菌在美国每年造成约1,600人的李氏杆菌病病例和250例死亡,代表美国约10%的所有因食物传染性疾病而死的病例。本发明可用于鉴别单核细胞增生李斯特氏菌蛋白,所述蛋白共同调节促进传播和毒力的基因。本发明的方法允许在低丰度蛋白质的复杂样品中以精确定量进行灵敏检测,所述蛋白质可以是调节网络的重要部分并且有助于限定由单核细胞增生李斯特氏菌调节蛋白控制的调节子。

[0214] 本发明可帮助测定艰难梭菌的毒力。艰难梭菌是形成孢子的革兰氏阳性厌氧菌,其导致在人和多种动物物种中的范围广泛的疾病。临床表现从无症状定殖和轻度腹泻到危及生命的假膜性结肠炎及死亡。虽然艰难梭菌毒素已得到充分研究,但对艰难梭菌毒力的理解还不透彻。本发明可用于确定在模拟体内感染的体外条件下艰难梭菌的时间依赖性蛋白质表达变化,目的是为了鉴别用于诊断或治疗措施的潜在候选者。

[0215] 克罗恩氏病(CD)是慢性、病理性、T细胞介导的自身免疫炎性病状,其涉及IL-12和IL-23,它们是响应于细胞内微生物侵入由抗原呈递细胞产生的细胞因子。然而,鉴于其在免疫中的作用,在疗法中直接靶向IL-12/IL-23可具有不希望有的后果,所以其它靶标可更有效。本发明的方法和装置可用于帮助鉴别来自CD患者血浆的细胞外可溶性因子,所述因子刺激IL-12/IL-23产生。

[0216] 本发明的另一方面涉及制造如本文所述的生物分子处理器的方法。制造过程一般由以下步骤组成:(1)制造在基底中包括微流体网络和纳米通道的微流体结构,(2)在基底上定位纳米线,(3)在纳米线中的纳米通道/纳米线接合处形成纳米间隙,(4)形成与纳米线传感器的微接触,以及(5)与盖板粘合以完成纳米流体传感器装置。步骤(1)-(3)的制造工艺顺序可任意地选择和/或其中两个可任意地组合。每个步骤如下更详细地描述。

[0217] 一个示例性制造工艺顺序展示于图12中。在此实施方案中,步骤1涉及使用单步骤纳米压印光刻(NIL)工艺,用预结构化的印模在基底(例如聚合物基底)中构建微米和纳米通道结构。在步骤2中,纳米线或纳米颗粒垂直于基底上所形成的通道定位,这些通道变成进料器和飞行时间通道。在步骤3中,纳米间隙是在定位的纳米线中形成,所述纳米线定位在通道上以形成与相应的飞行时间通道相交的一对感测电极。在步骤4中,感测电极连接至微接触上以便于外部引线连接。在步骤5中,将制造好的结构与盖板粘合以完成封闭式流体装置。

[0218] 替代的制造工艺顺序展示于图13中。在此实施方案中,步骤1涉及在纳米通道形成之前将纳米线或纳米颗粒定位在传感器位置处。在步骤2中,使用NIL形成微米和纳米流体结构、通道。纳米间隙可在NIL工艺期间或之后形成,或部分在NIL工艺期间形成且在后续处理过程中完成。形成纳米间隙之后,传感器电极连接至微接触垫用于外部引线连接(步骤3),且制造好的结构与盖板粘合以封闭流体装置(步骤4)。

[0219] 另一替代制造工艺顺序描绘于图14中。此工艺涉及使用横向液体离子芯电极构建

纳米流体装置,所述电极可随后用导电金属纳米颗粒填充以改善其导电率。NIL被用于制造微/纳米流体网络、结构和通道以及液芯纳米电极结构(步骤1)。制造好的结构随后与盖板粘合以封闭流体装置(步骤2),且将金属纳米颗粒添加到液芯电极中(步骤3)。

[0220] 不管采用的制造工艺顺序如何,尺寸缩小工艺步骤可被纳入在通过NIL制造纳米流体结构之后以缩小流体网络、特别是飞行时间通道的尺寸。例如,此工艺步骤可包括在低于基底聚合物的玻璃化转变温度(T_g)的高温下用盲压板加压,这将缩小纳米流体结构的尺寸并且降低纳米流体结构的侧壁粗糙度。

[0221] 对所有制造工艺通用的是将“底版印模”用于基于纳米压印光刻法形成微/纳米流体网络和结构。

[0222] 图15描绘在硅基底中底版印模结构的示例性制造工艺。在步骤1中,进行光刻法和铬蚀刻与Si蚀刻步骤组合(使用活性离子蚀刻或湿化学蚀刻)以界定和形成流体网络在硅基底中的微观结构。如果需要具有不同深度的微观结构,那么重复光刻法和Si蚀刻步骤。在步骤2中,聚焦离子束铣削或电子束光刻法可用于形成传感器装置的纳米结构,包括纳米通道(例如飞行时间通道)、纳米通道进口结构以及生物反应器室的支撑结构。

[0223] 将制造好的底版印模直接用作NIL印模以在聚合物基底中产生流体结构,或其被用作底版模具以产生具有负性结构(关于底版印模中的结构)的树脂NIL印模。关于后者,制造好的底版印模可经由热或UV NIL工艺被复制到各种材料中,这将在随后用作NIL印模以在聚合物基底中产生流体结构,如Wu等人,“Complete Plastic Nanofluidic Devices for DNA Analysis via Direct Imprinting with Polymer Stamps,”Lab Chip 11(17):2984-9(2011)所述,其在此以引用的方式整体并入。用于NIL树脂印模的材料包括例如热塑性聚合物片、涂覆在基底上的热塑性聚合物层、涂覆在基底上的UV可固化树脂以及涂覆在基底上的溶胶-凝胶材料。可使用的基底包括聚合物片、金属、硅、玻璃、石英以及复合材料。涂覆在聚合物基底上的UV可固化树脂的使用是一种优选材料,因为其通过在模制期间降低接触中的材料的杨氏模量而造成与模制聚合物基底的低粘附。热应力还由于UV树脂/聚合物印模和聚合物基底的类似的热膨胀系数而降低。

[0224] 图16A显示经由UV纳米压印光刻法从硅底版印模生产树脂印模的方法。在此方法中,首先将适量的低粘度可固化UV树脂分配或旋涂于玻璃基底上(步骤1)。随后将UV可固化树脂/基底压制于硅底版上(步骤2)。在树脂完全填充于硅底版的结构中之后,将UV光用于固化或聚合UV树脂。将固化的UV树脂/基底从硅底版上脱模以从硅底版印模完全制造UV树脂印模(步骤3)。

[0225] 图16B显示使用制造的树脂印模经由热NIL在热塑性聚合物基底中产生生物分子处理器的微/纳米流体网络结构的工艺。将印模/基底组件加热到基底聚合物的玻璃化转变温度(T_g)以上,通常在 T_g 之上10-100°C(步骤1)。接着将压力施加于印模表面之上(步骤2)。压力可通过使用两个平行压板的施力或通过压缩空气来施加。需要压缩空气以实现均匀施压以及NIL印模与聚合物基底之间的保形接触。在基底聚合物被完全地填充到印模结构中之后,将印模/基底组件冷却至 T_g 以下且将印模从模制基底中脱模(步骤3)。用于热NIL的适合的聚合物基底包括不限于PMMA、PC及COC。

[0226] 所有生物分子处理器制造工艺所共有的另一特征是在基底内制造和定位纳米线或纳米颗粒以用作传感器。适合的纳米线可使用本领域技术人员已知的工艺形成。一般制

造工艺的示意图显示于图17A中。简言之,使用电子束光刻法由塑料或氧化铝基底形成纳米多孔模板或膜,且导电基底例如银、金或氧化铟锡通过溅射被沉积在膜的底部上(步骤1)。纳米线是通过将金属或合金电沉积到具有导电基底的纳米多孔膜中而形成(步骤2)。溶解掉纳米多孔膜(步骤3)且除去导电基底(步骤4)以释放纳米线。图17B是经由此工艺形成的具有10nm直径的NiW纳米线的SEM。

[0227] 一种或多种不同种类的金属或合金向纳米多孔模板中电沉积以形成分节段纳米线通常使用两种不同的技术进行,例如双槽法,其中使用不同电解溶液和不同电镀条件;或单槽法,使用一种溶液和电流或电压的脉冲以形成不同的层(参见,例如Liu等人,“Perpendicular Giant Magnetoresistance of Multilayered Co/Cu Nanowires,”*Physical Review B* 51:7381-7384(1995);Blondel等人,“Wire-Shaped Magnetic Multilayers for ‘Current Perpendicular to Plane’ Magnetoresistance Measurements,”*Journal of Magnetism and Magnetic Materials* 148:317-318(1995);Blondel等人,“Comparative Study of the Magnetoresistance of Electrodeposited Co/Cu Multilayered Nanowires Made by Single and Dual Bath Techniques,”*Journal of Magnetism and Magnetic Materials* 165:34(1997);Piroux等人,“Arrays of Nanowires of Magnetic Metals and Multilayers:Perpendicular GMR and Magnetic Properties,”*Journal of Magnetism and Magnetic Materials* 175:127-136(1997);Evans等人,“Current Perpendicular to Plane Giant Magnetoresistance of Multilayered Nanowires Electrodeposited in Anodic GMR in Electrodeposited CoNiFe/Cu Multilayered Nanowires and Nanotube Aluminum Oxide Membranes,”*Applied Physics Letters* 76:481(2000);Davis等人,“GMR in Electrodeposited CoNiFe/Cu Multilayered Nanowires and Nanotube,”*ECS Transactions* 1:71-77(2005);以及Pratt等人,“Perpendicular Giant Magnetoresistances of Ag/Co Multilayers,”*Physical Review Letters* 66:3060-3063(1991),其在此以引用的方式整体并入)。这些技术的共同性在于界面区域经受断裂,这可用于形成生物分子处理器所需的必要的纳米间隙。

[0228] 退火步骤可有助于形成大粒且因此减小纳米线电阻,从而增强电极测量单个单元的能力。然而,挑战是匹配邻近区段的热膨胀系数。因此,申请者已开发使用具有受控组成的脉冲电沉积将FeNiCo电沉积为纳米线的条件。通过改变所施加的电流密度或添加添加剂如2-丁炔-1,4-二醇(BD),可改变沉积组成,由此调整热膨胀系数。金(或银)节段是由独立的电解质以双槽配置来电镀,且纳入硫酸处理以在Au与FeNiCo步骤之间形成稳固的分节段纳米线。

[0229] 图18是用于生成适用于本发明的生物分子处理器中的稳固的分节段纳米线的上述工艺和条件的示意图。在步骤1中,导电基底(例如银)通过溅射被沉积在纳米多孔膜的底部上。表面在FeNiCo的电化学脉冲沉积(步骤3)之前暴露于空气(步骤2)。可以50%工作循环(2秒开/2秒关)和等于或小于25mA/cm²的所施加的电流密度进行脉冲沉积。更高的电流密度可用于减少Fe的量,但脉冲沉积停止时间应增加。0.1g/L 2-丁炔-1,4-二醇(BD)的添加将在这些条件下产生具有更多Ni的合金。BD还用于调整热膨胀系数,并且防止在脉冲停止时间期间侵蚀。在步骤4中,Au是使用DC电镀由独立的电解质以电化学方式沉积。在后续

的FeNiCo电化学沉积(步骤6)之前,将表面用硫酸溶液(5体积%)处理(步骤5)。重复步骤4-6以形成完全分节段的线(步骤7)。溶解膜并释放纳米线(步骤8)。

[0230] 生物分子处理器的纳米线优选地具有在5-20nm范围内的直径。虽然涉及在纳米多孔膜中进行电沉积以便产生20-200nm线的方法在本领域中是已知的,但制造更薄的线需要纳米多孔膜的精化或替代技术。

[0231] 关于后者,图19是显示使用牺牲纳米管(即,改性的纳米多孔膜)制造超薄纳米线的工艺的示意图。将由不同的金属或聚合物如铜或聚吡咯组成的纳米管通过控制电镀条件在质量传递控制下沉积到纳米多孔模板中(步骤1)。线及其节段随后用电化学方法沉积到纳米管中(步骤2)。为了避免在孔口处的沉积,使用超填充添加剂来促进底部向上的填充。管-线复合物是通过溶解膜来释放(步骤3)且线是通过优先蚀刻管来释放(步骤4)。

[0232] 在替代实施方案中,使用图20中所描绘的电化学蚀刻方法形成超薄电极。将如上所述制造的分节段纳米线定位在装置上的传感器位置处(步骤1)。牺牲间隙区首先以化学方法在通道内蚀刻(步骤2)。剩余电极上的锐缘提供电化学蚀刻的优先区域,如果电流分布在很大程度上是主要的(例如,低瓦格纳数量)和/或蚀刻是由扩散驱动的话。使用构建在装置中的电接触蚀刻各侧面(步骤3)以形成适用于生物分子处理器中的超薄电极。

[0233] 制造工艺的另一个整体部分涉及将纳米线或纳米颗粒定位在生物分子处理器的传感器位置处的方法。当在低于20nm纳米线时,这尤其是一项挑战。在一个实施方案中,使用表面可编程组件。在表面可编程组件中,纳米线将特别“粘附”于的化学贴片是如图21所示在基底的表面上形成。将含有化学贴片的基底暴露于溶液分散的纳米线或纳米颗粒,所述溶液分散的纳米线或纳米颗粒本身定位在基底表面上的化学贴片位置处。

[0234] 化学贴片可使用剥离法分子组装图案化(MAPL)技术在基底表面上形成(参见,例如Falconnet等人,“A Novel Approach to Produce Protein Nanopatterns by Combining Nanoimprint Lithography and Molecular Self-Assembly,”Nano Lett.4(10):1909-1914(2004);Park等人,“Nanostructuring of Anti-Adhesive Layers by Hot Embossing Lithography,”Microelectron Eng.67-68:252-258(2003);Park等人,“Chemical Patterning of Sub-50-nm Half Pitches Via Nanoimprint Lithography,”Microelectron.Eng.78-79:682-688(2005);及Park等人,“Chemical Nanopatterns Via Nanoimprint Lithography for Simultaneous Control Over Azimuthal and Polar Alignment of Liquid Crystals,”Adv.Mater 17(11):1398-1402(2005),其在此以引用的方式整体并入)。概述此工艺的一般示意图显示于图22中,该图展示由以硫醇基封端的烷烃组成的化学贴片的生成作为一个实施例,所述化学贴片对基于金的材料具有高亲和力。对于图22中所描绘的此工艺,将具有与基底(高Tg材料)相比降低的玻璃化转变温度(Tg)的抗蚀剂旋涂于此基底上(步骤1)。在步骤2中,利用热NIL来将图案转移到此低Tg的抗蚀剂中。接着是击穿剩余抗蚀剂的反应性离子蚀刻步骤,其还在聚合物基底上形成官能团(步骤3)。侧挂的-COOH表面官能团与烷烃硫醇的化学反应(步骤4)之后,将剩余抗蚀剂从表面上剥离(步骤5)。这些硫醇用作金或银纳米线经由自组装的锚定点。

[0235] 在一个替代实施方案中,化学贴片如图23所示经由模版掩模在基底表面上形成。热塑性塑料如PMMA和PC是示例性基底。一般说来,此工艺涉及将纳米模版掩模施加于基底表面以掩蔽其中不需要化学贴片的表面区域(图23,步骤1)。暴露的表面经受反应性离子蚀

刻(RIE)和/或其它化学修饰以便添加化学贴片(步骤2和3)。一旦除去掩模(步骤4),便将化学贴片定位于所需位置处。替代的实施方案描绘于图24中。在此实施方案中,在施加纳米模版掩模(步骤1)和经由RIE和/或其它化学修饰添加化学贴片(步骤2)之后,沉积纳米线(步骤3),而掩模仍是完整的。然后在将纳米线定位在基底表面上之后除去掩模(步骤4)。

[0236] 纳米线还可在施加电场的不存在或存在下磁性针针对基底上的纳米感测位置,如分别在图25和26中所描绘。在图25中所示的实施方案中,纳米压印光刻法被用于在基底内形成微/纳米流体通道和结构。随后,将金接触(电极)沉积到基底表面上的纳米通道任一侧的传感器位置处(步骤1)。将铁沉积在金接触上以形成磁场(步骤2),并且将纳米线与通道垂直地对齐(步骤3)。在图26的实施方案中,将金接触(电极)沉积到基底表面上的纳米通道任一侧的传感器位置处(步骤1)。纳米线与在金接触沉积之后施加的外部磁场平行对齐(步骤2)。

[0237] 纳米间隙在定位线中的形成可如图27和28中所描绘以多种方式实现。图27显示使用纳米压印光刻法同时形成微/纳米流体网络(即通道和结构)和纳米间隙。在此实施方案中,使用上述任何适合的方法将纳米线定位于基底上(步骤1)。在步骤2中,将适用于形成纳米间隙和纳米通道的薄树脂印模用压力(例如压缩空气)施加于含有定位线的基底上。依赖于应力集中,实现完整或部分断裂。将微接触沉积在基底表面上且将盖板粘合以封闭纳米流体装置(步骤3)。只要实现部分断裂,便通过后续的电击穿(即,经由纳米线施加高电场)、化学蚀刻或其它合适的手段形成完全断裂(步骤4)。

[0238] 图28描绘当微/纳米流体网络(即通道和结构)在基底中预先形成时通过纳米压印光刻法形成纳米间隙的工艺。在此实施方案中,使用上述方法将纳米线定位于通道上(步骤1)。将定位的纳米线夹在薄聚合物与基底之间并且施加压力以实现纳米线跨通道的完全或部分断裂(步骤2)。将微接触沉积在基底表面上且将盖板粘合以封闭纳米流体装置(步骤3)。只要形成部分纳米间隙,便通过后续的电击穿、化学蚀刻或其它合适的手段形成完全断裂(步骤4)。

[0239] 纳米间隙还可通过本领域中已知的方法来制造,包括例如,扫描探针、线路交叉、模板合成材料、荫罩掩模蒸发、机械断裂接合技术、电镀及碳纳米管的局部氧化切割(Chen等人,“On-Wire Lithography-Generated Molecule-Based Transport Junctions:A New Testbed for Molecular Electronics,”J.Am.Chem.Soc.130(26):8166-8168(2008);Hu等人,“A Self-Assembled Nano Optical Switch and Transistor Based on a Rigid Conjugated Polymer,Thioacetyl-End-Functionalized Poly(para-phenylene ethynylene),”J.Am.Chem.Soc.127:2804-2805(2005);Kushmerick等人,“Effect of Bond-Length Alternation in Molecular Wires,”J.Am.Chem.Soc.124:10654-10655(2002);Mbindyo等人,“Synthesis and Assembly of Nanowires for Molecular Electronics,”Abstracts of Papers of the Am.Chem.Soc.223:155-PHYS(2002);Mbindyo等人,“Template Synthesis of Metal Nanowires Containing Monolayer Molecular Junctions,”J.Am.Chem.Soc.124:4020-4026(2002);Reed等人,“Conductance of a Molecular Junction,”Science 278:252-254(1997);Tang等人,“Encoding Molecular-Wire Formation Within Nanoscale Sockets,”Angew.Chem.Int.Ed.46:3892-3895(2007);Xu&Tao,“Measurement of Single-Molecule Resistance by Repeated

Formation of Molecular Junctions,” Science 301:1221-1223 (2003), 其在此以引用的方式整体并入)。

[0240] 微接触垫在基底表面上形成作为用于将纳米线或感测电极与将转换由电极检测的信号的外部电子部件联接的手段。可使用无电沉积工艺形成微接触垫,其中贵金属纳米颗粒被图案化到基底上以用作“种子”层,以便允许无电沉积Au(金)薄膜,举例来说,仅当纳米颗粒已经定位时。

[0241] 使用两个不同的可能工艺策略的微接触垫形成过程描绘于图29A和29B中。对于图29A中描绘的策略,通过常规的光掩模UV暴露(254nm)NIL聚合物基底(步骤1),之后NIL以形成流体网络并且定位纳米线和间隙形成。这仅在聚合物用UV光照射的位点处产生羧酸。在步骤2中,整个晶片在EDC/NHS存在下与乙二胺起反应-这通过在表面羧酸与乙二胺的一个胺末端之间生成酰胺键来形成胺封端的基团。胺封端的基团仅在暴露于UV辐射的位点处形成。然后,将晶片浸泡于含有Ag纳米颗粒的溶液中(步骤3),由此与胺基形成稳固的复合物,如文献中所详细记载。Ag纳米颗粒不会粘附到未暴露于UV辐射的聚合物基底,这是由于这些表面的疏水性和锚定点的缺乏,在此情况下是胺基。最终步骤涉及将晶片置于Au化学镀槽中并且形成所需的基于Au的微接触垫。

[0242] 如图29B中所描绘,在将光掩模置于聚合物基底上之后,各种直接沉积技术可用于形成本领域的技术人员熟知的金属微接触(步骤1)。这些技术包括但不限于电子束沉积、金属靶喷镀或磁控管沉积。沉积之后,除去模版掩模(步骤2)。

[0243] 在本发明的生物分子处理器的制造过程中的最终步骤涉及盖板的粘合以封闭微/纳米流体装置。盖板粘合工序描绘于图30A中。所述工艺涉及将盖板与模制基底装配在一起并且将已装配的产品置于真空密封袋中。真空密封通过整个热过程使结构变形减至最少。将脱气的真空密封袋置于压力加热器(图30B)中,加热至溶剂沸点,所述沸点接近基底和盖板的玻璃化转变温度,并且在粘合之后,冷却组件。沸腾的溶剂提供了均匀的温度和蒸气压,这确保了压力在整个装置上的均匀分布。使用此工艺的结果显示对于盖板与低纵横比的微观结构的粘合没有结构变形,甚至在纳米规模下也是如此(图30C)。

[0244] 可使用侧面、顶部或底部电子连接将在装配好的装置中的微接触与显示于图32中的适当的信号处理电子器件互连(图31)。侧面连接是通过将电线封接于直接在纳米感测装置上的微接触来实现。顶部或底部连接具有金属接触,所述金属接触是垂直于微接触做出并且包含在图32中所描绘的另外的电子芯片上。使用本领域技术人员常见的技术将信号处理所必需的电子部件保持在印刷电路板(PCB)上。

[0245] 集成双电极纳米间隙装置的等效电路显示于图32中。电读出的主要贡献因素是纳米间隙电阻(R_N)、与缓冲剂接触的纳米电极表面的电容(C_{ne})以及纳米电极的电阻(R_{ne})。

[0246] 因为纳米间隙是由 R_2 连接,所以结果是在入口纳米间隙处的干扰将同时在穿过 R_2 的出口纳米间隙处观察到。而且,跨每个纳米间隙产生的电流变得在各自之间以不确定和在不同时期变化的比率分配。因此,必须设计一个电子电路,其通过将跨纳米间隙的常见物(common)与另一者分离(浮动)而防止在穿过 R_2 的两个纳米间隙之间的串话(参见图32)。为了实现此目的,使用具有两个依赖性浮动($\pm 15V$ DC)电源的两个电流电压(I-E)转换器。此外,为了确保系统的放大端与数字化端的有效分离,使用在A/D和D/A转换器之前和之后的集成线性光隔离器(光耦合器)。光隔离器是电子装置,其被设计来通过利用光将电信号从

输入端转移到输出端。其提供与电绝缘的联接并且防止在电路一端处的变化影响其它。每个光隔离器电路包括光芯片和四个运算放大器。塑料屏蔽的BNC电缆和连接器被用于数字化仪与I-E转换器之间的所有连接中。

[0247] I-E转换器的核心部件是前置放大器。虽然任何流速仪都应具有低的输入阻抗,但I-E转换器中的全电子部件必须具有高输入阻抗,因为有待测量的电流极小(通常是皮安;pA)。二级放大器可用于实现1,000的增益并维持高带宽($\sim 80\text{kHz}$)。数位可选的反馈电容器 C_F (1pF)还可与 R_F 平行连接以防止任何振荡。放大器的不倒相输入经由屏蔽的三线电缆连接至纳米间隙电极(这提供了指令电压)、接收输出电流并且屏蔽来自任何电容耦合的信号线使其不受外部噪声的影响。使用连接到常见的一个放大器并与其它分离的双重法拉第笼实现外部噪声的屏蔽。

[0248] 上述电路设计可使用本领域技术人员的惯用做法在单一芯片中多路复用(参见图32)。为了使数据存储空间减到最小,峰值检测电路被用于单个芯片中,其中还记录了峰值振幅。多路复用芯片可与现场可编程门阵列(FPGA)相接以处理来自此多路复用芯片的输出。

实施例

[0249] 提供以下实施例用于说明本发明的实施方案,但并不意图限制本发明的范围。

[0250] 实施例1-5的材料和方法

[0251] 以下实施例和模拟结果呈现了使用电色谱法鉴别单个单核苷酸的飞行时间的过程。对于dNMP,使用CHARMM27力场(Foloppe&MacKerell, "All-Atom Empirical Force Field for Nucleic Acids:I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data," J. Comput. Chem. 21:86-104 (2000), 其在此以引用的方式整体并入), 并且对于水, 使用与刚性CHARMM TIP3P模型的离子相互作用。静电研究的短程截止点是1.0nm。在0.8nm与1.0nm之间的间隔下, 兰纳-琼斯相互作用(Lennard-Jones interaction)转变成零。对于长程静电相互作用, 使用了针对平板几何形状校正的三维粒子-粒子-网格法(Yeh&Berkowitz, "Ewald Summation for Systems with Slab Geometry," J. Chem. Phys. 111:3155-3162 (1999), 其在此以引用的方式整体并入), 其中在非周期性方向上的长度是在该方向上模拟箱长度的三倍。

[0252] 通过使用例如针对双链DNA的 λ -核酸外切酶切割完整双链或单链DNA, 可以得到在5'端上具有磷酸酯基的脱氧核苷酸单磷酸酯或dNMP(dAMP、dCMP、dGMP、dTMP)。基于酶活性的最佳pH值是7.5, 模拟的dNMP是非质子化的, 因为其pKa值是约6.8(Nucleic Acids in Chemistry and Biology; Blackburn编辑, IRL Press: Oxford, (1990), 其在此以引用的方式整体并入), 得到其净电荷是 $-2e$, 其中 e 是电子电荷。CHARMM27拓扑结构文件不包含DNA的在5'端上具有非质子化磷酸酯基的末端片段。使用补丁(patch), 通过增加或减少原子并调节局部电荷来调节现有的拓扑结构。拓扑结构文件包含DNA的在5'端上具有质子化磷酸酯基、在3'端上具有质子化磷酸酯基并且在3'端上具有非质子化磷酸酯基的末端补丁。因此, 在非质子化3'型式与质子化3'型式之间的受影响原子上局部电荷的差异增加了非质子化5'型式的局部电荷。受影响原子仅包括磷酸酯基及其直接连接的碳原子。

[0253] 为了能使用在5'端上具有非质子化磷酸酯基的脱氧核苷酸单磷酸酯进行分子动

力学 (molecular dynamic, MD) 模拟, 将以下“补丁”添加到现有的CHARMM力场中:

[0254] PRES 5P03-2.00!5'terminal PHOSPHATE patch!same shifts in charge as between 3PHO and 3P03

[0255] GROUP

[0256] ATOM C5'CN8B-0.18

[0257] ATOM 5'HN80.09

[0258] ATOM H5''HN80.09

[0259] ATOM P P 1.10

[0260] ATOM O1P ON3-0.90

[0261] ATOM O2P ON3-0.90

[0262] ATOM O5'ON2-0.40

[0263] ATOM O3P ON3-0.90

[0264] BOND O3P P

[0265] ACCE O3P P

[0266] !Built in B-DNA-like conformation (NF)

[0267] BILD C4'C5'O5'P 0.0000 000.00-146.00 000.00 0.0000

[0268] BILD C5'O5'P O3P 0.0000 000.00-46.90 000.00 0.0000

[0269] BILD O3P O5'*P O1P 0.0000 000.00-115.82 000.00 0.0000

[0270] BILD O3P O5'*P O2P 0.0000 000.00 115.90 000.00 0.0000

[0271] 图33显示了dNMP, 其中全部四个末端和受影响原子被标记。

[0272] 狭缝壁是由平行于xy平面并且位于 $z = \pm 1.5\text{nm}$ 处的两个平面界定; 壁原子的中心位于 $z \leq -1.5\text{nm}$ (底部) 和 $z \geq 1.5\text{nm}$ (顶部) 处。狭缝壁是通过模拟由原子构成的体相流体来构建, 其中对于碳原子, 利用兰纳-琼斯12-6参数 ($\epsilon_{\text{碳-碳}} = 0.11\text{kcal/mol}$, $\sigma_{\text{碳-碳}} = 0.4/21/6\text{nm}$)。原子质量从 12.011amu 增加到 14.30226amu , 使其等于PMMA的联合原子 (不包含氢原子, 而是连接氢原子的原子的质量增加) 呈现中原子的平均质量。这一模拟是在 5000K 下执行, 以确保其为流体, 并且密度接近非晶形PMMA的密度。所选模拟箱在x和y方向上的大小是 5.0nm , 以便匹配期望的狭缝壁大小。狭缝壁的最初配置是取自这一流体的 $5.0 \times 5.0 \times 1.2\text{nm}^3$ 矩形区域。如果原子中心在该矩形区域内, 那么其被包括在壁中。由此产生原子级光滑 (粗糙度小于原子大小) 但不均质的表面。在流体模拟期间, 从间隔 10ps 的不同时间取得每一组狭缝壁。所选狭缝宽度是 3nm , 以避免浪费模拟时间, 从而使dNMP在较长时间段内在狭缝中心周围扩散, 不过已经制造出尺寸小于 5nm 的实际通道 (Menard等人, Nano Lett. 11: 512-517 (2011), 其在此以引用的方式整体并入)。还应注意到, 由于水很难进入其中, 故使用小的、完全呈疏水性的纳米级几何结构是不切实际的。所选狭缝壁在物理上是不现实的, 而是在此初始研究中作为简易系统。

[0273] 一旦构建出壁, 就在狭缝壁之间移动dNMP, 并且使用VMD添加水和离子。对于所用几何形状, 38NPT模拟不能在LAMMPS中进行, 因此水量是由试验和误差决定。使用VMD中的溶剂化功能添加水, 其中设置如下: 边界=2.4, x和y范围= $\pm 25.65 \text{ \AA}$, z范围= $\pm 15 \text{ \AA}$ 。这些设置是由试验和误差决定, 由此在仅含水和狭缝壁的平衡模拟中得到水 (狭缝中心) 的合理容积密度。平衡时水的容积密度是约 1.015g/cm^3 。在dNMP溶剂化之后, 使用VMD中的自电离

功能去除四个水分子并用三个钠离子和一个氯离子代替。用于平衡和不平衡模拟的模拟系统以及DNA的化学结构示于图34A和34B中。

[0274] 以下描述了用于进行模拟的方法和常用参数。壁原子通过弹簧常数是83860kcal/mol·nm²的弹簧附接到其初始位置。对于流量模拟,使用贝伦德森恒温器(Berendsen thermostat)以0.1ps的时间常数,在300K下仅对壁原子进行温度调节。在稳态下流动期间流体的温度比由粘性流对流体加热而进行温度调节的壁的温度高约3K。对于平衡模拟,在相同时间常数下,对流体使用另外的恒温器。对于每一dNMP,用不同壁配置执行三次或四次模拟以减少由特定壁配置引起的任何偏差。通过在x方向上向每个原子施加恒定体积力 f_i 来诱导流动(参见图34A)。所选 f_i 的幅值应使得 $f_i = m_i a$,其中 m_i 是原子 i 的质量并且 a 是加速度,对于所有原子选择相同值。这一方法是压力驱动的流动(Zhu等人,“Pressure-induced Water Transport in Membrane Channels Studied by Molecular Dynamics,” Biophysical Journal 83:154-160 (2002), Carr等人, IEEE Trans. Nanotechnol. 10:75-82 (2011), 其在此以引用的方式整体并入)或由毛细管力产生的流动(Han等人, J. Colloid Interface Sci. 293:151-157 (2006), 其在此以引用的方式整体并入)的一种近似方法。由

对流体所有 n 个原子施加力 f_i 引起的沿流动方向的压力梯度是由下式表示: $\frac{dp}{dx} = \frac{\sum_{i=1}^n f_i}{A_{\text{流体}} L_{\text{箱}}}$,

其中 $L_{\text{箱}}$ 是模拟箱沿外加力的方向(x)的长度,以及 $A_{\text{流体}}$ 是在垂直于外加力方向的平面(yz 平面)上流体的面积。通过选择 $a = 263.592 \text{ nm/ns}^2$,并且使用描述模拟系统的参数(即,流动区域中原子的数量及其质量、垂直于流动的面积,以及模拟箱的长度)的特征值,在纳米狭缝中驱动流动的压力梯度是约 $dp/dx = 2.836 \text{ bar/nm}$ 。施加的力在 z 方向上产生近似抛物线的速度曲线,以约 1.5 m/s 的最大流动速度为特征。在MD模拟中,由于可达到相对较短的时间,通常仅数百纳秒,故相对于纳米级流动典型的速度,在此类较高速度下进行模拟是必要的。如稍后将论述的,尽管速度极高,但描述dNMP与壁的吸附和解吸附的特征参数被认为基本上不受流动的影响。分析时舍弃每次流动模拟的前15ns以便能达到稳态条件。在平衡模拟时,舍弃1ns的平衡期。

[0275] 用于确定核苷酸吸附和解吸附事件的算法。对于与下部壁的单次吸附和解吸附事件,用于测定吸附和解吸附时间的算法示于图35中,并且如下进行。首先,使用0.5ns的平均运行长度,使 d_w 变光滑(标为 $d_{\text{光滑}}$)。确定点,标为 d_1 ,其中 $d_{\text{光滑}}$ 横过 $\pm 0.8 \text{ nm}$ 。这近似于PMF曲线开始变平滑的情形(图36)。取决于这些点横过 $\pm 0.8 \text{ nm}$ 边界的方向,假定其进行吸附事件或遵循解吸附事件。其余步骤使用了初始未光滑化的 d_w 。在与 d_1 相同的时间,确定一组点(标为 d_2),但对应于 d_w 。从 d_2 开始,按时间正向(解吸附)或逆向(吸附)追踪数据,直到 d_w 值 $\geq 1.0 \text{ nm}$,即,远离PMF曲线的最小值(图36)。接着,倒转追踪方向并进行追踪,直到 $d_w \leq d_{\text{上部}}$ (此时接近上部壁), $d_w \leq d_{\text{下部}}$ (此时接近下部壁),或直到达到数据的开始或结束。在大多数情况下, $d_{\text{上部}}$ 和 $d_{\text{下部}}$ 被设置为 0.42 nm ,比PMF曲线的总体最小值略接近壁(图36)。在一些情况下,其中之一须略作调节,以免遗失明显的吸附和解吸附事件。这是必要的,因为每个壁中原子的构型都是不同的,因此每个壁上的吸附作用略有不同。当然,有可能设计出一种算法来测定初始 $d_{\text{上部}}$ 和 $d_{\text{下部}}$,由此无需对其进行手动调整以适当地捕获事件。然而,确切的初始值不影响最终的吸附和解吸附时间值。由此追踪得到的点表示为 d_3 。如果 d_3 存在重复,或时间等于数

据集中的第一时间或最后一个时间,那么去除这些点,因为这些点对应于不接近吸附或解吸附事件的 d_2 点。在去除这些点之后,在dNMP被吸附到上部壁和下部壁上时,使用 d_3 作为吸附和解吸附点来计算 d_w 的平均值。接着重复追踪步骤,不过用刚刚计算的平均值代替 $d_{上部}$ 和 $d_{下部}$,更新 d_3 ,并在吸附时再次计算 d_w 的平均值。在一系列迭代之后, d_3 收敛至一个稳态值。光滑化并选择 $d_w=0.8\text{nm}$ 来确定 d_1 将较短时间或与表面的距离偏离从解吸附事件分类中去除,并且初始选择 $d_{上部}$ 和 $d_{下部}$ 来确定初始 d_3 防止dNMP从表面弹回的事件被归类为吸附事件。

[0276] PMF的不确定性估计。图36中所示PMF的不确定性是通过将每个面元(bin)的制造时间分成三份来计算。另外,对于每个单核苷酸,接着由这三份时间产生三个PMF曲线。这允许评估所得平均化PMF曲线上每个点的平均值及与平均值的标准偏差。为了估计四个PMF中最大值位置以及平均吸附自由能的不确定性,使用以下数值误差传递法:i)假定曲线上每个点的平均值及与平均值的标准偏差确定每个面元的正态分布,ii)从每个点的分布取得100,000个随机样品以生成100,000个PMF曲线。从100,000个自由能曲线以100,000个最大值以及由这些曲线计算的平均吸附自由能获得PMF中最大值位置及平均吸附自由能的不确定性。所取得的不确定性是标准偏差的两倍。

[0277] 对于仅在dNMP被吸附到或未吸附到壁时计算的在流动方向和角度下的速度,如下估计不确定性。使用分块平均(block averaging),但不是按照每个分块具有相等长度的常规方式。而是针对每个dNMP的所有模拟的每个吸附或解吸附期都视作独立分块。使用每个分块中数据的平均值得到分块平均值 x_i 。通过对n个分块的平均值取加权平均值,获得总体平均值,其中权重 w_i 是每个分块中数据点的数量。 x_i 的加权平均值和方差是使用下式计算

$$[0278] \quad w'_i = w_i / \min(w_i) \quad (S1)$$

$$[0279] \quad \bar{x}_w = \sum_{i=1}^n w'_i x_i / \sum_{i=1}^n w'_i \quad (S2)$$

$$[0280] \quad s_w^2(x_i) = \frac{\sum_{i=1}^n w'_i (x_i - \bar{x}_w)^2}{\left(\sum_{i=1}^n w'_i\right) - 1} \quad (S3)$$

[0281] 加权平均值的方差是通过除以有效基点(b)(在权重相等的情况下缩减至n)来计算,并且所得不确定性(u)是加权平均值的标准偏差的两倍。

$$[0282] \quad s^2(\bar{x}_w) = s_w^2(x_i) / b \quad (S4)$$

$$[0283] \quad b = \left(\sum_{i=1}^n w_i\right)^2 / \sum_{i=1}^n w_i^2 \quad (S5)$$

$$[0284] \quad u(\bar{x}_w) = 2\sqrt{s^2(\bar{x}_w)} \quad (S6)$$

[0285] 使用下式计算在流动方向上的总体平均速度及其不确定性:

$$[0286] \quad \bar{v} = f_{ads} \bar{v}_{ads} + 0 - f_{des} \bar{v}_{des} \quad (S7)$$

$$[0287] \quad u(\bar{v}) = 2\sqrt{f_{ads}^2 s^2(\bar{v}_{ads}) + 0 - f_{des}^2 s^2(\bar{v}_{des})} \quad (S8)$$

[0288] 其中 f_{ads} 是所有模拟的吸附时间除以针对每个dNMP的所有模拟的总时间。

[0289] 平均吸附时间分数、吸附事件的平均频率及其不确定性也是用等式S1-S6计算的。唯一差别在于,“分块”是对每个dNMP的整体模拟,吸附时间分数的权重(w_i)是总模拟时间,并且吸附频率的权重是在流动方向上行进的总距离。对于吸附时的dNMP-壁能量,使用具有相等分块长度的分块平均化(在等式S1-S6中所有 w_i 都相等)。分块长度是1300个点,对应于0.52ns。

[0290] 轴2与壁表面的角度。如图41(针对吸附)和图42(针对解吸附)中所示,在平衡和不平衡情况下,图35中所确定的轴2与最近壁表面平面的平均角度随 d_w 变化。表7(见下文)显示了在dNMP被吸附时计算的平均轴2角度,不管这是在吸附期还是在解吸附期间。图41和图42中关于轴2所示的指定核苷酸的吸附和解吸附曲线彼此无法容易地区分,因为角度的变化与噪声相当。当核苷酸在吸附期间接近壁时,dNMP的疏水性最强的一侧一般更接近壁;dTMP例外。非平面的甲基使核苷酸带有甲基的一侧远离壁,即使甲基具有疏水性。壁附近轴2的定向在解吸附期间不够强。已知噪声水平,dAMP、dCMP或dGMP的轴2角度无显著最小值或最大值。对于dTMP,疏水性甲基-表面相互作用使得轴2角度在吸附和解吸附期间具有显著最大值。在平衡情况下,解吸附时dTMP的最大角度小于吸附时的最大角度。这一最大值在平衡情况下也小于不平衡的情况。对于dTMP的解吸附,在最大值之后也存在轴2角度的最小值。

[0291] 测定分离飞行时间分布所需的通道长度。有关测定分离单核苷酸飞行时间(TOF)分布所需的最小通道长度 L 的详情如下。假定飞行时间分布没有预先形成。而是通过使用梯形法则,对沿轨线区段经历0.5nm的飞行时间分布(d^{TOF})的概率柱状图进行数字积分来估计连续分布函数(CDF)。随后,通过线性插值法估计每个CDF达到值0.00135($\mu-\delta^-$)和0.99865($\mu+\delta^+$)的点,其中 μ 是平均TOF。这些点界定样品分布的宽度。在以下等式中,这些点被认为是TOF分布的边缘。数字0.00135和0.99865确定允许多少重叠并且可以改变以在分离分布时得到更高或更低准确性。为简单起见,在这一分析中,在归一化形式(即,全部除以 N 的随机变量TOF的值)下使用在主文档中确定的 $DTOF_{0.5 \times N}$ 。结果是,在归一化形式下,对于任何核苷酸 α , $DTOF_{0.5 \times N}$ 都以平均TOF值 μ_α 表征,并且其宽度随 N 的平方根而减小。使每对核苷酸的分布(α, β)不重叠所需的 $N_{\alpha, \beta}$ 最小值是由等式(S11)得出,该等式是等式(S9)和(S10)的组合。所需的最小通道长度(L)是由等式(S12)得出。

$$[0292] \quad \mu_\alpha + \frac{H[\mu_\beta - \mu_\alpha] \sigma_\beta^2 - H[\mu_\alpha - \mu_\beta] \sigma_\alpha^2}{\sqrt{N_\alpha}} = \mu_\beta + \frac{-H[\mu_\beta - \mu_\alpha] \sigma_\beta^2 + H[\mu_\alpha - \mu_\beta] \sigma_\alpha^2}{\sqrt{N_\beta}} \quad (S9)$$

$$[0293] \quad N_\beta = \frac{\sigma_\alpha^2}{\sigma_\beta^2} N_\alpha = N_\alpha = N_{\alpha, \beta} \quad (S10)$$

$$[0294] \quad N_{\alpha, \beta} = \frac{H[\mu_\beta - \mu_\alpha] \sigma_\beta^2 + H[\mu_\alpha - \mu_\beta] \sigma_\alpha^2}{(\mu_\alpha - \mu_\beta)^2} \quad (S11)$$

$$[0295] \quad L = \max [N_{\alpha, \delta^{TOF}}] \quad (S12)$$

[0296] 在这些等式中, μ 是平均值, δ^- 和 δ^+ 是平均值到CDF等于0.00135和0.99865的点的距离。 $H[x]$ 是赫维赛德阶跃函数(Heaviside step function)(对于 $x \leq 0$, $H[x] = 0$ 并且对于 $x > 0$, $H[x] = 1$),该函数用于区分两种可能的分布布置;类型 α 的分布平均值小于类型 β ($\mu_\alpha <$

μ_β), 或类型 α 的分布平均值大于类型 β ($\mu_\alpha > \mu_\beta$)。核苷酸类型 α (N_α) 和 β (N_β) 所需的样品数量具有满足等式(S1)的许多可能组合。然而, 这些组合受以下事实限制: 两个核苷酸必须行进相同的总体距离 ($d_\beta^{\text{TOF}} N_\beta = d_\alpha^{\text{TOF}} N_\alpha$)。由于用于计算类型 α (d_α^{TOF}) 和 β (d_β^{TOF}) 的飞行时间的距离都是 0.5nm (d^{TOF}), 故 $N_\alpha = N_\beta = N_{\alpha,\beta}$ 。现等式(S9)中仅 $N_{\alpha,\beta}$ 是未知的, 并且解出其值, 得到等式(S11)。 $N_{\alpha,\beta}$ 的最大值乘以 d^{TOF} 就是所需的通道长度。

[0297] 实施例1—平衡模拟与吸附自由能

[0298] 平衡模拟的目标是研究在不存在流动的情况下个别dNMP吸附和解吸附的机制和能量。对于每个dNMP, 在不同壁配置下执行数次模拟(参见表8)。每次模拟持续约65ns, 同时所有四种dNMP和所有不同壁配置的总模拟时间是约966ns。个别dNMP与狭缝壁的相互作用是通过整个纳米狭缝的自由能曲线完整描述。随反应坐标(平均力势或PMF)变化的相对于参考状态的自由能差异 ΔF 与dNMP位于指定反应坐标值处的概率 Pr 相关。由于dNMP被吸附到光滑表面或从光滑表面解吸附的概率是值得关注的, 故将反应坐标视为dNMP质量中心与位于 $z = \pm 1.5\text{nm}$ 处的最近狭缝壁平面的距离 d_w 。使用 $\text{Pr}(d_w)$ 值, 通过下式得出自由能曲线

$$[0299] \quad \Delta F(d_w) = -kT \ln \left[\frac{\text{Pr}(d_w)}{\text{Pr}(d_{w,\text{ref}})} \right] \quad (1)$$

[0300] 其中 k 是玻尔兹曼常数(Boltzmann's constant), T 是温度, 并且 $d_{w,\text{ref}}$ 是选择的在纳米狭缝中远离壁的参考状态($z=0$)。

[0301] 在平衡模拟期间, 全部四种dNMP都吸附到壁表面并从壁表面解吸附多次, 因此, 可以直接应用等式(1)。在dNMP吸附到狭缝表面上, 保持与总模拟时间相当的时间段(例如超过数纳秒)的情况下, 计算自由能可能需要偏差平衡模拟(Biased equilibrium simulation), 如伞形取样(Torrie&Valleau, Chem. Phys. Lett. 28:578-581 (1974); Novak等人, "Umbrella Sampling Simulations of Biotin Carboxylase: Is a Structure With an Open ATP Grasp Domain Stable in Solution?" J. Phys. Chem. B 113:10097-10103 (2009); 及Novak等人, "Behavior of the ATP Grasp Domain of Biotin Carboxylase Monomers and Dimers Studied Using Molecular Dynamics Simulations," Proteins: Struct. Funct. Bioinform. 79:622-632 (2011), 其在此以引用的方式整体并入)。多次吸附和解吸附事件还允许在待取样的不均质表面上存在不同吸附位点, 因此等式1中的 ΔF 是这些不同位点的平均值。事实上, dNMP在与壁平面相切的方向上移动, 即使是在吸附到壁上时, 这也使得其能从较多壁表面取样。由此指示, dNMP在壁表面上不存在任何强吸附位点并且在整个表面上吸附强度存在极少变化。这些是使基于飞行时间的测序可行所希望的壁表面特征。然而, 壁表面不是由实际材料制成并且比实际表面光滑。一经发现, 在非晶形二氧化硅表面上的不同位点处的吸附强度在0到10kT间变化(Carr等人, J. Phys. Chem. Lett. 2:1804-1807 (2011), 其在此以引用的方式整体并入)。PMMA表面上dNMP吸附强度的变化是未来研究的主题。

[0302] 图36显示了全部四种dNMP的自由能随 d_w 变化的曲线。这些曲线是类似的并且以充分确定的最小值的存在为特征。深度至多是热能的数倍 kT 的势阱的存在反映了以下事实: dNMP吸附到疏水性壁上, 但不够强。四条曲线的最小值近似地位于相同距离处, 距壁约 0.475nm 。如下一节中所论述, 当dNMP被吸附到壁上并处于其最小自由能状态时, 其主要分

子单元中的两个,即含氮碱基和糖基与壁表面具有类似的几何布置,并因此,其质量中心都位于距壁大致相同的位置处。全部四条曲线都显示在吸附区与邻近狭缝中心的主体溶剂之间存在具有多达 kT 的小能垒。此处,吸附区定义为邻近壁的区域,在其中dNMP经历朝向其势阱最小值的拉力。如图36中所见,针对四种核苷酸的吸附区的向外程度可以通过小能垒的相应峰值的位置 d_w^M 来定界。尽管自由能和相应整体最小值 $\Delta F_{\min}$ 的特征(表6中所示)提供了对dNMP与壁相互作用的强度的直接度量,标量参数 ΔF_{ads} ,即所谓的净吸附自由能更适于与实验比较(Raut等人,“Molecular Dynamics Simulations of Peptide-Surface Interactions,”*Langmuir*21:1629-1639 (2005),其在此以引用的方式整体并入)。根据自由能特征,通过对表征处于吸附状态的dNMP的所有状态的自由能的加权和进行评价,来获得 ΔF_{ads} ,如下所示:

$$[0303] \quad \Delta F_{\text{ads}} = -\ln\left(\frac{C_{\text{ads}}}{C_{\text{容}}}\right) = \int_0^{\infty} PD(d_w) \Delta F(d_w) d(d_w) \approx \sum PD_i \Delta F_i(\Delta d_w)_i = \sum Pr_i \Delta F_i \quad (2)$$

[0304] 在这一等式中, ΔF_{ads} 和 ΔF 都是以 kT 为单位表示。 C_{ads} 是吸附区中的浓度,并且 $C_{\text{容}}$ 是远离壁的本体溶液的浓度。下标 i 表示面元数。 $(d_w)_i$ 是面元 i 的位置,其宽度是 $(\Delta d_w)_i$ 。 Pr_i 表示dNMP处于指定面元中的概率。 $PD_i = Pr_i / (\Delta d_w)_i$ 表示指定面元中归一化的概率密度。表6提供了表征dNMP吸附到狭缝疏水性壁上的吸附区的能量和范围的相关量的概述。

[0305] 表6:四种dNMP与疏水性狭缝壁相互作用的自由能特征的最小值($\Delta F_{\min}$)及其平均吸附自由能 ΔF_{ads} ,以及用于对吸附区定界的自由能最大值的位置(d_w^M)

核苷酸	$\Delta F_{\min}/kT$	$\Delta F_{\text{ads}}/kT$	d_w^M (nm)
dAMP	-1.6 ± 0.4	-0.69 ± 0.07	0.71 ± 0.09
dCMP	-1.0 ± 0.7	-0.21 ± 0.05	0.71 ± 0.05
dGMP	-2.2 ± 0.7	-1.19 ± 0.03	0.70 ± 0.04
dTMP	-3.4 ± 0.4	-2.62 ± 0.04	0.88 ± 0.03

[0307] 实施例2—吸附和解吸附动力学

[0308] 将吸附和解吸附统计学特性和动力学特性与如飞行时间等其它特性相关联是值得特别关注的。关于模拟结果的分析指示,当dNMP被吸附时,相对疏水性的核碱基的环往往是在表面上大致呈平面,亲水性磷酸酯基从表面指向外,并且糖也与表面具有某种接触(参见图34A)。先前已经在通过氮化硅纳米孔模拟输送DNA链时观察到核碱基与表面的这种疏水性粘附作用(Aksimentiev等人,*Biophys. J.* 87:2086-2097 (2004),其在此以引用的方式整体并入),并且在通过石墨烯纳米孔模拟输送DNA链时观察到与石墨烯板的疏水性粘附作用(Wells等人,*Nano Lett.* 12:4117-4123 (2012),其在此以引用的方式整体并入)。图37显示在流动模拟期间dNMP质量中心轨线的 z 分量的50ns区段。平衡模拟的轨线是类似的。如通过吸附事件的频率和长度的差异所指示,四种dNMP的轨线之间存在明显差异。为了定量吸附和解吸附统计,引入了一种用于测定吸附和解吸附时间的方法并且已于上文详细解释。

[0309] PMF曲线的整体最小值和小能垒与吸附和解吸附机制相关。由于dNMP的平面环往

往吸附到壁上,故在核碱基平面和壁表面平面中所界定的两个轴之间的角度可用于观察吸附/解吸附机制。图38(针对吸附)和图39(对于解吸附)中显示了在平衡和不平衡情况下,从核碱基中将糖附接到核碱基另一端的原子或其附近指出的轴(轴1)与最近壁表面平面的平均角度随 d_w 的变化。图41(针对吸附)和图42(对于解吸附)中显示了在平衡和不平衡情况下,垂直于轴1并且在核碱基平面中的轴(轴2)与最近壁表面平面的平均角度随 d_w 的变化。这些轴示于图33(轴1)和图40(轴2)中以及插图上。为了计算这些平均角度,吸附期是从前一次的解吸附时间与指定吸附时间之间的一半时间长度到吸附时间与下一次的解吸附时间之间的一半时间长度。使用了一种类似方法来测定解吸附期。表7显示了在dNMP被吸附时计算的平均轴1和轴2角度,不管这是在吸附期还是在解吸附期间。

[0310] 表7:在平衡和不平衡情况下,当dNMP被吸附时轴1和轴2与壁表面的平均角度。不确定性是平均值的标准偏差的两倍并且如以上所描述计算。

	核苷酸	平衡		不平衡	
		轴1	轴2	轴1	轴2
[0311]	dAMP	-8.60 ± 1.15	-0.63 ± 1.66	-9.00 ± 1.10	-0.67 ± 0.83
	dCMP	-7.86 ± 1.33	-0.81 ± 1.29	-6.16 ± 3.25	-2.72 ± 3.04
	dGMP	-4.63 ± 0.58	0.82 ± 0.82	-4.37 ± 2.09	0.72 ± 0.70
	dTMP	-8.84 ± 0.93	-3.18 ± 0.89	-8.91 ± 0.66	-2.99 ± 0.66

[0312] 图38和图39中的曲线显示,核碱基附接到糖的末端最后吸附并且最先解吸附。这是因为糖和磷酸酯基的亲水性比核碱基强,而这更有利于与疏水性狭缝壁相互作用。轴1角度的整体最小量值遵循与核碱基的疏水性相同的趋势($G < C < A < T$) (Shih等人, “Hydrophobicities of the Nucleic Acid Bases: Distribution Coefficients from Water to Cyclohexane,” J.Mol.Biol.280:421-430 (1998); 及Munoz-Muriedas等人, “A Hydrophobic Similarity Analysis of Solvation Effects on Nucleic Acid Bases,” J.Mol.Model.13:357-365 (2007), 其在此以引用的方式整体并入)。对于疏水性更接近于糖的核碱基,轴1角度的最小值不太严格。当核碱基的糖端与表面的距离略远于另一端,但仍被吸附时,出现PMF曲线的最小值,约 $d_w = 0.475\text{nm}$ 。糖端从表面分离使自由能增加。PMF曲线在 $d_w = d_w^M$ 处的能垒(参见表6)与核碱基上与糖端相对的一端(dNMP离开表面的最后一部分)的吸附和解吸附关联。

[0313] 图38和图39中关于轴1所示的指定dNMP的吸附和解吸附曲线不相同。对于dAMP、dCMP和dGMP,解吸附时轴1角度的整体最小量值大于吸附时的整体最小量值。对于dTMP,解吸附时轴1角度的整体最小量值小于吸附时的整体最小量值,并且在解吸附曲线中存在显著最大值。dTMP的最小量值较大是因疏水性甲基粘于表面所致。

[0314] 图41和图42中关于轴2所示的指定核苷酸的吸附和解吸附曲线彼此无法容易地区分,因为角度的变化与噪声相当。当核苷酸在吸附期间接近壁时,dNMP的疏水性最强的一侧一般更接近壁;dTMP例外。非平面的甲基使核苷酸带有甲基的一侧远离壁,即使甲基具有疏水性。壁附近轴2的定向在解吸附期间不够强。已知噪声水平,dAMP、dCMP或dGMP的轴2角度

无显著最小值或最大值。对于dTMP,疏水性甲基-表面相互作用使得轴2角度在吸附和解吸附期间具有显著最大值。在平衡情况下,解吸附时dTMP的最大角度小于吸附时的最大角度。对于dTMP的解吸附,在最大值之后也存在轴2角度的最小值。

[0315] 轴1与壁表面的角度通常具有相同性质,并且不平衡情况与平衡情况通常在数量上具有类似表现。性质表现的例外是,在嘧啶碱基(dAMP和dGMP)解吸附期间,在不平衡情况下存在最大值,但在平衡情况下则不存在。数量上的例外是在平衡情况下,在dTMP解吸附期间轴2角度的最大值较小。

[0316] 实施例3—dNMP与通道壁相互作用对各种平衡和流动特征的影响

[0317] 表8显示了通过模拟计算的各种量。这些量特别显示:i)对于四种dNMP中的每一种,在不平衡情况下的吸附时间分数与在平衡情况下大致相同。由此指示,通过这些模拟中使用相对较高的流动速度不会显著改变dNMP的吸附/解吸附行为。相较于每秒约数百米的分子最大瞬时热速度,所使用的流动速度仍较小。ii) dCMP和dTMP的飞行时间即使仅经历50nm的行进距离仍充分分离,这意味着不带电的疏水性表面(如这些模拟中所使用的表面)足以区分这两种dNMP。这可以归因于核碱基疏水性的差异:G<C<A<T(Shih等人,“Hydrophobicities of the Nucleic Acid Bases:Distribution Coefficients from Water to Cyclohexane,”J.Mol.Biol.280:421-430(1998);及Munoz-Muriedas等人,“A Hydrophobic Similarity Analysis of Solvation Effects on Nucleic Acid Bases,”J.Mol.Model.13:357-365(2007),其在此以引用的方式整体并入)。胸腺嘧啶具有疏水性甲基,而胞嘧啶具有亲水性胺基。iii)疏水性递增的趋势与飞行时间的递增趋势相配,不过dGMP和dCMP的次序颠倒。这是由dNMP-壁相互作用所致;吸附时dGMP-壁能量量值($V_{\text{nuc-壁}}$)大于吸附时dCMP-壁能量量值。由于核碱基上的不同基团与壁表面之间不存在较强的特异性相互作用,并且由于核碱基在平面上大致呈平面,故 $V_{\text{nuc-壁}}$ 遵循与核碱基的溶剂可及表面积($A_{\text{碱基}}$)相同的趋势。由于鸟嘌呤的面积大于胞嘧啶,故dGMP与壁的接触面积大于dCMP并且具有较大的 $V_{\text{nuc-壁}}$ 。iv) dAMP的吸附事件数量是dGMP的两倍,但它仅吸附略微较短的时间分数,并且具有比dGMP长的飞行时间。这是因为dGMP在再次解吸附或吸附之前在吸附或离开壁上花费了较长的时间量,并且dGMP具有略大的壁滑行速度。将表8中所示的数据与表6中所示的数据相比较显示,通过平衡模拟计算的自由能显示出预期的趋势,较低的最小或平均吸附自由能的负值得到较长的飞行时间。dGMP和dAMP的自由能是个例外。

[0318] 表8:通过流动模拟计算的量,最后三行除外。碱基的面积是由单一构型得到,并且最后两行是由平衡模拟得到。括号中的不确定性是平均值的标准偏差的两倍并且如以上所描述计算。

	dNMP	dCMP	dGMP	dAMP	dTMP
[0319]	N_{sim}^1	3	4	4	4
	$T_{\text{sim}}(\text{ns})^2$	155.9	252.9	213.1	305.2
	$D(\text{nm})^3$	50.0	48.0	32.0	32.0
	n_{ads}^4	0.21(0.06)	0.08(0.03)	0.18(0.03)	0.09(0.06)
	$f_{\text{ads,neq}}^5$	0.38(0.15)	0.68(0.23)	0.63(0.10)	0.93(0.03)
	$t_{\text{飞行}}(\text{ns})^6$	52(7)	66(22)	77(19)	121(30)
	$v_{\text{all}}(\text{m/s})^7$	0.99(0.23)	0.80(0.17)	0.63(0.18)	0.41(0.13)
	$v_{\text{滑动}}(\text{m/s})^8$	0.68(0.61)	0.42(0.24)	0.33(0.28)	0.38(0.14)
	$V_{\text{nuc-壁}}/(\text{kT})^9$	-10.9(0.4)	-14.5(0.3)	-13.2(0.3)	-12.4(0.2)
	$A_{\text{碱基}}(\text{nm}^2)^{10}$	2.4270	2.7963	2.6446	2.6197
	$f_{\text{ads,eq}}^{11}$	0.38(0.09)	0.68(0.14)	0.60(0.05)	0.90(0.09)

[0320] ¹壁配置(模拟)的数量

[0321] ²所有配置的总模拟时间

[0322] ³对于每一配置,在流动方向上行进的距离

[0323] ⁴每行进一纳米,吸附事件的数量

[0324] ⁵在不平衡情况下的吸附时间分数

[0325] ⁶经历50nm距离的飞行时间

[0326] ⁷在流动方向上的平均速度

[0327] ⁸吸附时在流动方向上的平均速度

[0328] ⁹吸附时dNMP-壁势能

[0329] ¹⁰核碱基的溶剂可及表面积

[0330] ¹¹平衡情况下的吸附时间分数

[0331] 实施例4—基于不平衡MD模拟分离dNMP所需的通道长度

[0332] 实现dNMP飞行时间的可靠分离所需的通道长度是由dNMP经历从相应MD轨线的总长度取样的0.5nm区段的飞行时间分布(DTOF_{0.5})计算。这些轨线的总长度对于dTMP和dAMP是128nm,对于dCMP是150nm并且对于dGMP是192nm(参见表8,第1行和第3行)。0.5nm是任意选择的,并且使用使用介于0.5nm与50.0nm之间的任何其它长度得到大致相同的结果。

[0333] 使用DTOF_{0.5}并采用经典的中央极限定理(Wasserman L.W., ALL OF STATISTICS:A CONCISE COURSE IN STATISTICAL INFERENCE (2004), 其在此以引用的方式整体并入), 可以计算出每种dNMP经历更长距离 $d=0.5N$ nm的飞行时间分布(DTOF_{0.5N}), 其中N是正整数。对于每种dNMP, 使用DTOF_{0.5}可以构建出相应DTOF_{0.5N}随所描述的随机变量分布 $T_N=t_1+t_2+\dots+t_N$ 的变化, 其中 $\{t_1, t_2, \dots, t_N\}$ 是从DTOF_{0.5}随机选取的一组独立的飞行时间值 t_i ($i=1, N$)。每一新的随机变量 T_N 表示dNMP经历 $d=0.5N$ nm的飞行时间。根据中央极限定理, 如果 μ 和 σ^2 分别是平均值和方差, 用于表征DTOF_{0.5} (不必呈正态或甚至对称分布), 那么在较大N值的极限情况下, DTOF_{0.5N}将接近正态分布并将通过平均值 $N\mu$ 和方差 $N\sigma^2$ (或等于 $(N)^{1/2}\sigma$ 的标准偏差) 表征。相较于随N线性增加的平均值, 当区段数量N增加时, DTOF_{0.5N}的宽度仅随N的平方根而增加。因此, N增加(在此情况下转换成飞行距离的增加) 引起DTOF_{0.5N}的更好分离。很明显, 当估计分离所需的最小通道长度 $d_{\min}=0.5N$ nm时, 需要界定可接受的分布重叠水

平。在所述方法中,当N使得任一平均值分布中小于0.27%与其它平均值分布之一重叠时,估计 $d_{\min}=0.5N$ nm是dNMP行进的距离。对于正态分布,这一百分比对应于3倍的平均值标准偏差。对于C和T的误判率是0.27%,并且对于A和G的误判率是0.46%。有关N的计算和所需长度的详情描述于上。

[0334] 表9显示了分别分离六个不同dNMP对的飞行时间所需的距离值。完全区分所有dNMP所需的长度是分离dAMP和dGMP的飞行时间所需的距离,即5.9 μm 。尽管是近似值,并且对于在真实系统中难以复制的模拟条件,这一数值可以用作比较不同表面的基础。

[0335] 表9:分离dNMP对的飞行时间所需的距离(μm)

	dCMP	dGMP	dAMP	dTMP
dCMP	∞	1.31	0.43	0.06
dGMP	1.31	∞	5.94	0.24
dAMP	0.43	5.94	∞	0.56
dTMP	0.06	0.24	0.56	∞

[0337] 图43显示了经历5.9 μm 的飞行时间分布,其中每一飞行时间是通过来自0.5nm的飞行时间分布的 $N=N_{\text{dAMP,dGMP}}$ 样品进行加和来产生。在这一距离下,dAMP和dGMP的分布刚好分离。全组分布的宽度是每一dNMP的最短分析时间($\Delta t_{\text{分析}}$),因为进入通道的dNMP之间的时间必须至少是此长度,以防止混乱。

[0338] 实施例5—实际流动速率的估计。

[0339] 尽管在MD模拟中,可以在每一dNMP约10 μs 的最短分析时间内经历约数微米的较短距离对dNMP进行分离,但实际流动速率将会小2-3个数量级。在这些条件下,由于扩散,飞行时间分布将会明显变宽。在这一小节中,将估计扩散对所需通道长度和每一dNMP的最短分析时间的影响。设想如果不在较低流动速率下进行极其昂贵的MD模拟,那么估计实际流动速率的影响的最佳方式是使用Carr等人所介绍的策略,该策略涉及在全部三个笛卡尔维度(Cartesian dimension)中计算dNMP与壁之间的平均力势(PMF),然后在布朗动力学(BD)模拟中使用所得到的力与位置的函数(Carr等人,Lab Chip 11:3766-3773 (2011),其在此以引用的方式整体并入)。尽管通过平衡PMF计算得到的力在不平衡条件下不是绝对准确,但Carr等人显示,在压力驱动流下,BD模拟与MD模拟之间存在良好的一致性(Carr等人,Lab Chip 11:3766-3773 (2011),其在此以引用的方式整体并入)。这种3D PMF+BD策略的优势在于,尽管PMF计算很昂贵,但BD模拟仅涉及允许比MD模拟的时间范围和通道宽度大得多的单一dNMP。

[0340] 使用有关dNMP的扩散能力和速度曲线以及观察dNMP随距壁的距离而变化的概率(通过等式1与PMF关联)的一些假定,可以获得通道长度和每一dNMP的分析时间的数量级,而无需计算3D PMF。使用以下假定:

[0341] (1) 所有dNMP在距壁任何距离处的扩散能力与环状AMP的扩散能力($4.44 \times 10^{-6} \text{cm}^2/\text{s}$)相同并且相等(Dworkin等人,J.Biol.Chem.252:864-865 (1977),其在此以引用的方

式整体并入)。

[0342] (2) 在最大速度 ($v_{\max}$) 0.048cm/s (该速度是dNMP在1000V/cm电场下通过电泳作用在主体水溶液中进行近似的速度) 和速度零 (针对小于或等于吸附时的平均dNMP位置的 d_w 值) 下, dNMP的速度曲线呈抛物线形。假定零滑行速度可能不够准确, 而在较低速度下并且利用较粗糙的壁, 滑行速度与平均速度的比率将比在MD模拟中低得多。假定不同形状的速度曲线将略微改变平均速度, 而不改变估计值的数量级。

[0343] (3) 在流动方向 (x) 上的平均速度是

$$\begin{aligned} \bar{v}_x &= \int_0^{d_w^{\text{cl}}} PD(d_w) v_x(d_w) dd_w \\ &\approx \sum_i PD_{i, v_x}(\Delta d_w)_i = \sum_i Pr_i v_{x,i} \end{aligned} \quad (3)$$

[0345] 该等式类似于等式2并且除 $v_{x,i}$ 和 d_w^{cl} 外, 使用了相同的符号, $v_{x,i}$ 是在面元i中在流动方向上的速度, 并且 d_w^{cl} 是在狭缝中心线处距壁的距离。等式3的使用假定, 在流动模拟中的概率与在平衡模拟中的概率相同。由于即使在不平衡MD模拟中, 在极高速度下, 吸附时间分数也是类似的, 故这是一个有效的假定。

[0346] (4) 对于指定通道长度L、平均速度 $\bar{v}_x$ 及扩散率D, 飞行时间概率密度分布是

$$p(t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{(L - \bar{v}_x t)^2}{4Dt}\right] \quad (4)$$

[0348] 刚好是在时间依赖性平均位置 $\bar{v}_x t$ 下一维扩散的分布。

[0349] (5) 容许的分布重叠与MD模拟分析中所使用的分布重叠相同。

[0350] 如果这些假定是适当的, 那么目标是发现不会违反所希望的分布重叠的最小L, 以及相应的 $\Delta t_{\text{分析}}$ 。估计在 $v_{\max}=0.048\text{cm/s}$ 下测定的L和 $\Delta t_{\text{分析}}$ 值分别是约2.5mm和39s。图44显示了经历2.5mm距离的飞行时间分布。尽管已经制造出1cm长并且尺寸在20nm以下的通道 (Liang等人, Nano Lett. 7:3774-3780 (2007), 其在此以引用的方式整体并入), 但考虑到可能需要对数百万或数十亿dNMP (人类基因组具有约30亿个碱基对) 进行测序, 分析每一dNMP的时间则太长。在一天内以所述速率对人类基因组进行测序将同时需要约140万个通道, 还不包括任何冗余。幸运的是, 当L仅与速度成反比时, $\Delta t_{\text{分析}}$ 与速度的平方成反比。如果速度增加10倍, 那么所需长度将是250 μm 并且 $\Delta t_{\text{分析}}$ 将是390ms。

[0351] 对于指定 $v_{\max}$, 可以估计出最小可能 $\Delta t_{\text{分析}}$ 。除上述假定1、4及5以外, 还假定一类dNMP具有 $\bar{v}_{x,1}=v_{\max}$; 假定它不受壁影响。假定第二类dNMP具有 $\bar{v}_{x,2}=\bar{v}_{x,1}-\Delta v_x$, 并且第三类和第四类具有甚至更小的速度 ($\bar{v}_{x,4}<\bar{v}_{x,3}<\bar{v}_{x,2}$), 因此, 根据以上假定5, 所有分布彼此尽可能地聚集在一起。通过改变 Δv_x , 可以发现 $\Delta t_{\text{分析}}$ 的最小值。对于 $v_{\max}=0.048\text{cm/s}$, $\Delta t_{\text{分析}}$ 的最小值是约5.0s并且L是约1.29mm。对于 $v_{\max}=0.48\text{cm/s}$, $\Delta t_{\text{分析}}$ 的最小值是约50ms并且L是约129 μm 。

[0352] 关于实施例1-5的讨论

[0353] 在水溶液中, 在由无序的兰纳-琼斯碳原子构成的3nm宽的纳米狭缝中模拟DNA中常见的四种dNMP 5'-单磷酸酯, 以便比较其飞行时间和动力学。通过体积力驱动所述溶液,

已知这些体积力可以诱导与压力差或毛细管力所诱导类似的流动。为进行比较,并且为了计算将dNMP吸附到狭缝壁上的自由能,同时考虑无外加力的情形。dNMP在数纳秒内实现吸附和解吸附,即使在无流动情况下也是如此。该流动对dNMP的吸附和解吸附行为不具有较大影响,并且dNMP吸附时的时间分数在平衡模拟中与在不平衡模拟中大致相同。使用疏水性碳狭缝壁,容易分离疏水性最强的dNMP (dTMP) 及一种亲水性最强的dNMP (dCMP) 的飞行时间,由此指示,调节壁材料的润湿性将是基于通过纳米通道的飞行时间高度可靠地区别各dNMP的一种颇具前景的方式,不过dAMP和dGMP中特定化学基团与壁表面上的基团的相互作用对于其分离也很重要,因为其飞行时间大致相同。关于dAMP和dGMP的重要问题是,相较于平衡条件,在流动条件下在解吸附期间其性质上不同的表现是否在分离其飞行时间的能力方面起到作用,并且这一表现在实际装置中所使用的低得多的速度下是否能保持。有关在实际流动速率下扩散的影响的分析指示,dNMP速度应相对较高,以便得到每一dNMP的合理的最短分析时间。

[0354] 实施例6— λ -核酸外切酶(λ -Exo)的固相固定

[0355] 将 λ -核酸外切酶(λ -Exo)固定于热塑性(PMMA)固相表面上。 λ -Exo是在10X反应缓冲液(670mM甘氨酸-KOH,pH 9.4;25mM $MgCl_2$;0.1% (v/v) Triton X-100) (Fermentas Life Sciences,Glen Burnie,MD) 存在下提供。使用前,未进行纯化步骤。

[0356] 在将PMMA盖板热熔粘接到固定 λ -Exo的PMMA基底上(Henry等人,“Surface Modification of Plastics Used in the Fabrication of Microanalytical Devices,” Analytical Chemistry 72:5331-5337 (2000);McCarley等人,“Resist-Free Patterning of Surface Architectures in Polymer-Based Microanalytical Devices,” J.Am.Chem.Soc.127:842-843 (2005),其在此以引用的方式整体并入)之后,形成琥珀酰亚胺酯以便于酶附接。这是通过在室温下,用200mM的3-(3-二甲基氨基丙基)碳化二亚胺(EDC)和50mM的N-羟基琥珀酰亚胺(NHS)于0.1M的2-(4-吗啉基)-乙烷磺酸中的溶液(pH 5.1) (MES,Fisher Biotech,Fair Lawn,NJ) 填充固相反应器,保持15分钟来进行。接着,用由0.03 μ g/ μ L λ -Exo酶组成的溶液在水动力下置换EDC/NHS试剂。使反应在4 $^{\circ}$ C进行整夜。用1X λ -Exo反应缓冲液冲洗酶功能化的装置以从PMMA表面去除未结合的试剂。图45是概述 λ -Exo固定过程的示意图。

[0357] 使用Asylum Research MFP3D原子力显微镜检(AFM),在1.00Hz扫描速率下以AC(轻敲)模式来表征表面固定的 λ -Exo。用于AFM的样品由固定有 λ -Exo酶的PMMA板组成。将PMMA板切割成1.7cm x 1.7cm的正方形并且厚度是3mm。切割后,如先前所描述,对基底进行清洁、干燥及UV活化。如以上所描述,在酶添加完成后,使这些板经历EDC/NHS偶联。在AFM分析之前,用反应缓冲液、ddH₂O冲洗样品,并在压缩空气下小心干燥。

[0358] 经由EDC/NHS介导的外部蛋白质上的胺基与基底表面上产生的羧酸基团的偶联,将 λ -Exo固定于PMMA。通过AFM分析确定 λ -Exo成功附接到PMMA上。为了确定任何酶的非特异性吸附,PMMA基底用UV光进行活化,并且在不存在EDC/NHS偶联剂情况下,与 λ -Exo酶溶液一起在4 $^{\circ}$ C下孵育整夜。图46A是在不存在EDC/NHS偶联剂情况下进行的PMMA/ λ -Exo反应的AFM扫描图。这一扫描未显示出与 λ -Exo酶的大小和形状相符的表面纹理迹象,证明在当前这些条件下未发生聚合物表面上酶的物理吸附。图46B是在EDC/NHS偶联剂存在下进行的PMMA/ λ -Exo反应的AFM图像。与图46A的扫描图相比较,图46B的扫描图显示出与 λ -Exo的大小相符

的表面特征。相较于存在酶并且无EDC/NHS偶联剂情况下 $0.34 \pm 0.01 \text{ nm}$ 的基底平均RMS粗糙度,共价附接蛋白质的基底的平均RMS粗糙度是 $1.58 \pm 0.18 \text{ nm}$ 。因此,仅在EDC/NHS偶联剂存在下观察到酶共价附接到PMMA表面,并且未发生物理吸附。从在基底不同区域($n=6$)上越过 $15 \mu\text{m}$ 区域进行的多次扫描可知,表面特征的AFM高度曲线表明,这些特征的平均高度是 15 nm ,对应于根据x射线结晶数据所报导的 λ -Exo的外径(约 $150 \text{ \AA}$)。图46C是所产生的用于测定这些特征的大小分布的特征高度的柱状图。从这些数据可知,所检测的特征中有71%在 $14.3 \pm 2.3 \text{ nm}$ 的高度范围内。高度变化可能是由于在分析期间为了减少样品破坏或置换所使用的不接触成像模式。这些数据指示,大多数特征与活化表面上 λ -Exo的单点附接相符。因此,大多数固定的酶是与其用于纳入dsDNA的接取孔一起大致垂直于基底表面定向。在柱状图中不存在特征高度的双重分布表明,在用于固定反应的UV剂量和酶浓度下,酶不会发生表面交联,其中酶与其接取孔将与表面平行,从而使其基于表面空间考虑而无法接近dsDNA。由固定得到的AFM数据也证实,实现了完整同三聚体的偶联,并且没有观察到由附接过程引起的酶解离成其单体单元。

[0359] 实施例7—使用 λ -核酸外切酶固相消化 λ -DNA

[0360] 在 λ -Exo改性的反应器中孵育 λ -DNA双链体(48,502bp) (New England Biolabs, Ipswich, MA) 不同反应时间,通过将 λ -DNA溶液以适当速度水动力抽吸穿过固相反应器以提供希望的滞留时间来进行选择。水动力抽吸是使用机械泵 (PHD2000注射泵, Harvard Apparatus, Holliston, MA) 进行控制。经由定制的热电偶加热台将反应器检验的温度控制在 37°C 。在装置的出口处收集流出物用于下游分析。

[0361] 在通过酶反应器传代之后,使用 **PicoGreen[®]** (Life Technologies, Grand Island, NY; 一种双链DNA (dsDNA) 结合染料) 确定残留dsDNA的量。**Picrogreen[®]**对结合于dsDNA具有高度特异性,并且产生荧光增强(当结合于dsDNA时期荧光量子效率有约1,000倍增加)。由于所述染料不结合于单链DNA (ssDNA) 或单核苷酸并伴随荧光增加,故它适合于特异性确定 λ -Exo固相反应中dsDNA的含量,该固相反应应当是由ssDNA、dsDNA及单核苷酸组成的。所述染料是在消化后添加,以用于定量全部dsDNA的量。消化后添加与消化前添加相对,因为先前的报导提出在核染色之后, λ -Exo的酶活性降低 (Subramanian等人, "The enzyme bias of processivity in lambda exonuclease," *Nucleic Acids Research* 31: 1585-96 (2003), 其在此以引用的方式整体并入)。在480nm下激发染料标记的样品,并且使用Fluorolog-3分光荧光计 (Horiba JobinYvon, Edison, NJ) 和DataMax软件2.20收集从490nm到700nm的荧光谱。

[0362] 图47显示了自由溶液中 λ -Exo消化的 λ -DNA和固相反应器中 λ -Exo消化的 λ -DNA (有效反应时间相同) 的荧光谱。为了确定反应器中DNA消化的程度,将**Picrogreen[®]** (一种核染色染料) 添加到反应床流出物或自由酶反应溶液中,并获取荧光谱,从而指示溶液中残留的dsDNA的量。如图47的图中所示,所观察的自由溶液消化的荧光量远高于通过固相反应器消化检测到的荧光量,这表明在自由溶液消化中存在的dsDNA多于固相消化。

[0363] 实施例8—实时监测表面固定的 λ -核酸外切酶的消化作用

[0364] 为了测定固定的 λ -Exo的消化速率,采用实时荧光监测。这些研究中使用的显微镜由配备Zeiss 100x/1.3NA油浸显物镜和Andor iXon3EMCCD相机的Zeiss Aviovert 200M倒

置显微镜(75W的Xe灯;Zeiss,Germany)组成。分析期间,对定制底座进行机械加工以将组装的PMMA装置固定在显微镜台上。收集所有图像并且使用MetaMorph Advanced 7.7.6.0软件(Molecular Devices LLC,Sunnyvale,CA)和ImageJ 1.46(National Institutes of Health,Bethesda,MD)进行分析。在这些实验中,制备不含辅因子的 λ -Exo 10X反应缓冲液(含甘氨酸-KOH的超纯水,pH 9.4;0.1% (v/v) Triton X-100)用于控制对酶消化dsDNA的起始。在 λ -Exo酶附接到PMMA之后,在反应器中,将1:50染料与碱基对比率的YOYO[®]-1标记的 λ -DNA孵育约3小时,以产生酶-DNA复合物。将含有25mM的MgCl₂辅因子的反应缓冲液(1X)引入装置中并加热到37℃,随后获取图像,用于实时监测酶活性(参见图48)。为了能将试剂和dsDNA抽吸通过反应器,使用PHD2000注射泵(Harvard Apparatus,Holliston,MA)。通过用环氧树脂将peak管密封于进口/出口储槽并使用鲁尔锁连接器(luer-lock connector)将进口管连接到注射器,来将泵连接到反应器。

[0365] 基于 λ -DNA的碱基对总数(48,502bp)和总消化时间,测定表面固定的 λ -Exo的平均消化速率是 1135 ± 124 个核苷酸/秒($n=4$)。这一消化速率与先前所报导的那些相符(Matsuura等人,“Real Time Observation of a Single DNA Digestion by Lambda Exonuclease under a Fluorescence Microscope,”Nucleic Acids Research 29:e79 (2001),其在此以引用的方式整体并入),但具有增加的消化速率(Ayub等人,“Nanopore Based Identification of Individual Mononucleotides for Direct RNA Sequencing,”Nano Letters 13:6144-6150 (2013);Matsuura等人,“Real Time Observation of a Single DNA Digestion by Lambda Exonuclease under a Fluorescence Microscope,”Nucleic Acids Research 29:e79 (2001),其在此以引用的方式整体并入)。这一信息表明, λ -Exo的进行性行为消化了>30,000个核苷酸,该值是先前报导值的10倍。为了确定荧光强度降低是由消化作用而非光漂白或光刻作用引起,通过在不存在辅因子情况下,将成串的 λ -DNA分子暴露于激发光,持续比所观察的总消化时间更长的时间,由此进行对照实验。如图48中所描绘,当未用酶辅因子增进反应(即,酶活性等于0.0)时,在典型消化实验的持续时间内,观察到相对恒定的荧光。

[0366] 实施例9—单核苷酸的微芯片(PMMA)毛细管电泳

[0367] 使用Beckman Coulter CE仪器进行毛细管电泳(CE),并在254nm下进行UV检测。对于CE,使用了来自Molex Polymicro Technologies(Phoenix,AZ)的具有50 μ m内径的未涂渍熔融二氧化硅毛细管(总长度=33cm,有效长度是20cm)。CE柱用0.1M NaOH预先调节30分钟,并通过用过滤过的0.5X TBE缓冲液(含45mM Tris、45mM硼酸盐、1mM EDTA的超纯水,pH 8.3;18M Ω)冲洗来进行漂洗。最后,用在0.5X TBE(pH 8.3)中含有2% (w/v) 聚乙烯吡咯烷酮(PVP,Mr=40,000;Sigma Aldrich St.Louis,MO)的动态涂布液处理毛细管表面,以抑制电渗流。CE电泳缓冲液还含有阳离子表面活性剂,在这一情况下是鲸蜡基三甲基溴化铵(CTAB)。通过在10kV下电动进样来引入样品,持续180秒。对1 μ M浓度的天然脱氧核苷酸单磷酸酯(dNMP,4种,包括甲基化的C)进行CE。

[0368] 使用本文所描述的CE条件,实现了天然dNMP的基线分离,如图49中所描绘。不过,这是使用约20cm长的柱实现,并且需要8分钟的显影时间。对于飞行时间分析,分离时间将是约1ms并且柱长度将是约50 μ m。尽管显影时间和柱长都较短,但分离并未降级,因为该分离仅取决于各dNMP的峰宽以及电泳迁移率的差异。电泳迁移率与电场强度、柱长度及CE显

影时间无关。

[0369] 实施例10—表面处理对PMMA纳米通道中热塑性电渗流的影响

[0370] 使用纳米压印光刻法(NIL)制造PMMA纳米流体装置。简单点说,基于纳米通道的流体装置由三个主要部分组成:(a)微流体网络,其包括储槽及75 μm 宽并且20 μm 深的输送通道;(b)在纳米通道的进口处的梯度界面,漏斗状进口由16个梯形棱柱组成,这些棱柱具有从30 μm 递减到75nm的宽度及从10 μm 递减到75nm的深度;及(c)75nm宽并且75nm深的纳米通道阵列。

[0371] 制造步骤可以分成四个关键步骤:(1)形成具有凹进的纳米通道和分级微米级流体网络的硅母模;(2)经由UV-NIL将硅母模图案转印到环烯烃共聚物(COC)基底上涂布的UV可固化树脂层中,以产生具有凸出结构的聚合物印记;(3)将UV-树脂印记纳米压印到PMMA基底中,以产生纳米流体结构;及(4)将PMMA盖片粘接于基底上以形成流体系统。

[0372] 为了构建硅母模,组合硅湿式蚀刻和聚焦离子束(FIB)研磨以构建分级结构。通过常规光刻法界定储槽和输送通道,并且在30%KOH溶液存在下,在硅表面上进行各向异性蚀刻。然后,用FIB(FEI Quanta3D FEG)逐步研磨漏斗状进口和纳米通道阵列。最后,在真空室中,由气相涂布1H,1H,2H,2H-全氟癸基三氯硅烷单层。

[0373] 选择一种含有69wt%三丙二醇二丙烯酸酯(TPGDA)作为基料、29wt%三羟甲基丙烷三丙烯酸酯作为交联剂及2wt%Irgacure 651作为光引发剂的UV可固化聚合物掺混物作为聚合物印记的材料。通过用吸液管分配UV树脂来涂布硅模具。然后,将COC基底(COC-TOPAS 6017, TOPAS Advanced Polymers, Florence KY)放在涂有UV树脂的印模上并轻柔地按压以确保树脂完全填入模具腔中。随后使其暴露于UV光,保持20秒,用于固化。用于固化的UV灯的强度是1.8W/cm²;在遮光角下的宽带波长是400nm并且在约365nm下具有最大波长强度。固化之后,轻柔地将UV可固化树脂从硅模具中取出,由此在UV可固化树脂上得到负拷贝。

[0374] 然后,使用图案化的UV可固化树脂作为印模,以热压印到具有先前研磨用于储槽的孔的3mm厚PMMA板(Lucite CP)中。压印是在130℃和20巴下通过NIL(Obducat纳米压印系统)进行,持续5分钟,该方法是用压缩空气向印模和基底施加压力,确保整个压印区域内的压力均一。压力是在印模和基底在希望的模制温度下预热30秒之后施加,并且在压印过程中保持,直到系统冷却到70℃。在印模和基底冷却到室温之后,可以将PMMA复本容易地从UV-树脂印模取出。

[0375] 使用125mm厚的PMMA板(Goodfellow Corporation)作为盖片。图案化的PMMA板和盖片都通过氧等离子体在15W下处理15秒,以活化表面。在等离子体处理之后立即进行热粘接。在不影响PMMA上的结构的情况下,使用20巴压力、80℃温度和10分钟时间来达到希望的粘接强度。采用NIL机器在粘接过程中产生均匀压力。

[0376] 在用1 μM KCl溶液填充纳米通道之后,使用不连续缓冲系统测量电渗流 EOF,其中阳极储槽填充有2 μM KCl,而阴极储槽填充有1 μM KCl。在施加200V/cm电场进行电泳期间,使用AxoPatch电流放大器连续地监测电流。除非另作陈述,否则归因于通道壁上的负电荷,EOF流向阴极储槽。对于胺改性的PMMA纳米通道壁,EOF是沿阳极方向流动。

[0377] 对于未经过处理的PMMA纳米通道(图50A),EOF在所有pH值下都沿阴极方向移动。可以看出,EOF明显与溶液pH无关,这与先前关于微米通道的结果相符。然而,当用O₂等离子

体处理PMMA纳米通道时,EOF在 $\text{pH}>3$ 下变得显著更高。EOF的这一显著增加可以归因于表面固定的羧酸的去质子化,因为 O_2 等离子体诱导PMMA表面发生氧化反应,产生表面固定的羧酸基团。当这些基团去质子化时,表面电荷密度增加,使得EOF显著改变。因此,聚合物纳米通道上的表面电荷密度可以使用 O_2 等离子体处理进行调节。当PMMA纳米通道经过等离子体氧化,随后在EDC/NHS偶联剂存在下用乙二胺(EDA)处理(图50B)时,EDA的一个胺基共价锚接到在PMMA纳米通道表面上产生的侧接羧酸基团。这一改性引起EOF的逆转(从阴极向阳极行进),表明表面电荷是正电荷,与在从阳极向阴极行进的EOF情况下的负电荷形成对比。在这些数据中观察到的趋势与先前在PMMA微米通道中所观察到的趋势类似。

[0378] 实施例11—使用聚合物纳米柱评价SME:追踪荧光标记的dNMP

[0379] 如以上所描述,制备纳米通道。不过,在这一情形中,这些通道是在环烯烃共聚物(COC)中使用热NIL制造。纳米通道是 $100\text{nm} \times 100\text{nm}$ 并且长度是 $100\mu\text{m}$ 。用 O_2 等离子体处理这些通道和盖板(也由COC制造),并且在 78°C 温度下热组装所述装置,该温度略低于天然热塑性材料的玻璃化转变温度。在这些实验中,使用暗视野显微镜追踪单个分子。这些分子由单个银(Ag)纳米粒子(NP)组成,这些纳米粒子具有 10nm 直径并且涂有柠檬酸根离子,使Ag-NP上带有负电荷。将Ag-NP放在阴极储槽中,并使其在不同电场强度下经电泳进入阳极储槽中。电泳是在TBE缓冲液($1\times$, $\text{pH}=8.0$)中进行,该缓冲液还含有 10mM KCl以减小纳米通道中电双层的厚度。

[0380] 图51B是单个Ag-NP的图像(3D图),显示粒子在放入(稳定的)聚合物纳米通道中时具有高强度。这些单个分子在不同电场强度下经电泳穿过聚合物纳米通道,并且通过光学追踪其运动。绘制易位时间随电场强度变化的柱状图以确定所测量的飞行时间的变化。如图51A中所示,在最高电场强度($4000\text{V}/\text{cm}$)下观察到最小变化。从图51C可以看出,Ag-NP的电泳迁移率(10nm)与所施加的电场无关,而板数与飞行时间分布的变化成反比,在较高电场强度下增加。这符合以下事实:所述变化主要是由纵向扩散而非与易位事件相关的任何其它人为因素引起,因为扩散性散布与 $(2Dt)^{1/2}$ 成比例,其中D是Ag-NP的扩散系数并且t是时间。较高的电场强度引起较短的易位时间,并因此具有较短的扩散性散布时间。

[0381] 实施例12—T4DNA和 λ DNA的纵向电流阻断事件

[0382] 如以上所描述,制造PMMA纳米流体装置。使用倒置显微镜(Zeiss Axiovert 200),通过所述装置对DNA易位事件进行光学测量。荧光显微镜配备有相机(iXon3888Andor)。使用两个法拉第笼(Faraday Cage)屏蔽流体芯片以进行电力测量,一个法拉第笼用于屏蔽样品台,并且另一个覆盖整个显微镜。记录下视频图像并且通过MetaMorph (Molecular Devices)进行处理。在视频记录期间,EMCCD使用重叠模式,并且曝光时间是 5ms 到 20ms 。

[0383] 使用Axopatch 200B放大器和Digidata 1440A (Molecular Devices)以全细胞模式进行电流测量。数据采集和分析是使用pCLAMP 10软件进行的。以 100kHz 速率对信号进行采样,其中内部低通Bessel滤波器设置为 10kHz 。使用Ag/AgCl线作为电极。对于所有测量,在膜片钳施加的 1V 下驱动DNA。

[0384] 在 $1\times$ TBE缓冲液(89mM Tris-硼酸盐和 2mM EDTA, $\text{pH} 8.3$)中制备T4DNA溶液($0.5\mu\text{g}/\text{ml}$)。利用YOYO[®]-1,以 $20:1$ 比率(碱基对/染料)对DNA进行染色。添加 β -巯基乙醇(3%)作为抗光漂白剂,同时向缓冲液中添加由 $0.2\text{mg}/\text{ml}$ 葡萄糖氧化酶、 $0.04\text{mg}/\text{mL}$ 过氧化氢酶和 $4\text{mg}/\text{mL}$ β -D-葡萄糖组成的酶促除氧系统以去除氧。使用真空泵先润湿通道并且将DNA溶液

引入装置中。

[0385] 图52A-H是关于T4DNA在电作用下移动通过PMMA纳米通道的DNA易位事件的光学和电学描绘。这些易位事件产生电流反应,并且如通过光学观察,当DNA易位通过纳米通道时,正好出现尖峰。在易位事件的开始,DNA分子由于低电场强度而在微米通道中缓慢移动。当DNA分子通过位于纳米通道输入端的纳米柱阵列中时,该分子通过与纳米柱碰撞并钩住纳米柱而伸长。一旦到达纳米通道,DNA在强电场作用下迅速地通过纳米通道,所述强电场因纳米通道的高电阻而集中在纳米通道内部。在离开纳米通道之后,DNA分子减速并且快速地重绕。在这些易位过程中,只有DNA通过纳米通道的过程产生离子电流降低。从离子电流在DNA进入纳米通道的那一刻降低可知,其花费1-3ms来使电流从基线值降低到尖峰的底部。离子电流在一定时间内保持约15pA的电流降低,该时间与DNA分子的长度成比例。如荧光图像中所示,将整个T4DNA链转移通过纳米通道花费2~3帧(40~60ms)。所述时间与I-t曲线中的读数相配,其中对于完整T4DNA分子,尖峰宽度是52ms(图52G和52H)。

[0386] 此处 t_T 代表电流降低尖峰的持续时间(DNA易位时间), I_D 代表电流降低的幅度。图52A-52C显示了由短DNA分子易位产生的离子电流尖峰。这些小DNA区段通常是由T4DNA在荧光激发光下通过光刻作用或在样品制备期间的剪切力作用而破坏所产生。通过荧光图像测量的这些DNA区段的长度通常小于5 μm 。较快的电泳迁移率和较小长度使得这些DNA在较短时间内易位通过纳米通道;Error!Reference source not found.=5~10ms并且Error!Reference source not found.在5-10pA间变化。

[0387] 图52D-52F显示了较长DNA易位通过纳米通道的情况。这些DNA在20~60ms内通过并且Error!Reference source not found.是约15pA。假定这些DNA是带切口的DNA分子,而不是对应于全长T4DNA分子。图52G-52I中显示了全长T4DNA分子易位通过纳米通道的情况。此处,电流瞬变的持续时间更符合全长T4DNA分子。

[0388] 这些数据表明,DNA分子完全可以通过电测量进行检测,并且DNA的长度与所产生的电流瞬变的长度相关。在这些测量值中,电流瞬变是由纵向测量值推导,其中未使用纳米电极,而是使用了位于装置的阳极储槽和阴极储槽处的微电极。这些纵向电流测量值可以在电驱动的易位事件期间得到。从这些测量值可知,瞬变的持续时间可以与DNA长度相关。尽管在这些情况下进行了光学测量,但这些测量对于电转导来说不是必需的。进行光学测量只是为了使电学测量的性质具体化。

[0389] 实施例13—使用电化学法在分层纳米线中产生纳米间隙

[0390] 已开发出一种两步骤方法在分节段纳米线中产生纳米间隙。这些步骤包括:(1)电沉积多节段纳米线;及(2)由包埋的牺牲层电化学蚀刻间隙。为了描述电沉积技术,使用带纳米孔的聚碳酸酯模板(Whatman,约 10^8 个纳米孔/ cm^2 ,并且指定孔径是100nm)在两个电极电池中作为阴极。为了产生导电电极,用金对聚碳酸酯模板的一侧进行溅镀。阳极是铂网。按以下次序制造五个纳米线区段:FeNiCo/Au/Cu/Au/FeNiCo。FeNiCo纳米线区段在纳米线的顶部和底部,并且充当以磁力操作所得纳米线的一种方式。FeNiCo是使用 $-25\text{mA}/\text{cm}^2$ 的脉冲沉积(20%孔隙率)、0.5的占空比及2秒的通断时间,在pH 2和40℃温度下由含有以下各物的电解质沉积得到的富含Fe的合金(约64wt%Fe、31wt%Ni、5wt%Co):0.72M氨基磺酸镍、0.155M硫酸亚铁、0.005M硫酸钴、0.5M硼酸、0.001M十二烷基硫酸钠、0.011M抗坏血酸。金区段是在 $-1\text{mA}/\text{cm}^2$ 的电流密度和60℃温度下,通过使用商购的金溶液(TG-25E RTU,

Technic Inc.) 制造。铜中间层是在 $-3\text{mA}/\text{cm}^2$ 下趋电性沉积。铜电解质含有0.15M硫酸铜、0.01M硫酸钠、0.45M柠檬酸钠及0.40M硼酸。沉积后,将膜溶解于二氯甲烷中,然后用乙醇洗涤三次。铜充当牺牲层并且沉积时间与间隙大小成比例。在第二个步骤中,通过对放在导电碳表面上的纳米线进行蚀刻,以电化学方式控制铜层的溶解。相对于SCE,所施加的工作电位是+0.1V,并且溶液是与用于镀铜相同的柠檬酸盐-硼酸电解质。

[0391] 所述程序成功地产生了带有间隙的纳米线,这些间隙可以控制在不同尺寸(参见图53A-53B)。一个重要特征在于,基于所施加的电位的选择,在蚀刻期间,FeNiCo区和Au区未受影响。Au未被溶解是因为其惰性比铜强,并且FeNiCo未受影响是因为所选电位和电解质在其表面上以核-壳型配置产生钝化膜。在所述技术的开发中使用了较厚、较大直径的纳米线,并且将扩展到更小直径的线。为了产生更分散的间隙和更窄的间隙大小,应当用脉冲电镀和反向脉冲电镀方案检查铜层的沉积。

[0392] 尽管在本文中已经详细描绘并描述了优选实施方案,但相关领域的技术人员应显而易见的是,在不偏离本发明精神的情况下可以作出各种修改、添加、取代等,并因此,这些都被视为在如所附权利要求书中所界定的本发明的范围内。

序列表

<110> 康奈尔大学
路易斯安那州州立大学及农业机械学院管理委员会
北卡罗来纳大学查佩尔山分校
美国东北大学

<120> 生物分子处理平台及其用途

<130> 29543.7201 (6013-02-PC)

<150> US 61/762,671

<151> 2013-02-08

<160> 6

<170> PatentIn 3.5版

<210> 1

<211> 36

<212> DNA

<213> 人工

[0001]

<220>

<223> 寡核苷酸适配体

<220>

<221> misc_feature

<222> (3)..(3)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (5)..(5)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (7)..(7)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n是a、c、g或t

[0002]

<220>
<221> misc_feature
<222> (11).. (20)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (27).. (30)
<223> n是a、c、g或t

<400> 1
cgntngnanc nnnnnnnnnn ccgagcnnnn gctcgg

36

<210> 2
<211> 11
<212> DNA
<213> 人工

<220>
<223> 寡核苷酸适配体

<220>
<221> misc_feature
<222> (2).. (2)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (4).. (4)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (6).. (6)
<223> n是a、c、g或t

<220>
<221> misc_feature
<222> (8).. (8)
<223> n是a、c、g或t

<400> 2
gntncnancg t

11

[0003]

<210> 3
 <211> 38
 <212> DNA
 <213> 人工

<220>
 <223> 寡核苷酸适配体

<220>
 <221> misc_feature
 <222> (3)..(12)
 <223> n是a、c、g或t

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n是a、c、g或t

<220>
 <221> misc_feature
 <222> (17)..(17)
 <223> n是a、c、g或t

<220>
 <221> misc_feature
 <222> (19)..(19)
 <223> n是a、c、g或t

<220>
 <221> misc_feature
 <222> (21)..(21)
 <223> n是a、c、g或t

<220>
 <221> misc_feature
 <222> (29)..(32)
 <223> n是a、c、g或t

<400> 3
 ggnnnnnnnn nncgntngna ncccgagcnn nngctcgg

38

<210> 4
 <211> 13
 <212> DNA

[0004]

<213> 人工

<220>

<223> 寡核苷酸适配体

<220>

<221> misc_feature

<222> (5)..(5)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (7)..(7)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (11)..(11)

<223> n是a、c、g或t

<400> 4

tttgntncna ncg

13

<210> 5

<211> 25

<212> DNA

<213> 人工

<220>

<223> 寡核苷酸适配体

<220>

<221> misc_feature

<222> (5)..(5)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (7)..(7)

[0005]

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (9).. (9)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (11).. (11)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (14).. (23)

<223> n是a、c、g或t

<400> 5

tttgntnena negnnnnnnn nnncc

25

<210> 6

<211> 26

<212> DNA

<213> 人工

<220>

<223> 寡核苷酸适配体

<220>

<221> misc_feature

<222> (5).. (5)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (7).. (7)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (9).. (9)

<223> n是a、c、g或t

<220>

<221> misc_feature

[0006]

<222> (11).. (11)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (14).. (23)

<223> n是a、c、g或t

<400> 6

tttgntncna ncgnnnnnnn nnncca

26

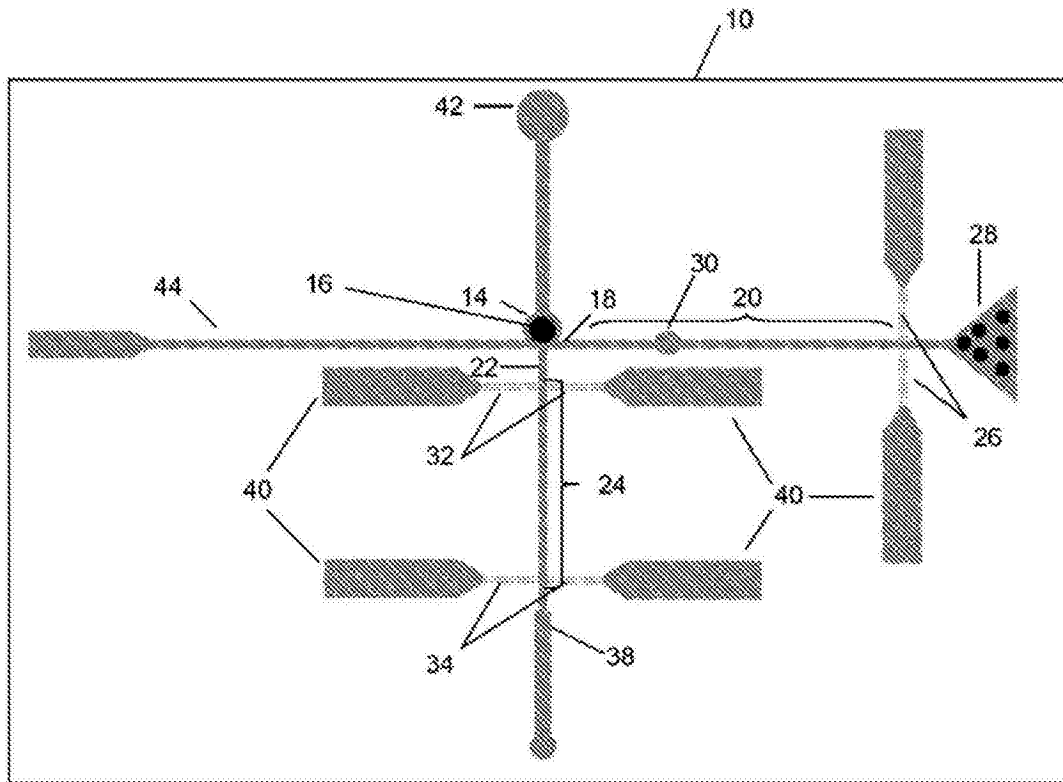


图1

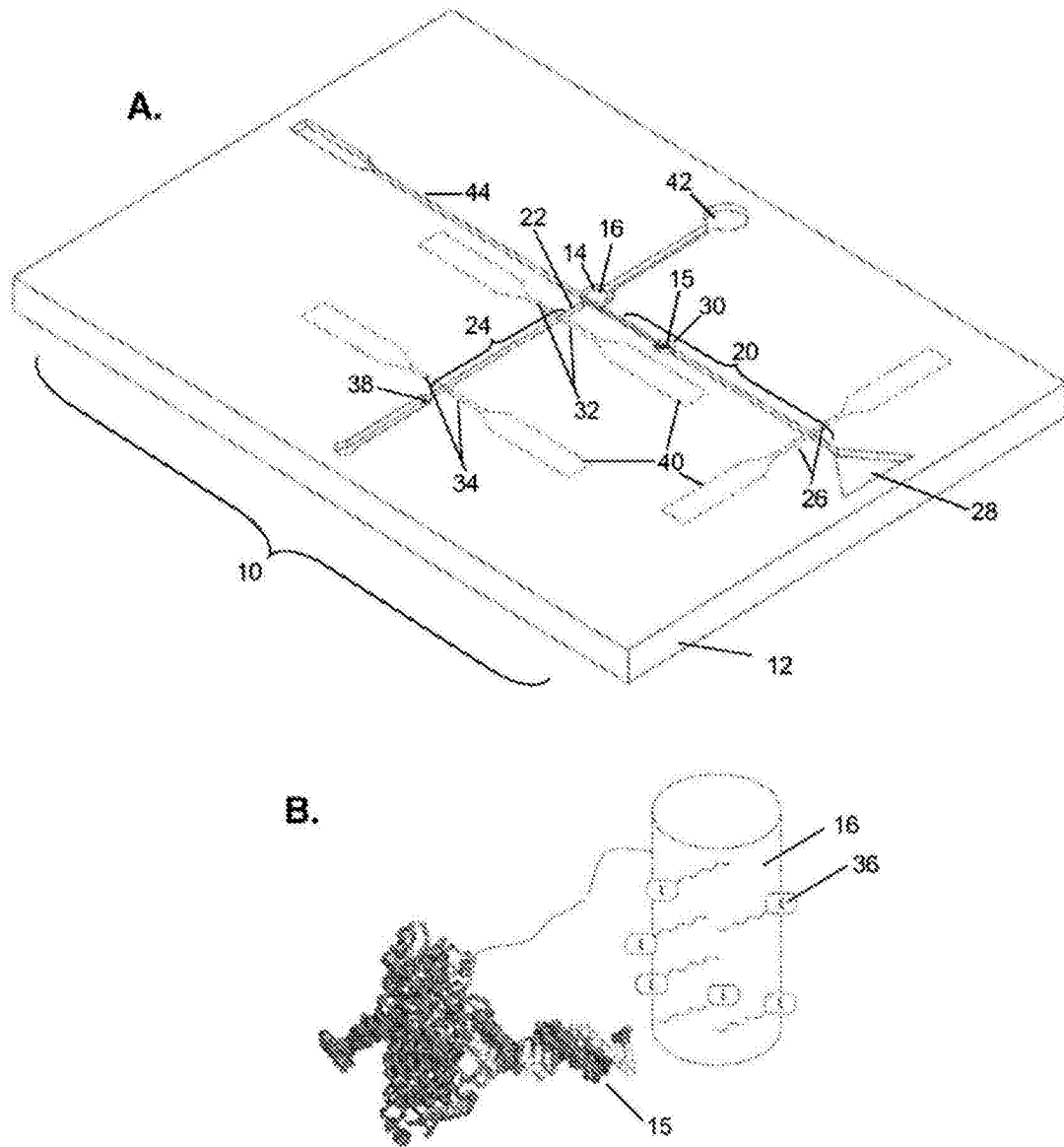


图2A-2B

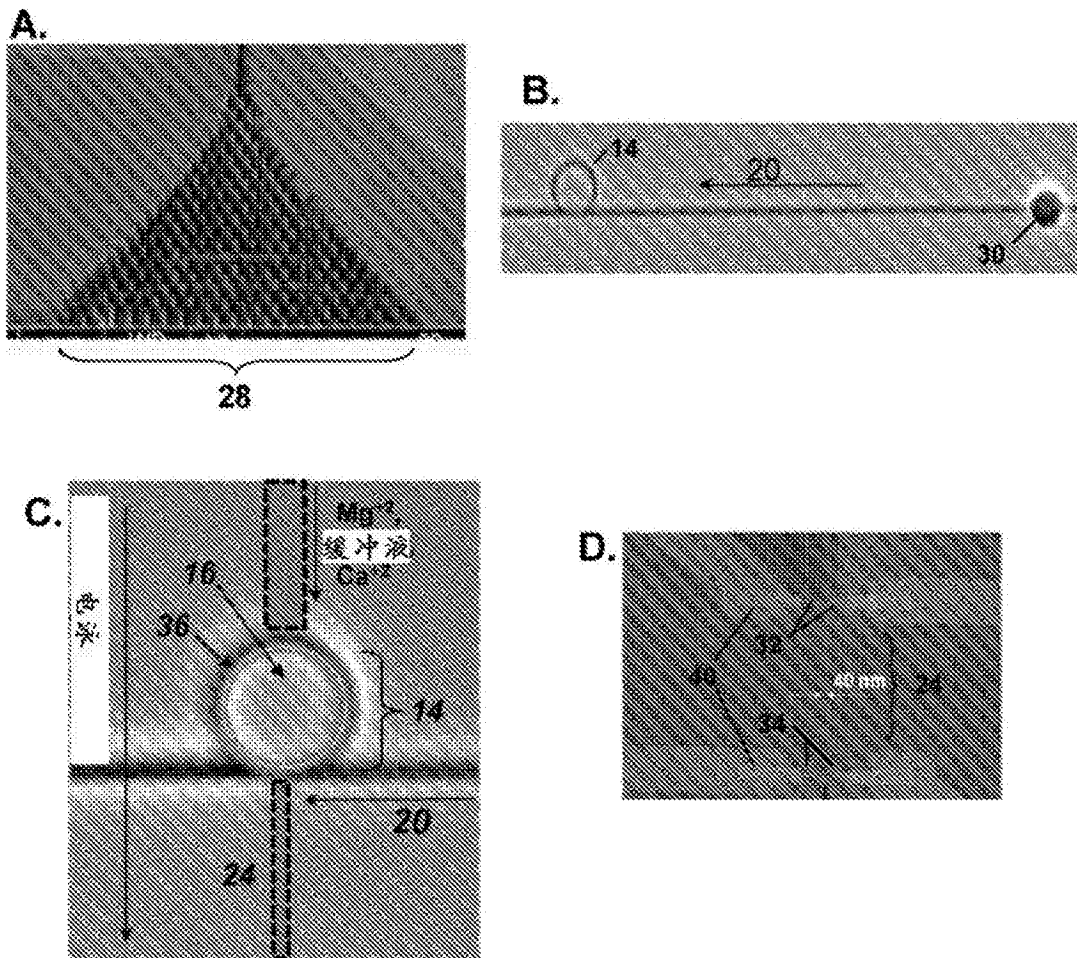


图3A-3D

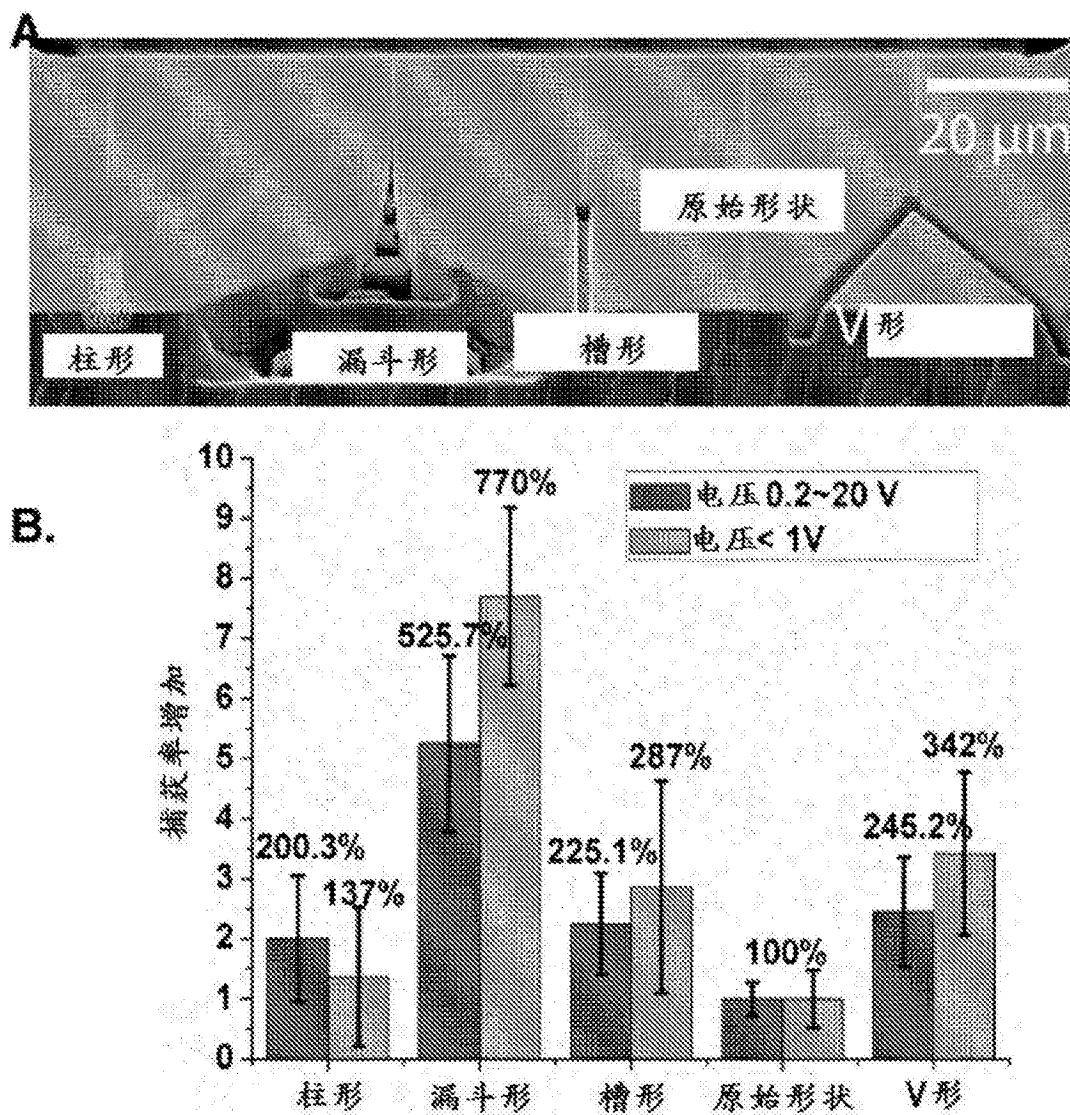


图4A-4B

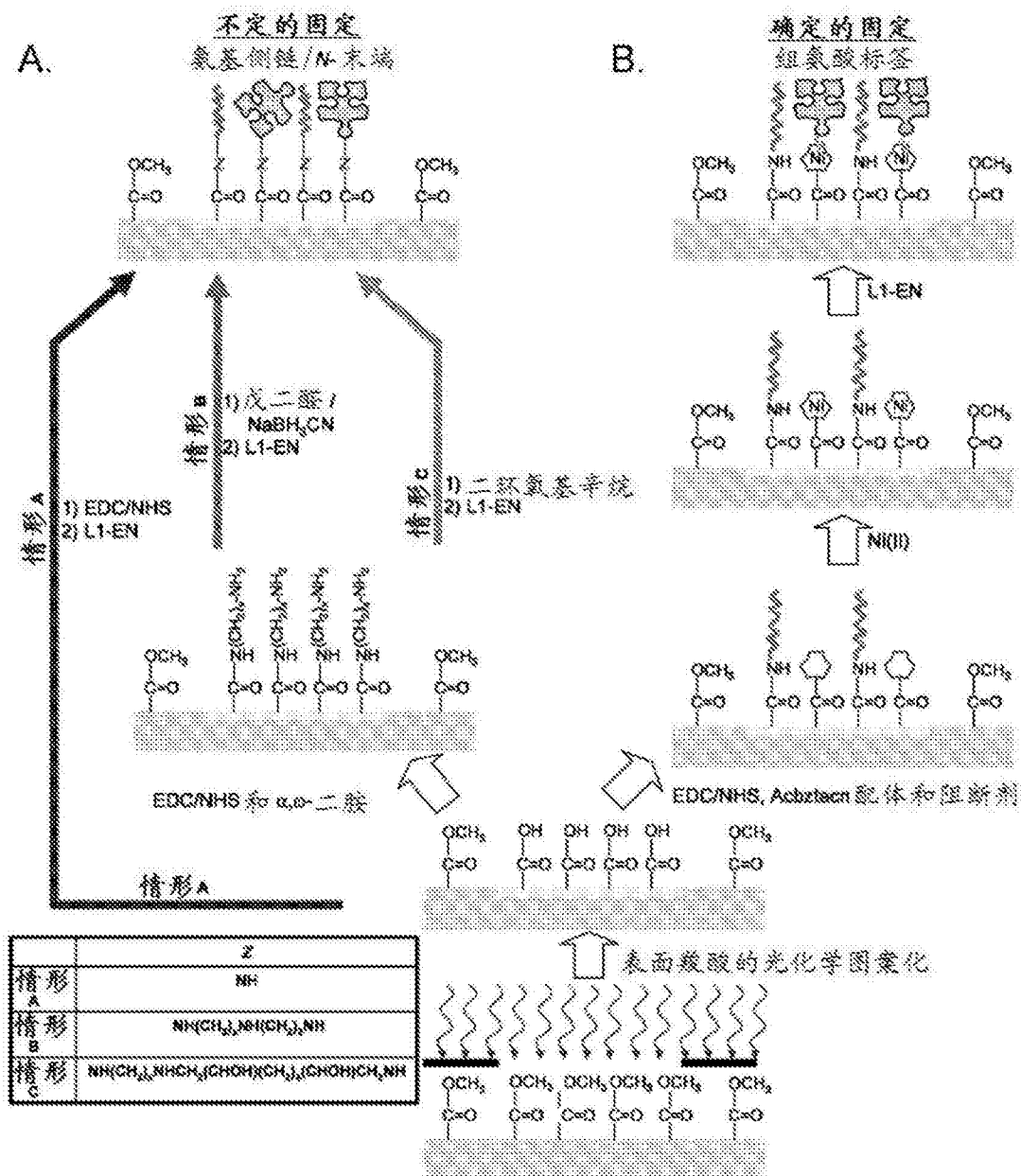


图5A-5B

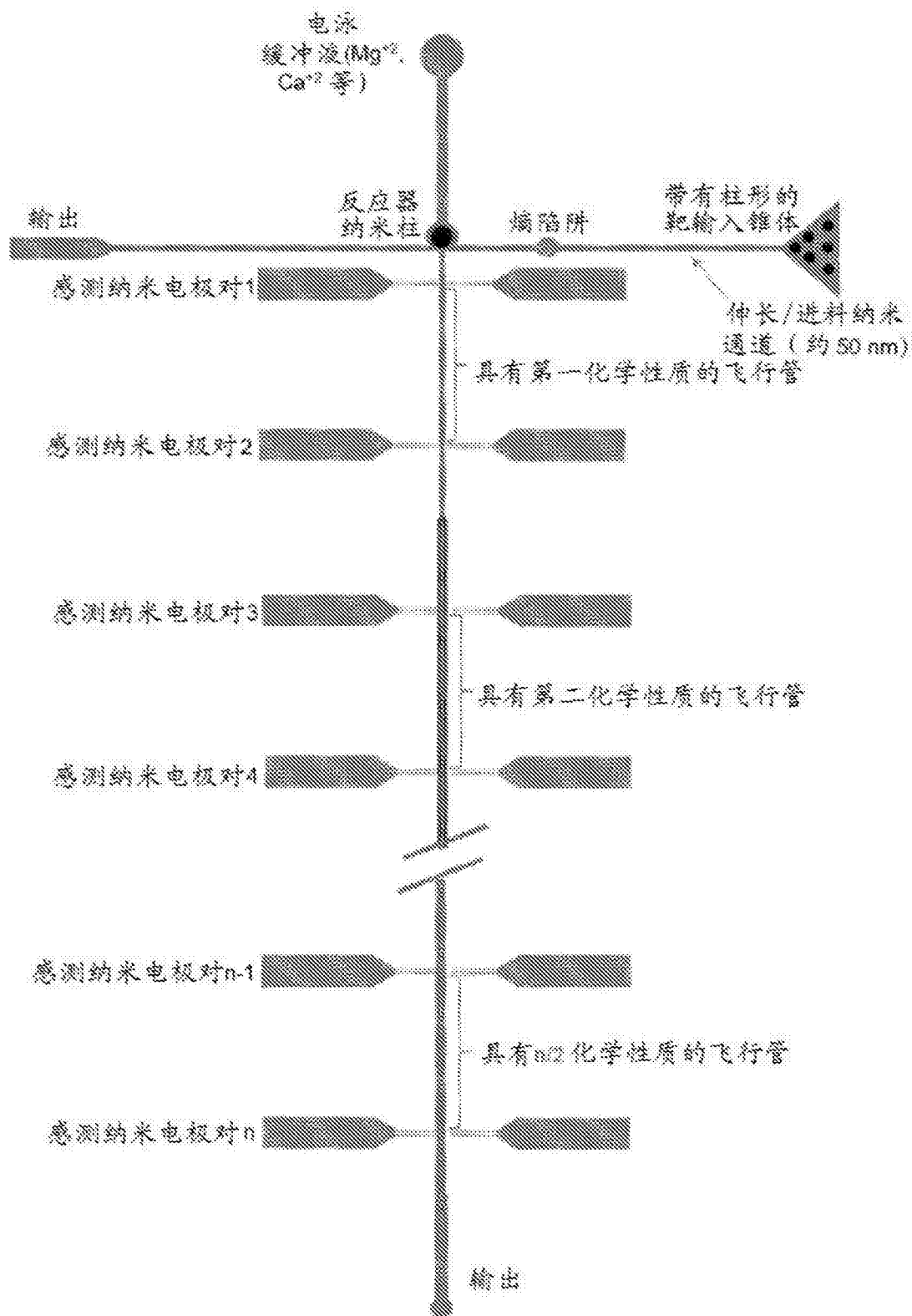
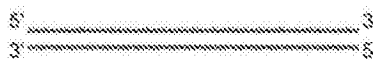


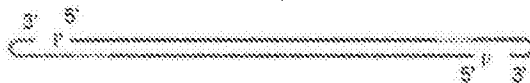
图7

1. (仅显示出一个基因组双链DNA片段。)



2. 将通用适配体与每个DNA片段连接。

适配体含有可选用的链ID和短发夹区。



3. 3' 发夹末端使用带有链置换活性的聚合酶延伸。所得双链DNA特别适合于X-TOF测序。测序链在两端上含有链ID，从而允许其起源的独特验证。此外，由于对原始基因组链进行了测序，故所有碱基修饰都得到保留。

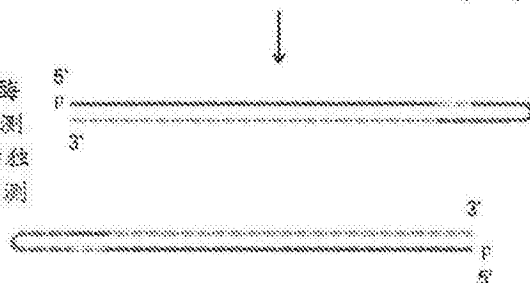
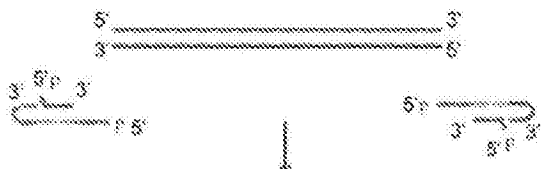


图8A

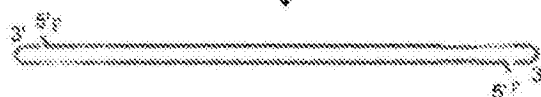
1. (仅显示出一个基因组双链DNA片段。)

对通用适配体进行延伸并加尾。



2. 将通用适配体与每个DNA片段连接。

适配体含有可选用的链ID和短发夹区。



3. 3' 发夹末端使用带有链置换活性的聚合酶延伸。所得双链DNA特别适合于X-TOF测序。测序链在两端上含有链ID，从而允许其起源的独特验证。此外，由于对原始基因组链进行了测序，故所有碱基修饰都得到保留。

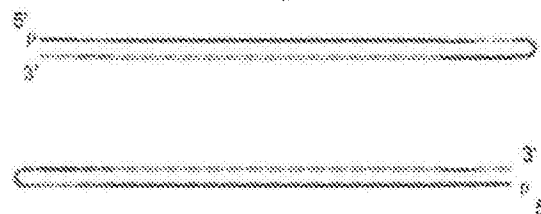
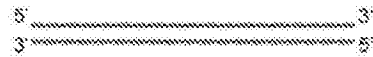
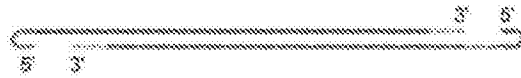


图8B

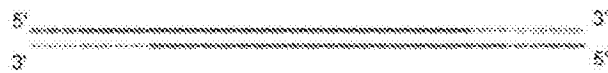
1. (仅显示出富集相关靶双链DNA。)



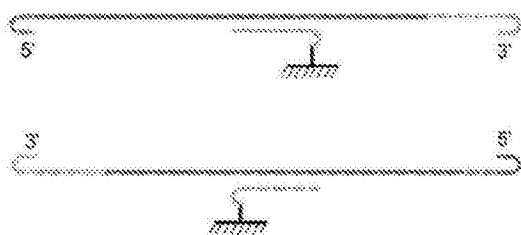
2. 将通用适配体与每个DNA片段连接。适配体含有可选用的患者ID、链ID和短发夹区。



3. 3'连接子端使用缺乏5'-3'活性的聚合酶延伸通过发夹区。



4. 使双链DNA解链并形成单链DNA。通过与固定于固体支撑物的互补链杂交来捕获上部靶链和下部靶链。互补链在空间上分离，由此互补链和靶链彼此将无法再杂交。



5. 3'发夹状末端使用带有5'-3'核酸外切酶活性的聚合酶延伸，由此从固体支撑物释放出双链靶DNA并且还释放出5'磷酸酯末端，特别适合于X-TOF测序。测序链在两端上含有患者ID和链ID，从而允许其起源的独特验证。此外，由于对原始基因组链进行了测序，故所有碱基修饰都得到保留。

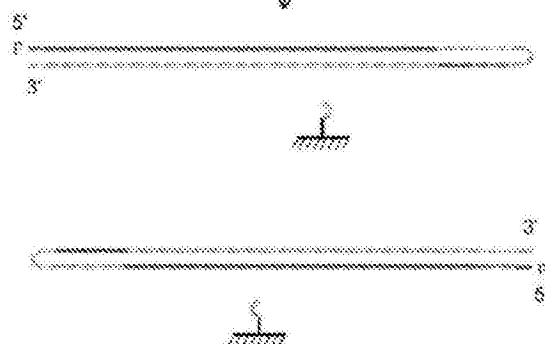
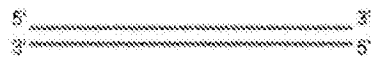
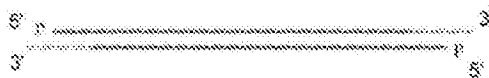


图9

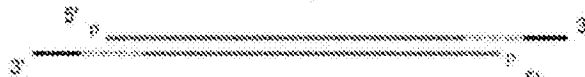
1. (仅显示出一个基因组双链DNA片段。)



2. 将通用适配体与每个DNA片段连接。适配体含有可选用的链ID。



3. 3'连接器端使用末端转移酶加尾。



4. 使引物与加尾的适配体链的3'端杂交，并且引物使用聚合酶延伸。由此产生特别适合于X-TOF测序的双链DNA。测序链在两端上含有链ID，从而允许其起源的独特验证。此外，由于对原始基因组链进行了测序，故所有碱基修饰都得到保留。

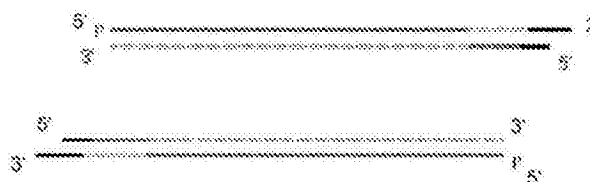
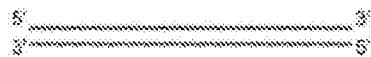
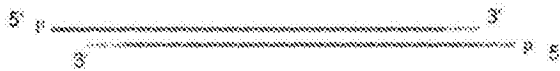


图10

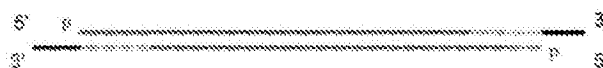
1 (仅显示出富集相关靶双链DNA,)



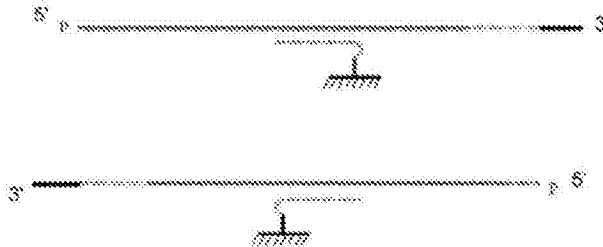
2. 将通用适配体与每个DNA片段连接。
适配体含有可选用的患者ID和链ID。



3. 3'连接于端用聚合酶延伸并使用
末端转移酶加尾。



4. 使双链DNA解链并形成单链DNA。通
过与固定于固体支撑物的互补链杂
交来捕获上部靶链和下部靶链。互
补链在空间上分离,由此互补链和
靶链彼此将无法再杂交。



5. 使引物与靶单链的3'端杂交,
并且引物使用聚合酶延伸,由此
产生特别适合于X-TOF测序的双
链靶DNA。测序链在两端上含有
患者ID和链ID,从而允许其起源
的独特验证。此外,由于对原始
基因组链进行了测序,故所有碱
基修饰都得到保留。

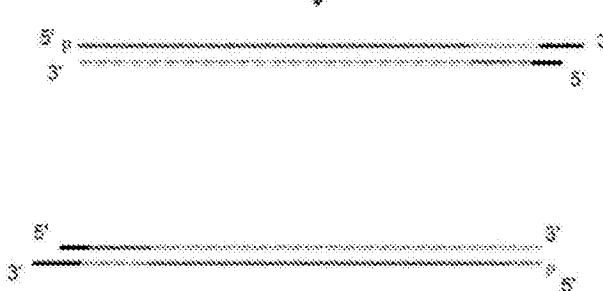


图11

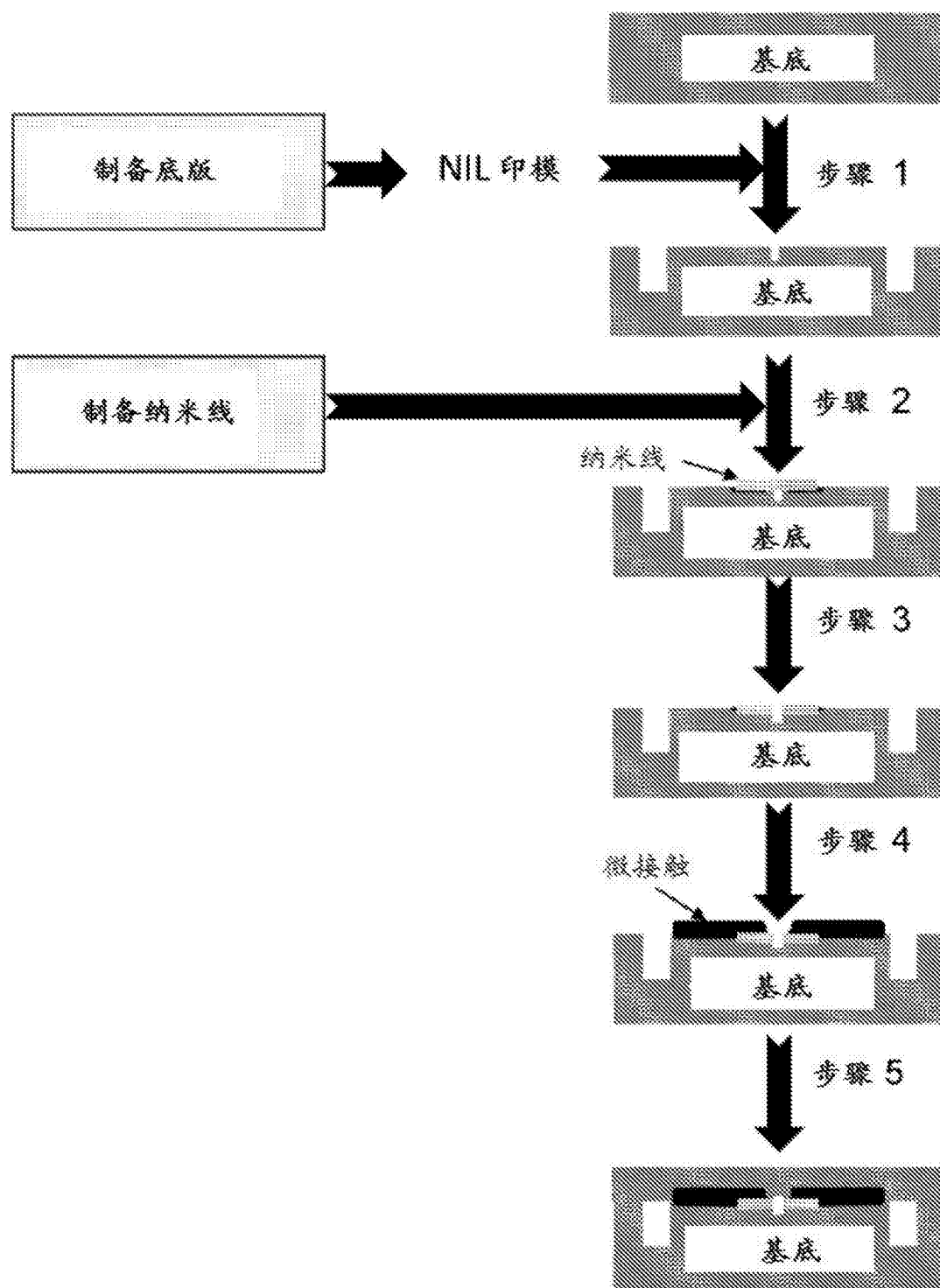


图12

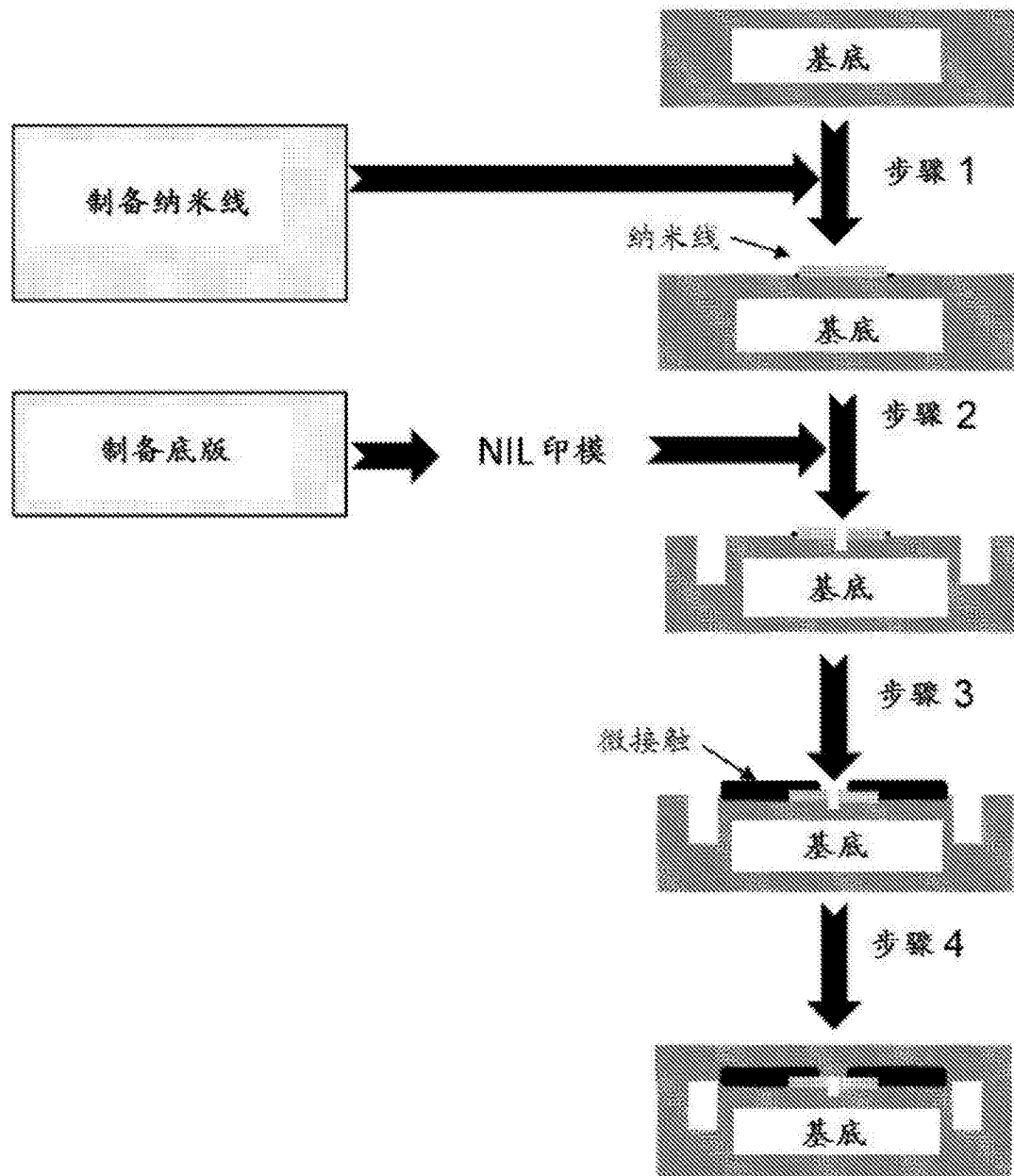


图13

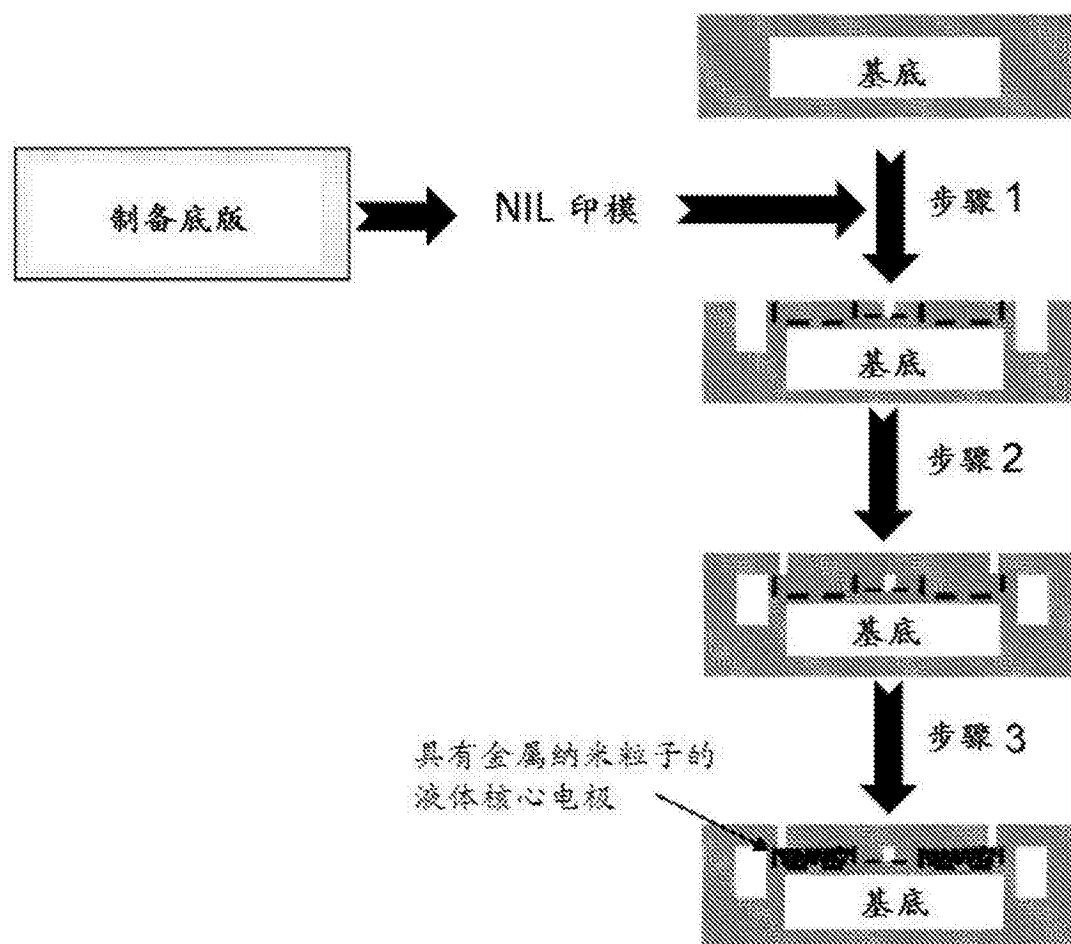


图14

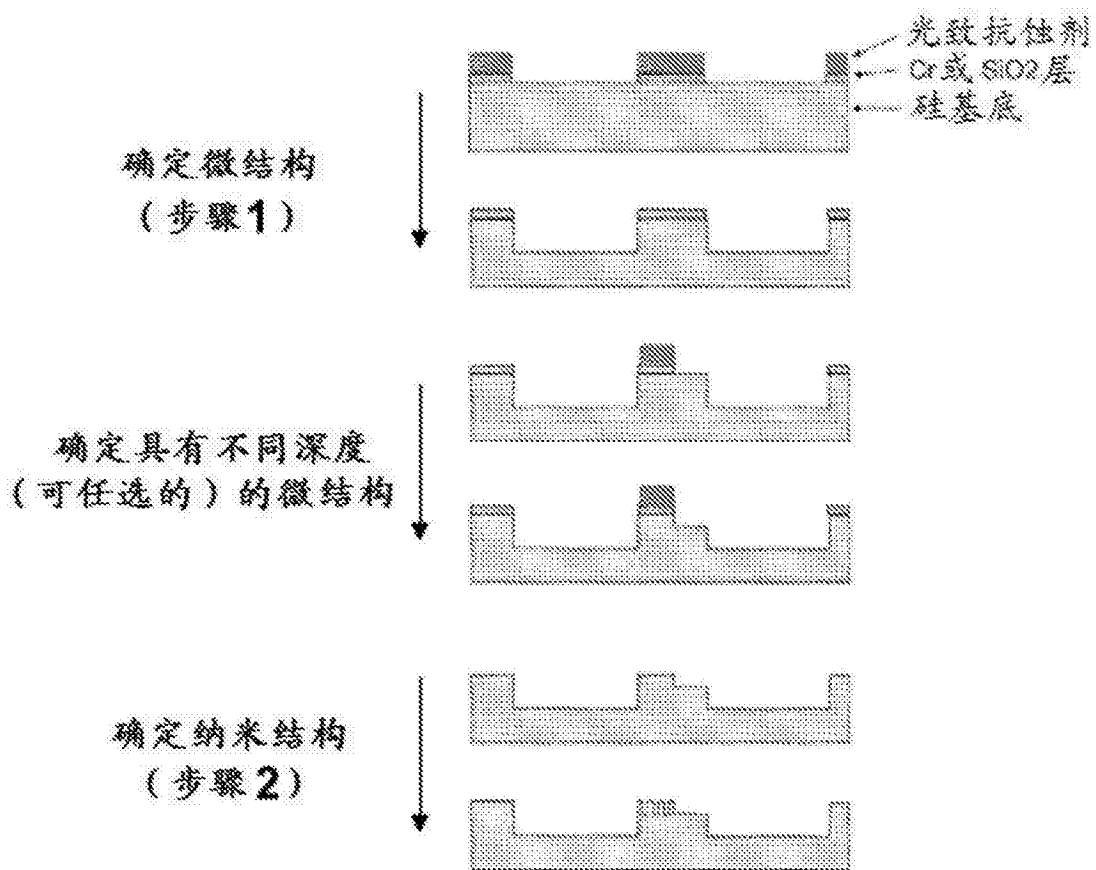


图15

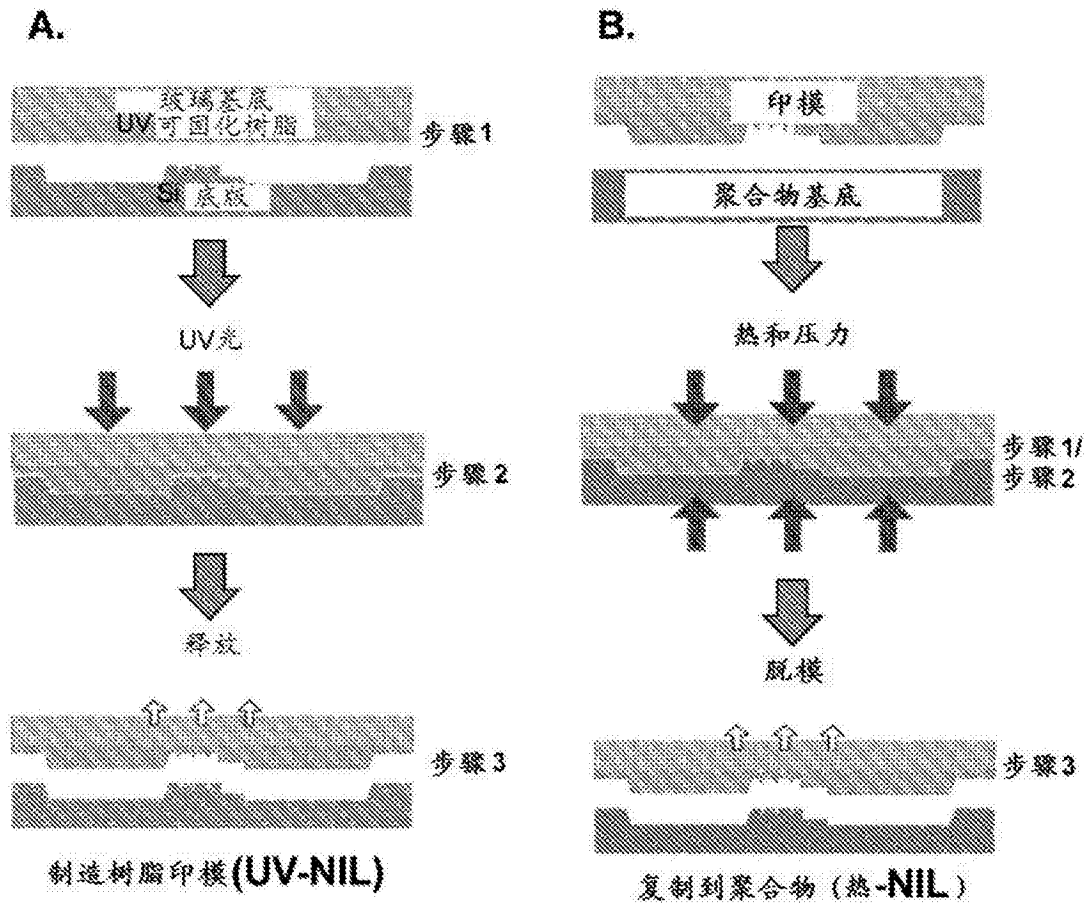


图16A-16B

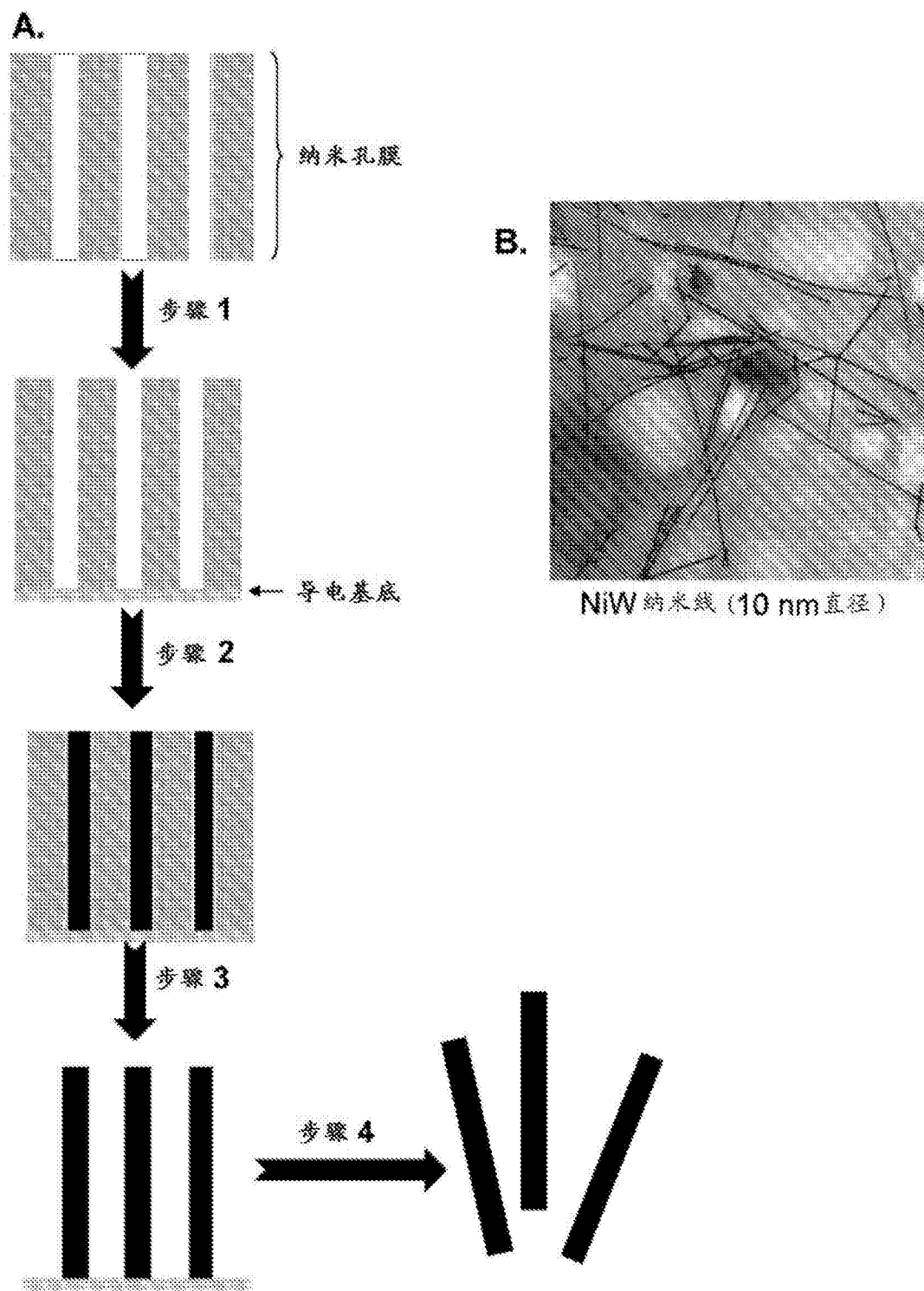


图17A-17B

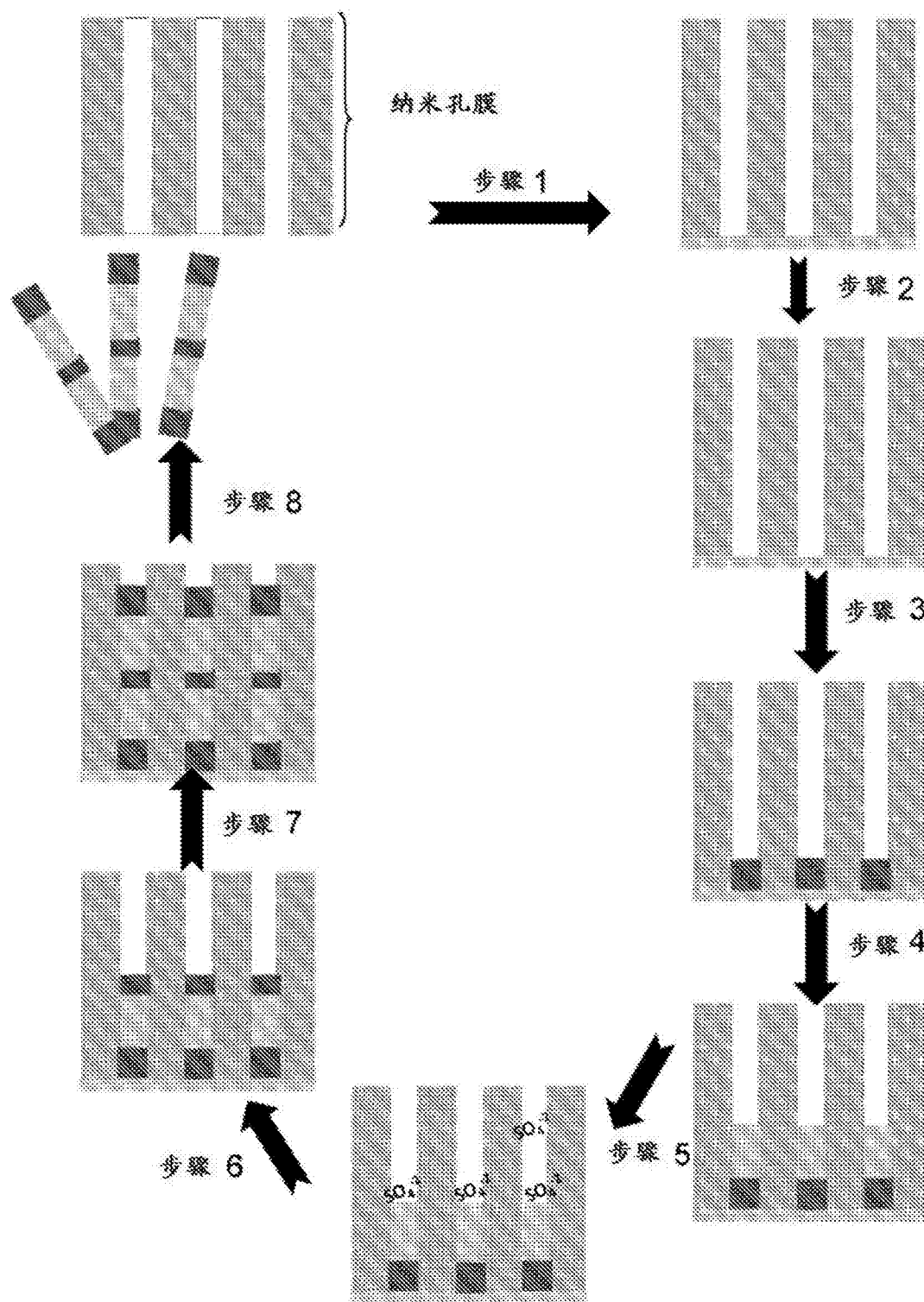


图18

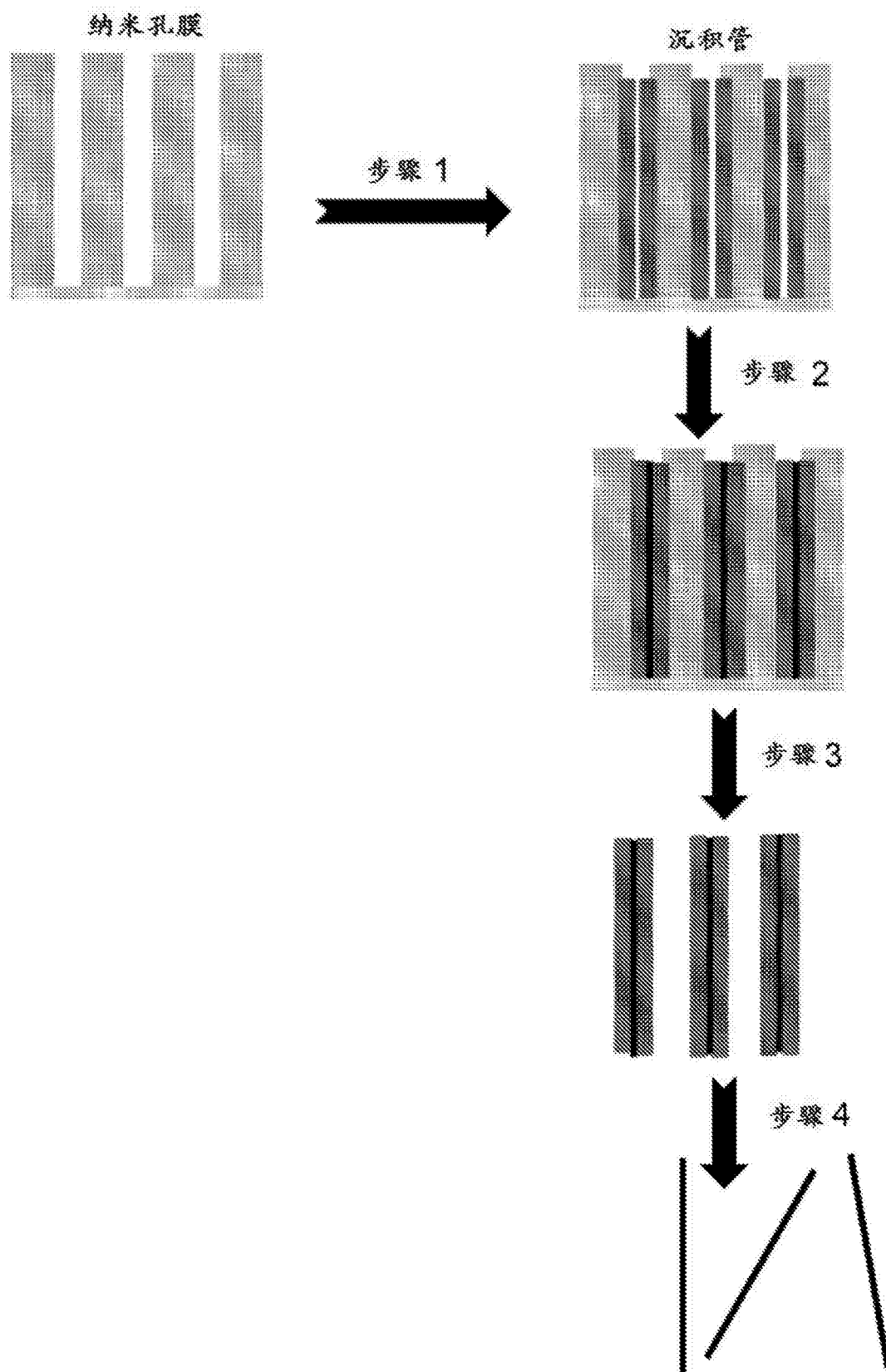


图19

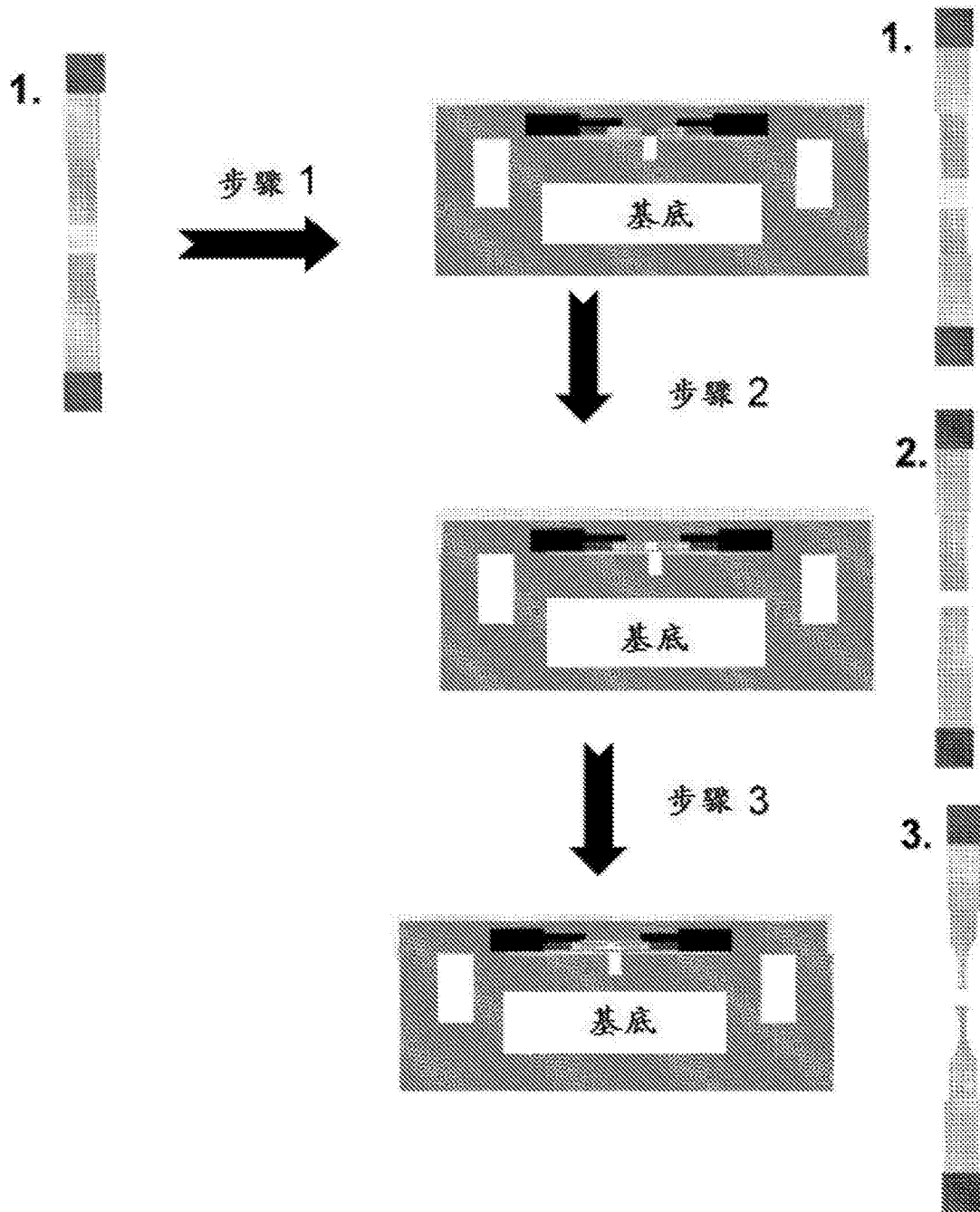


图20

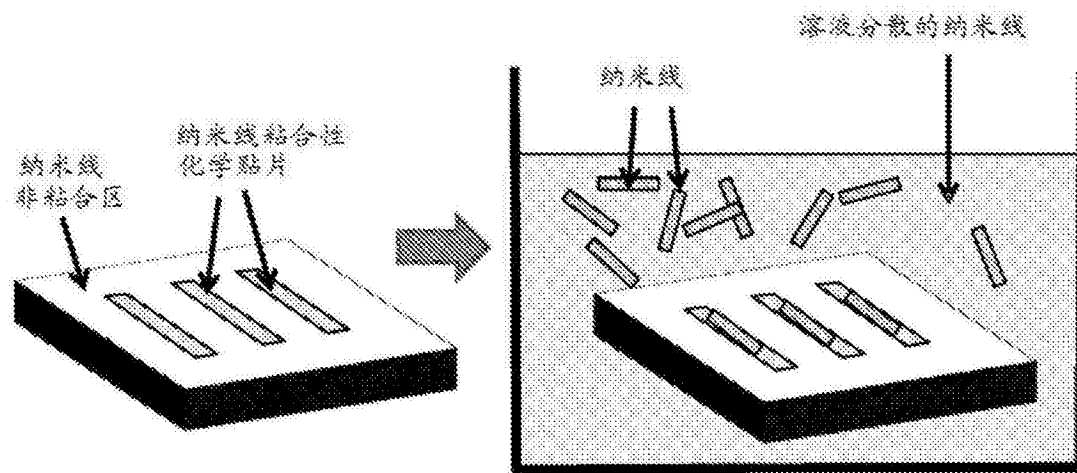


图21

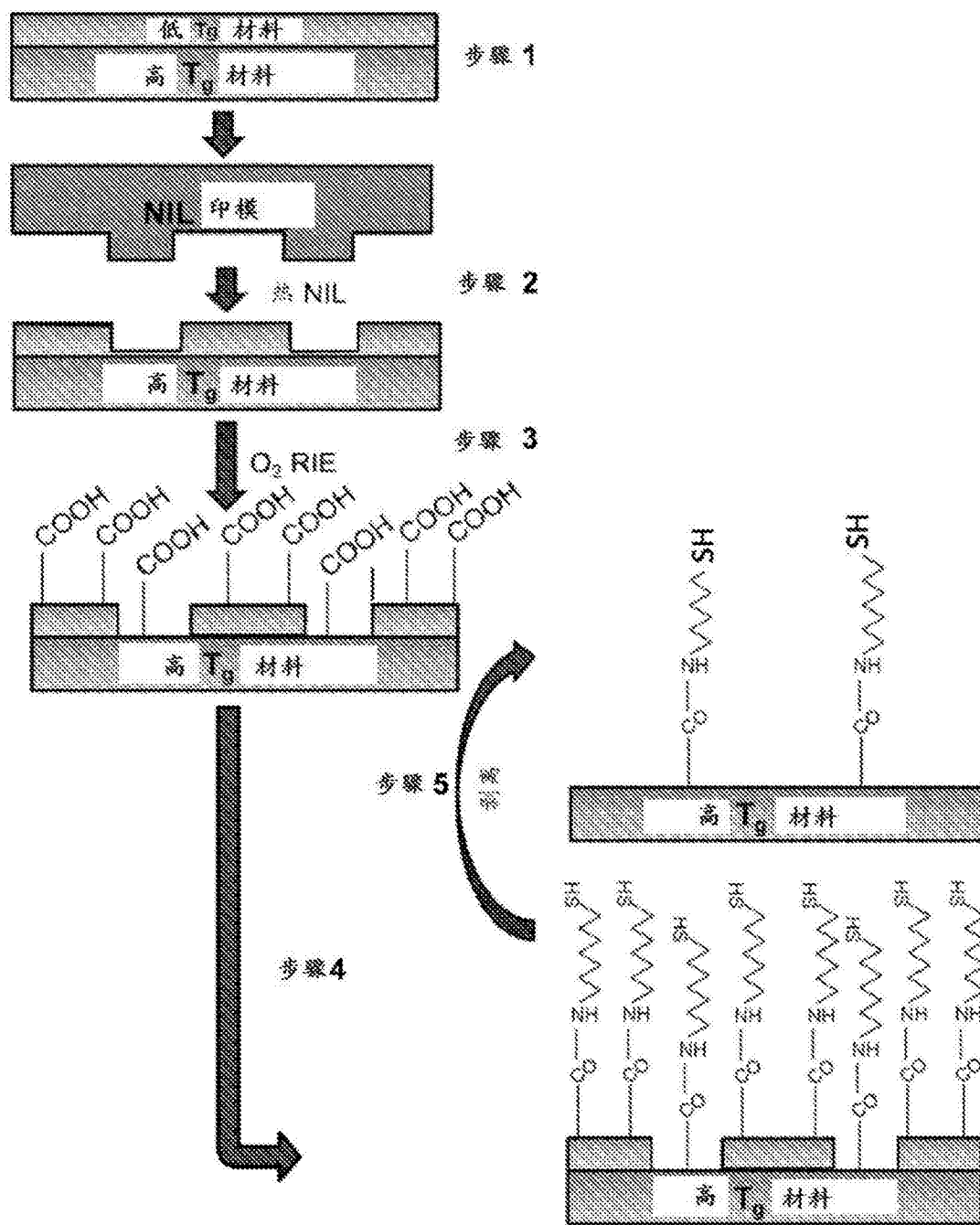


图22

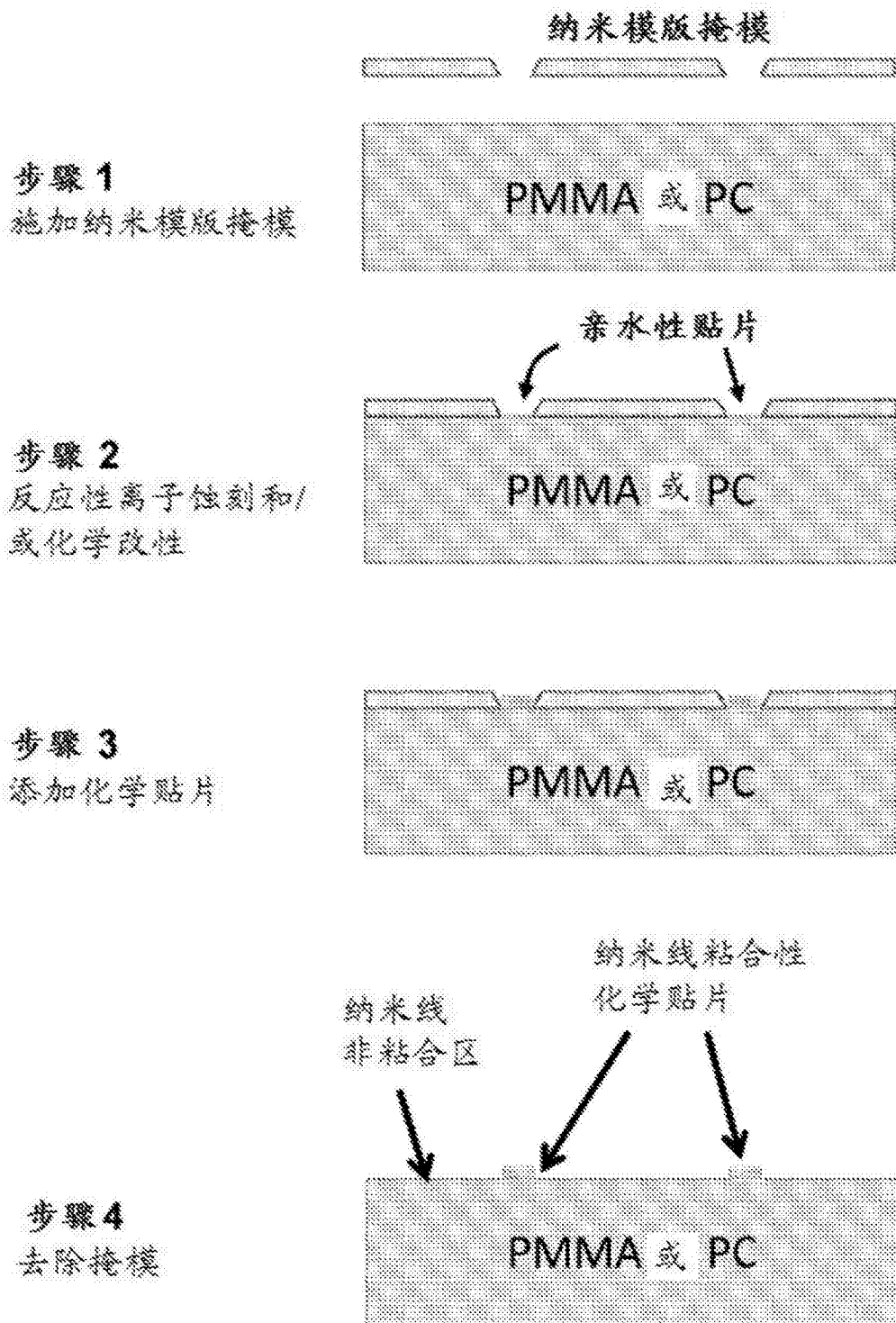


图23

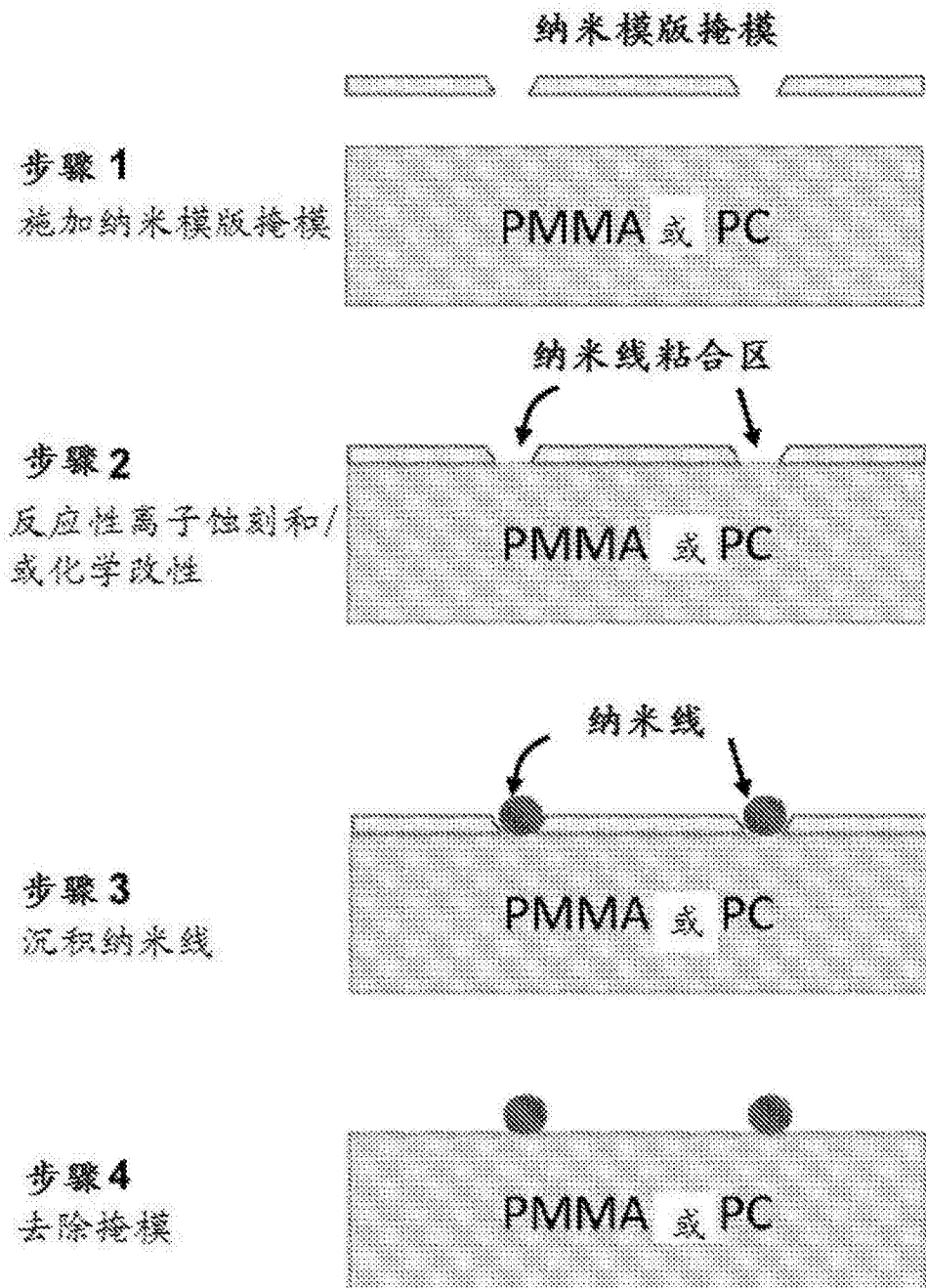


图24

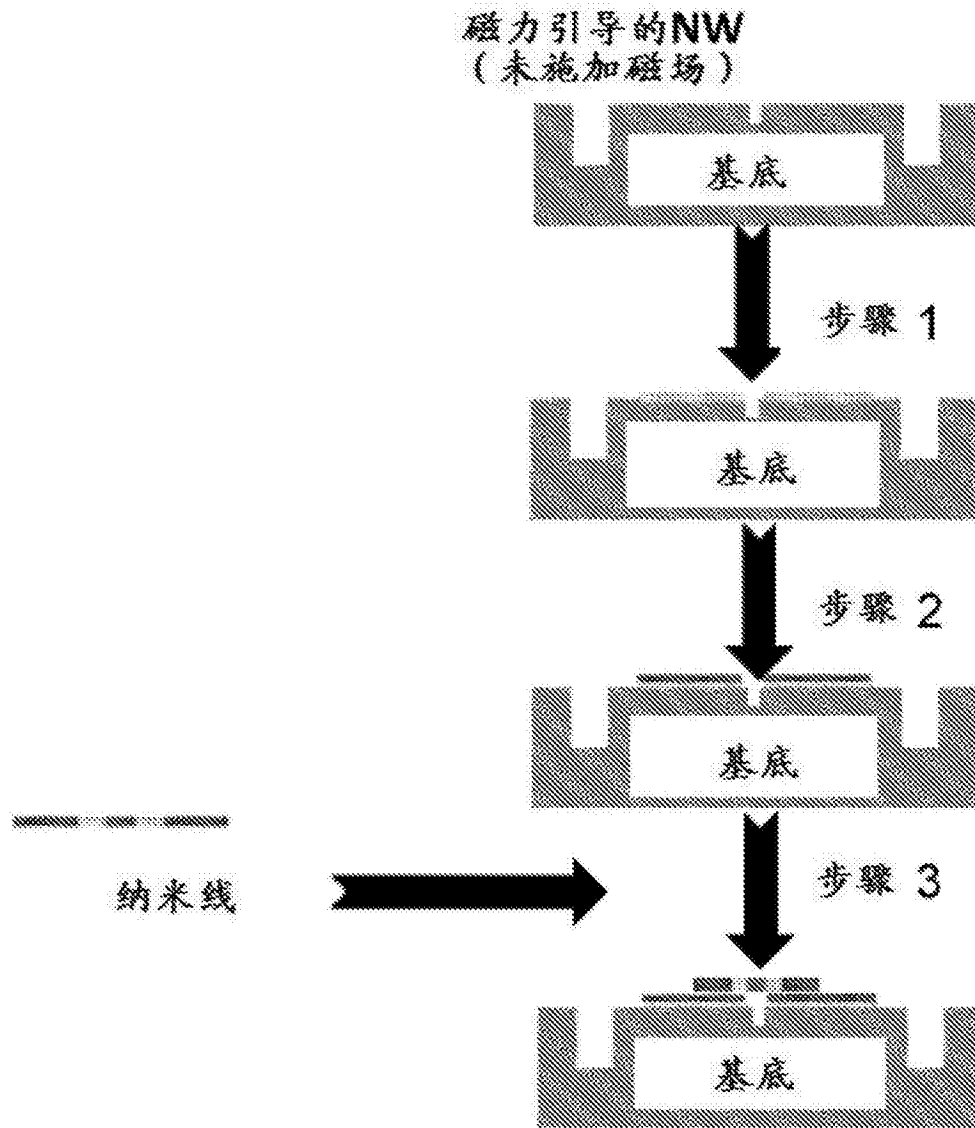


图25

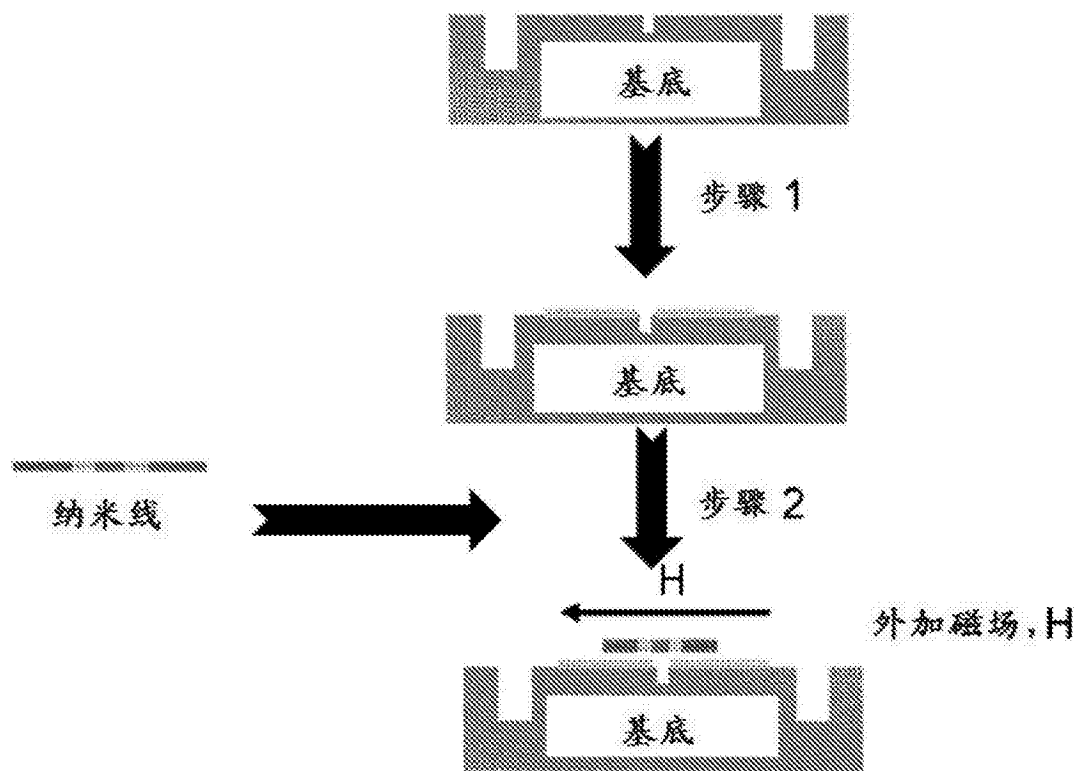


图26

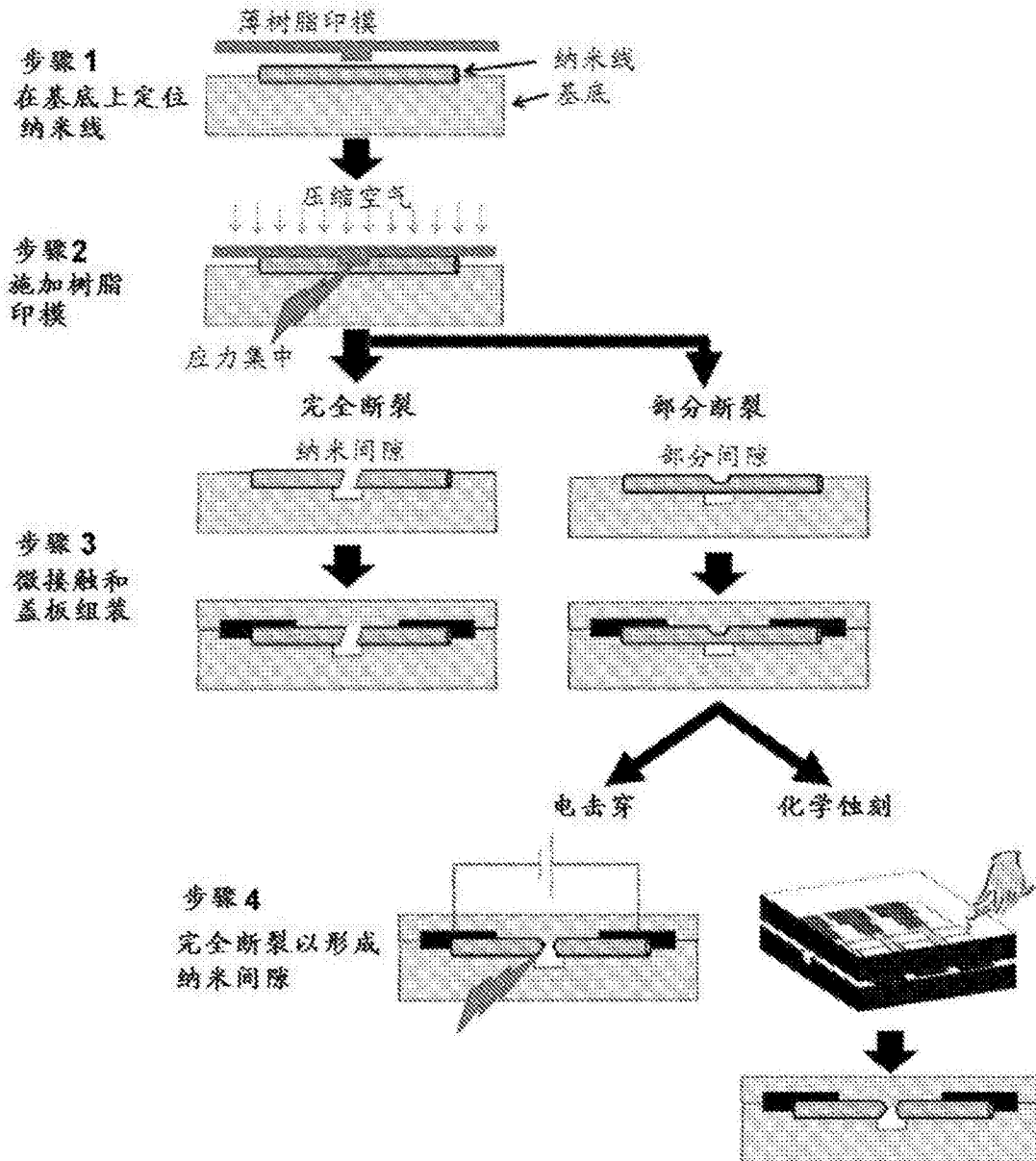


图27

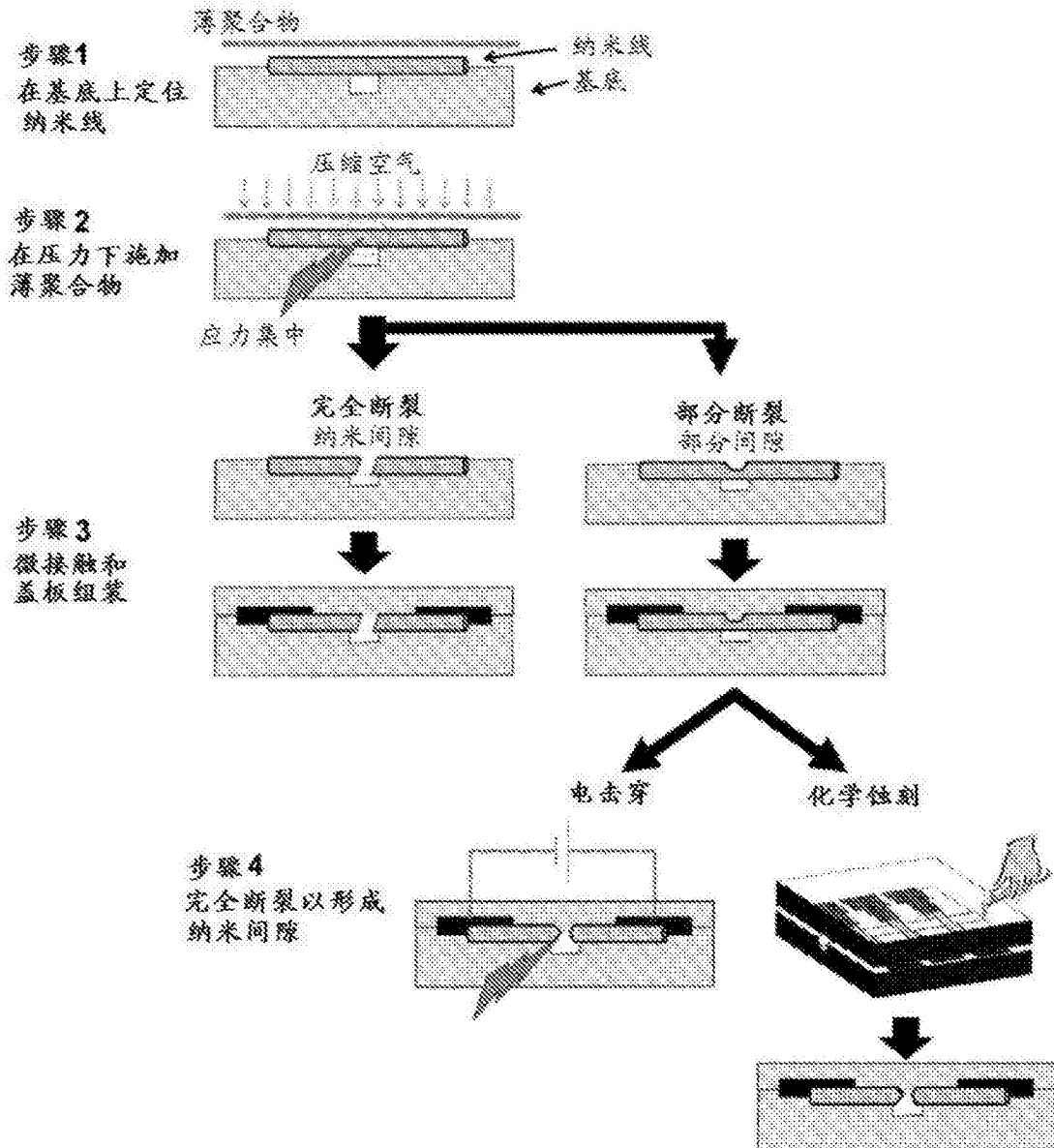


图28

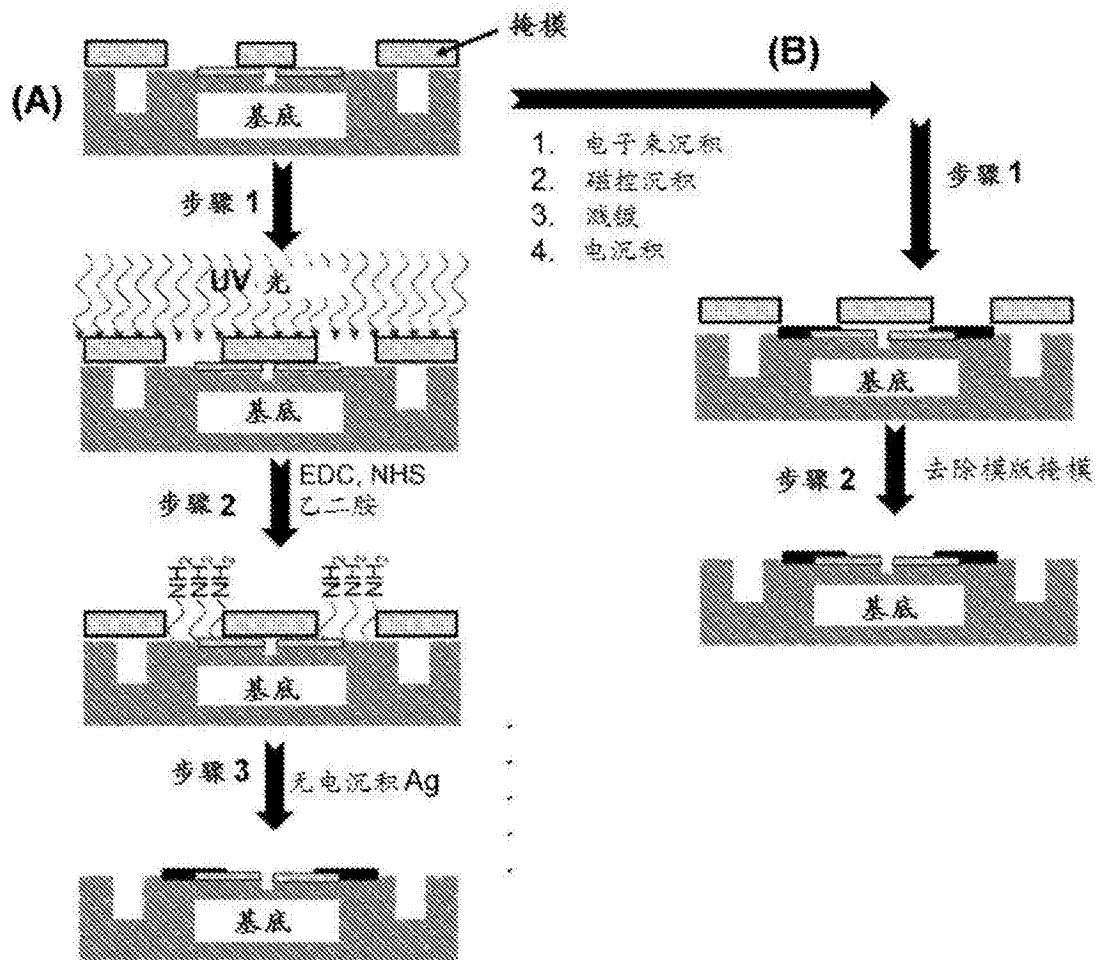


图29A-29B

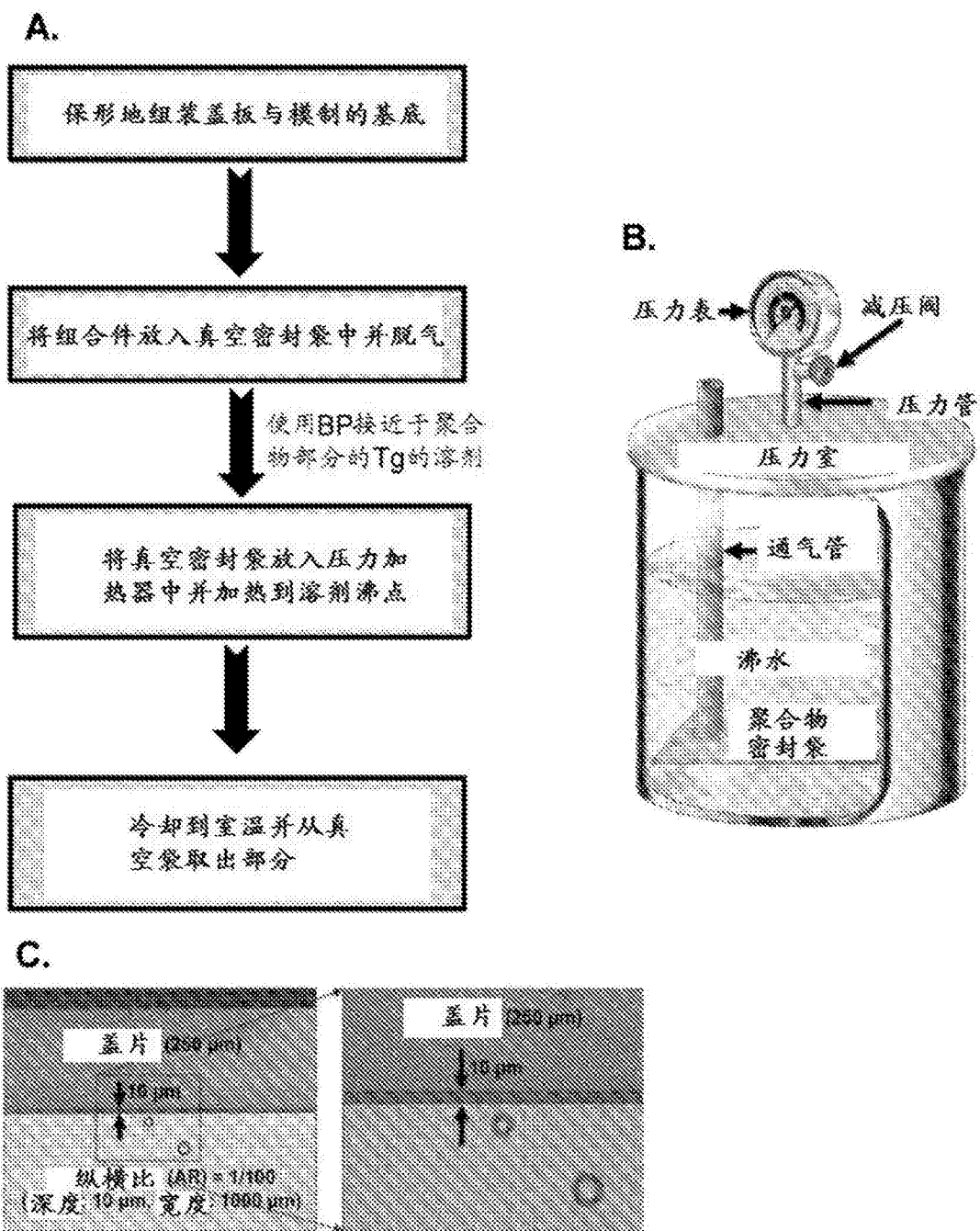


图30A-30C

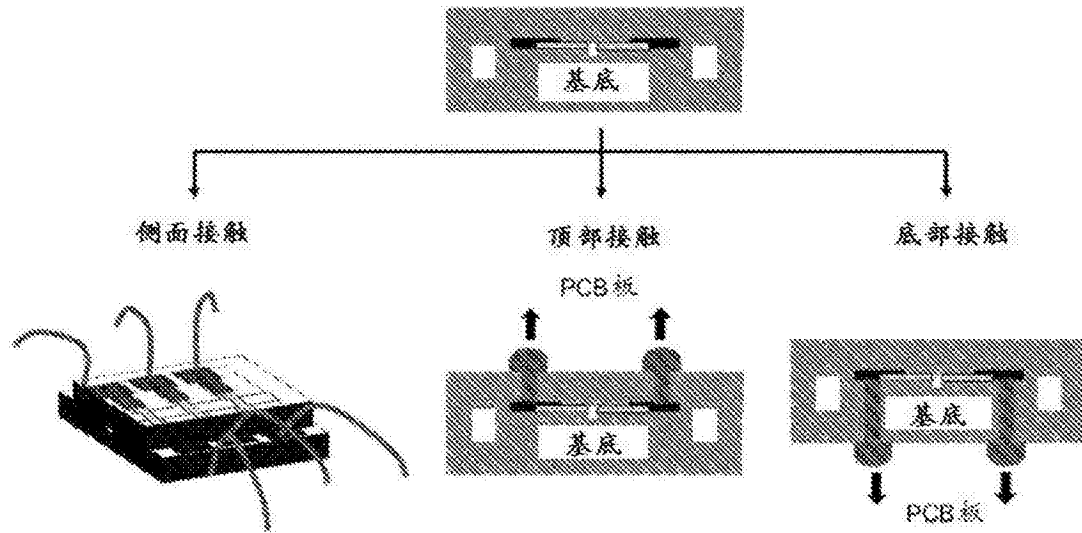


图31

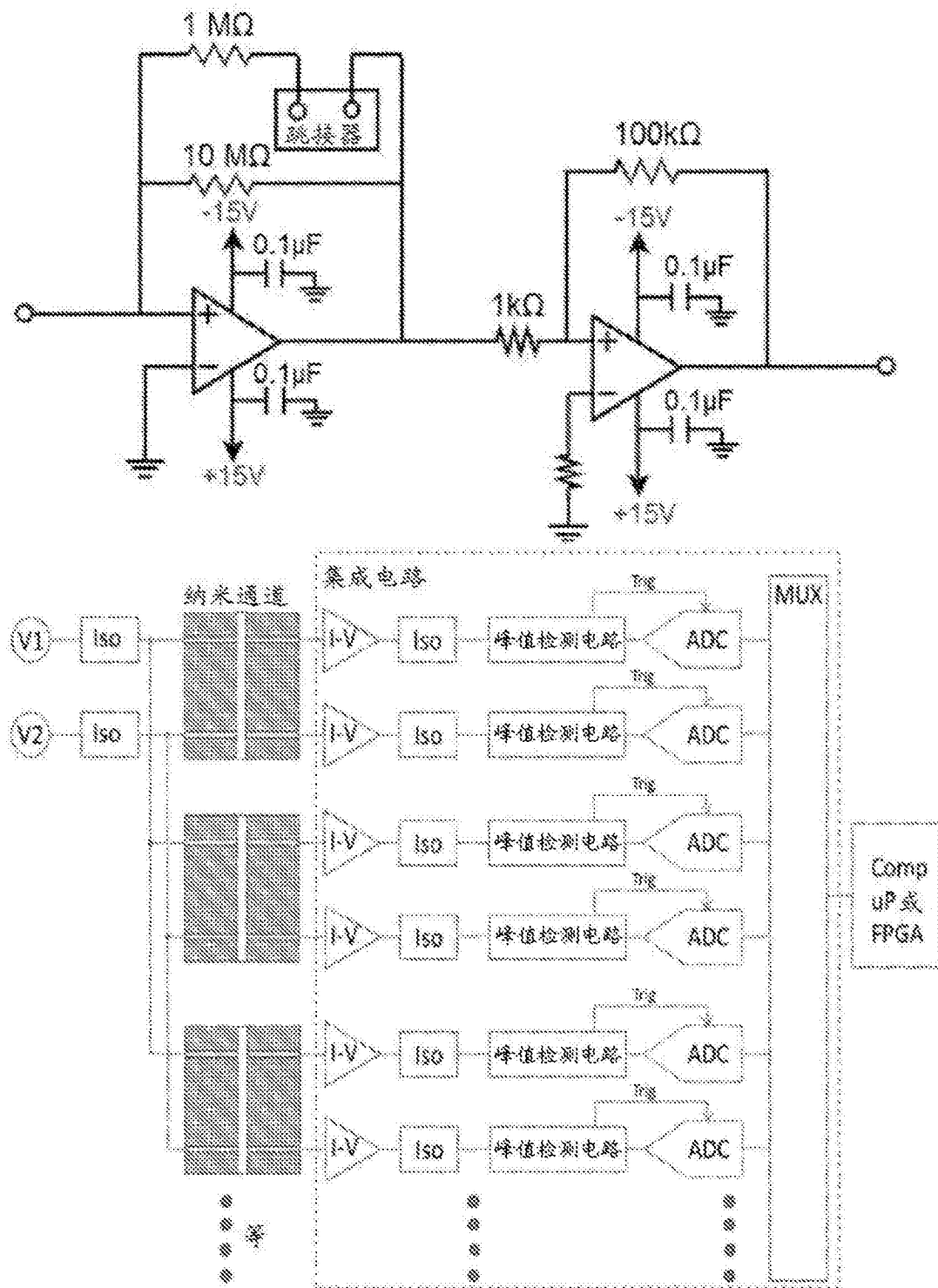


图32

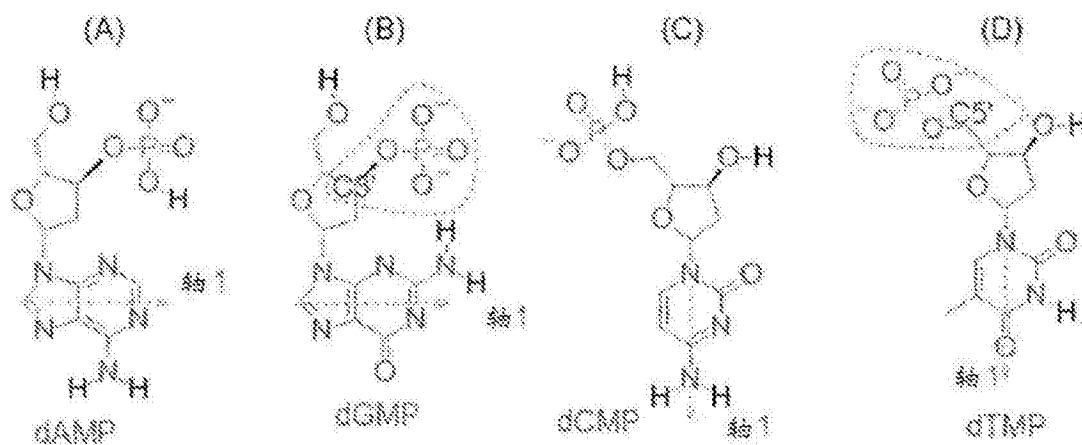


图33A-33D

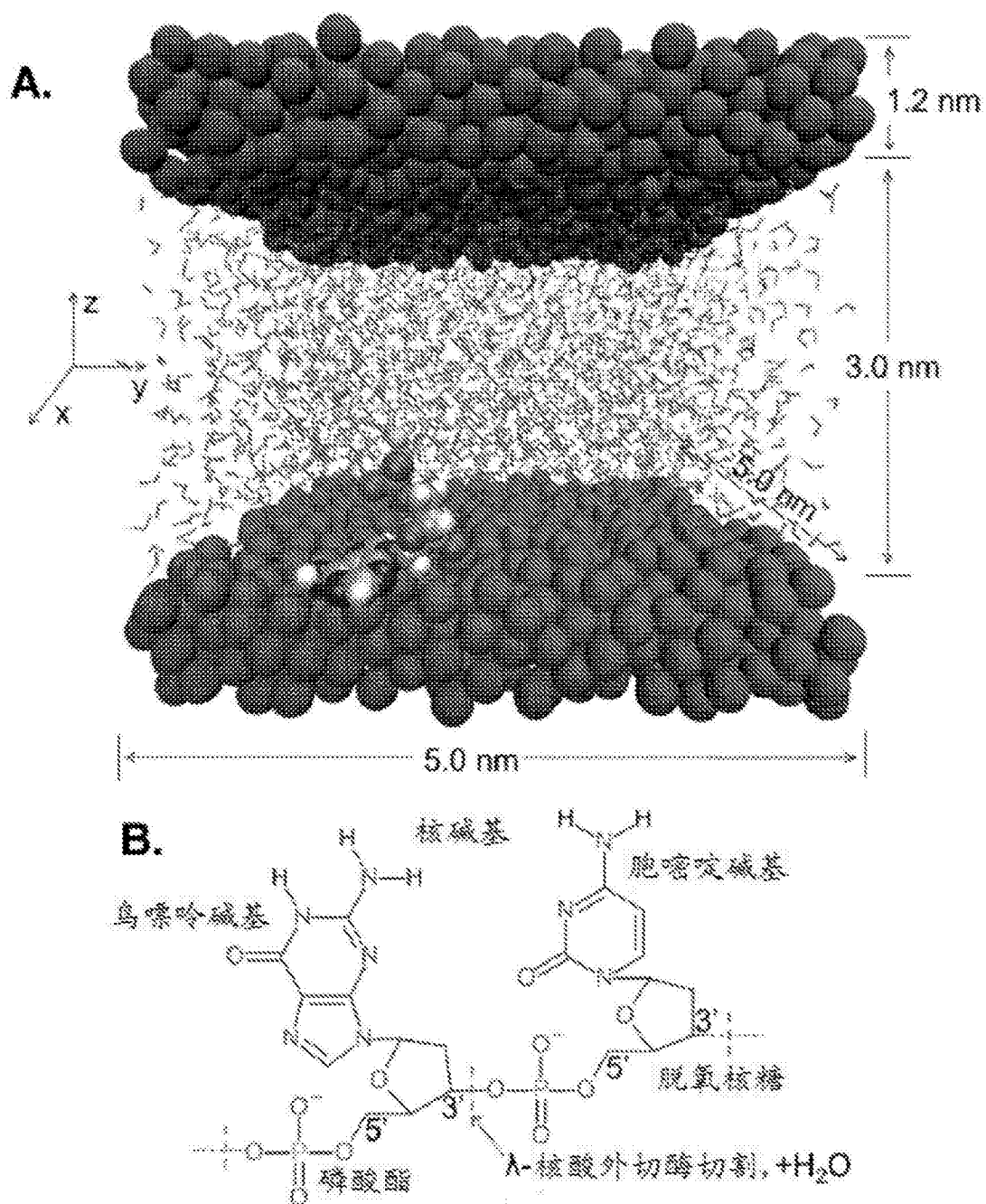


图34A-34B

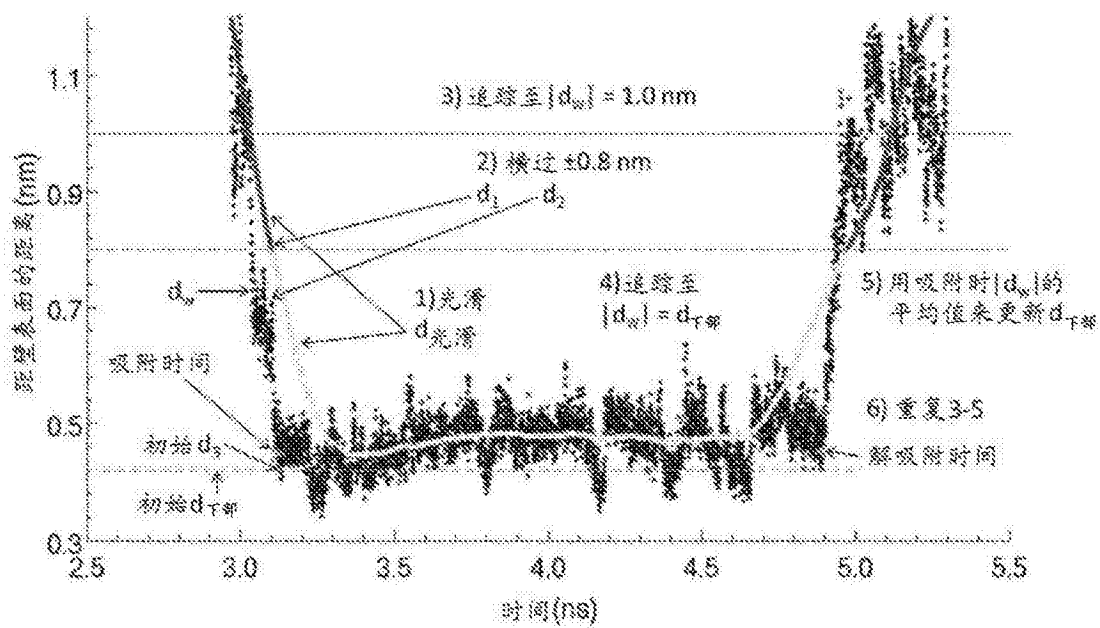


图35

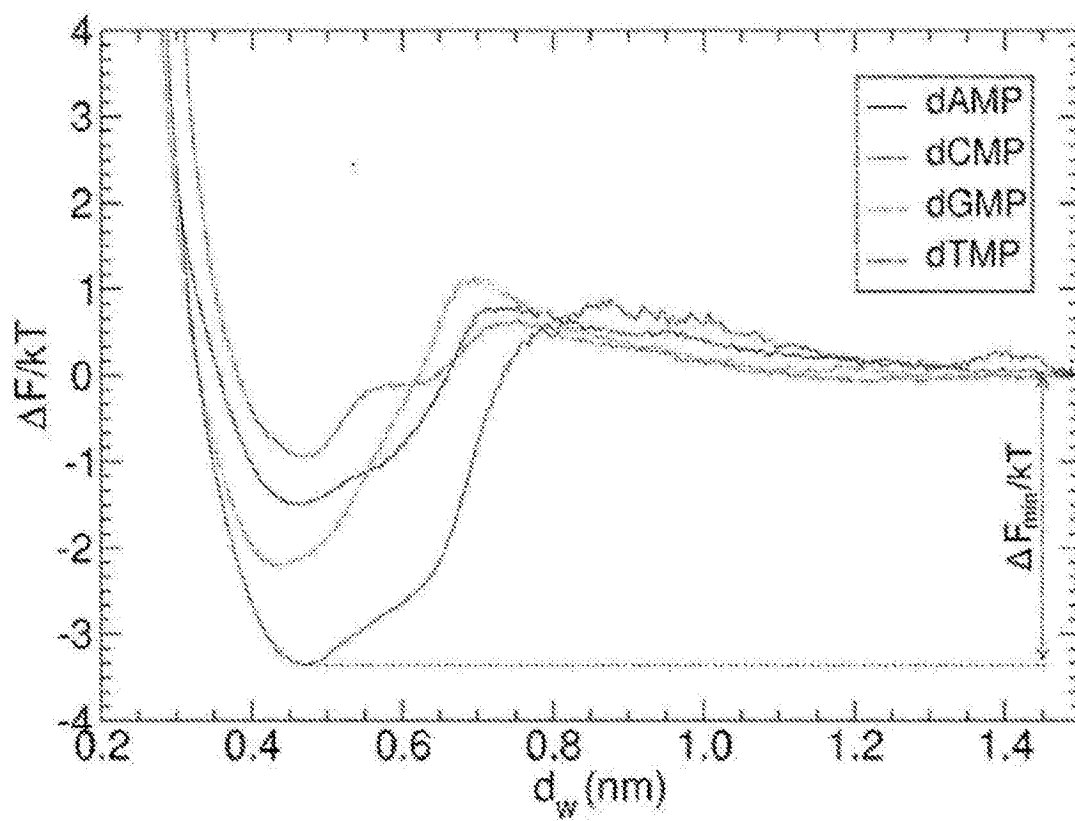


图36

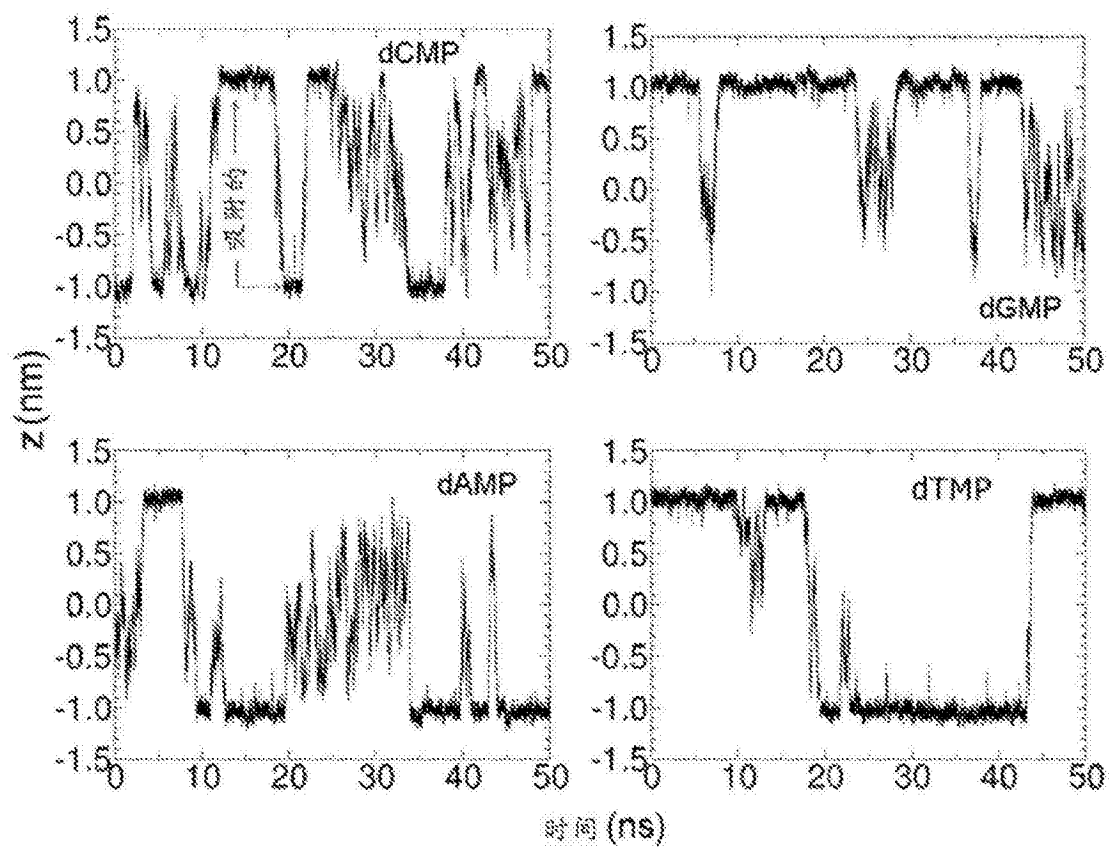


图37

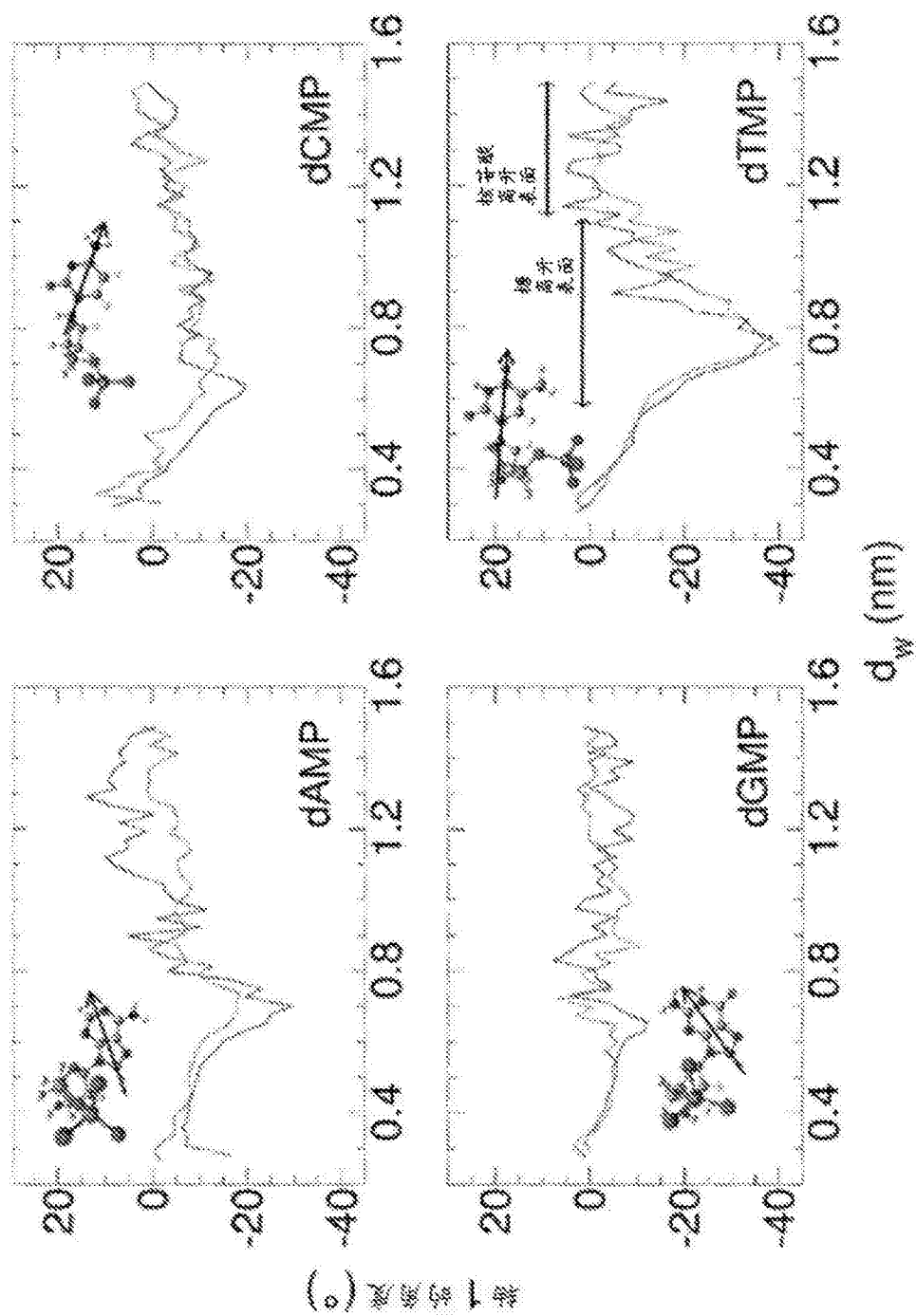


图38

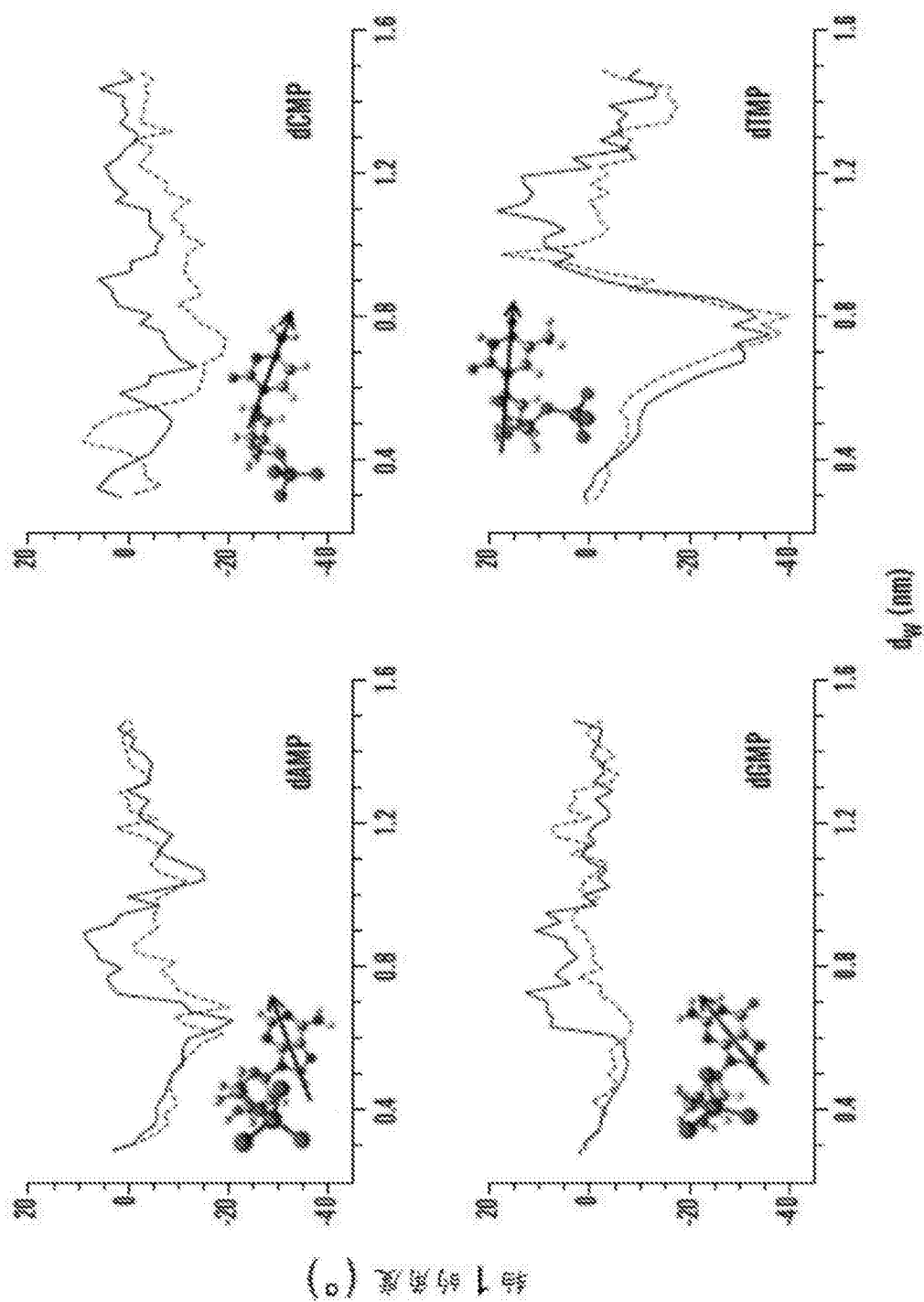


图39

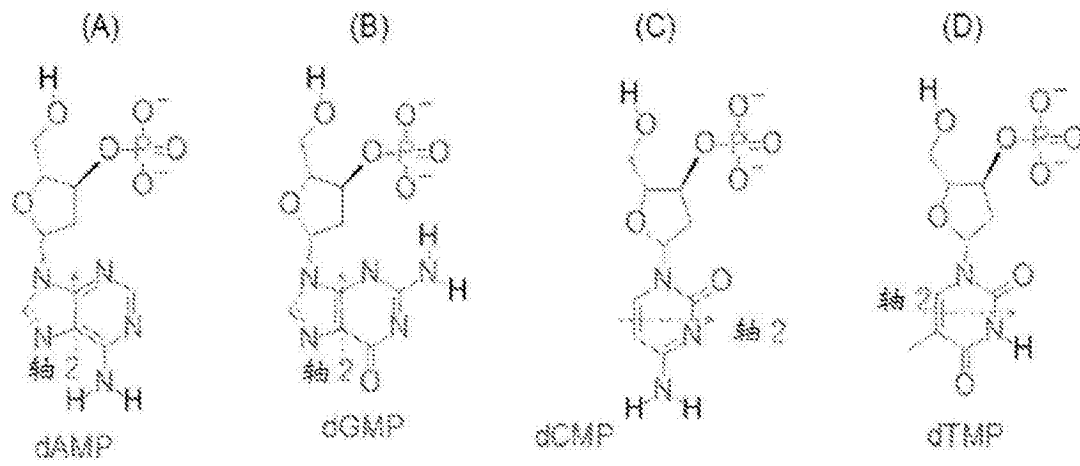


图40A-40D

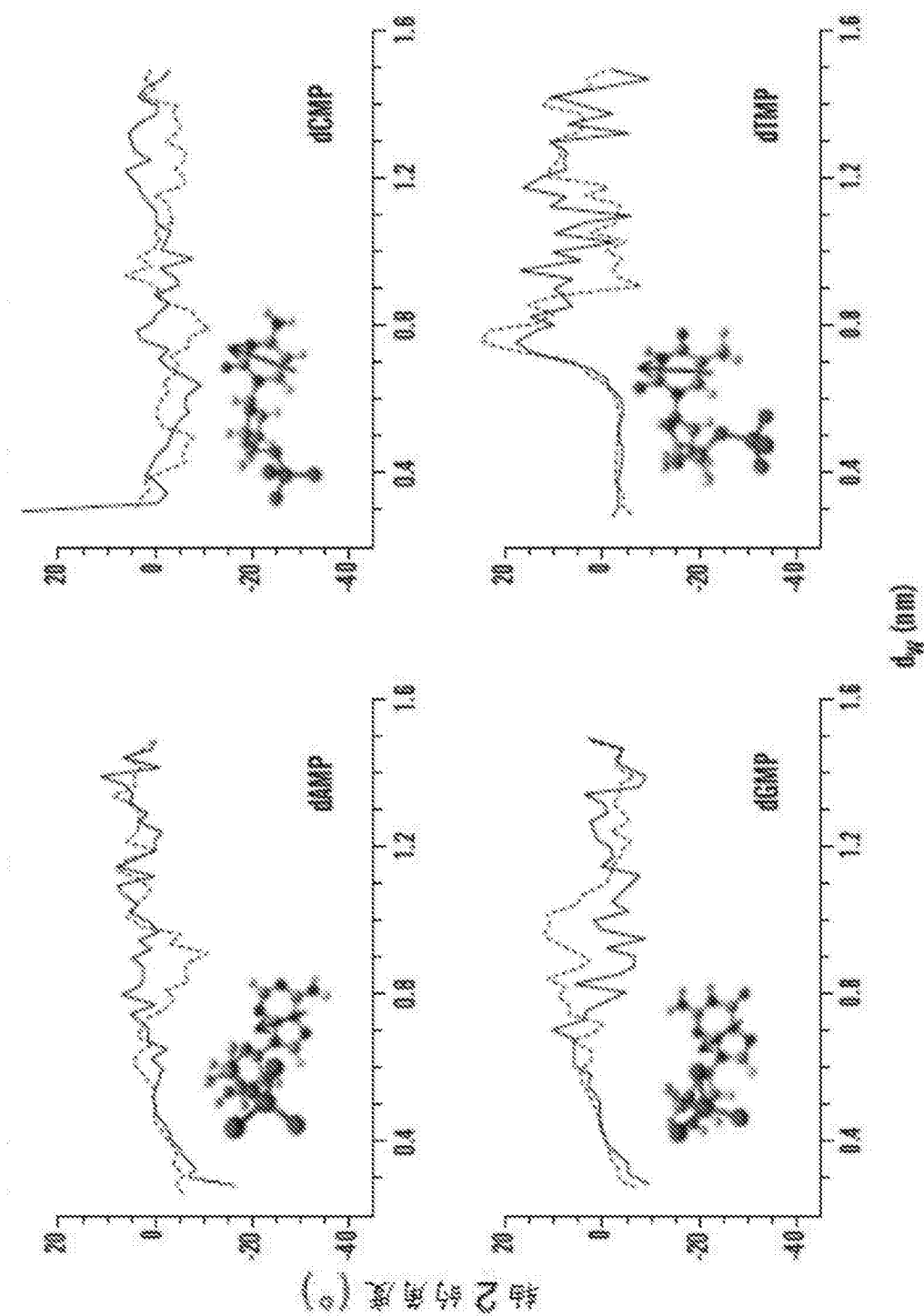


图41

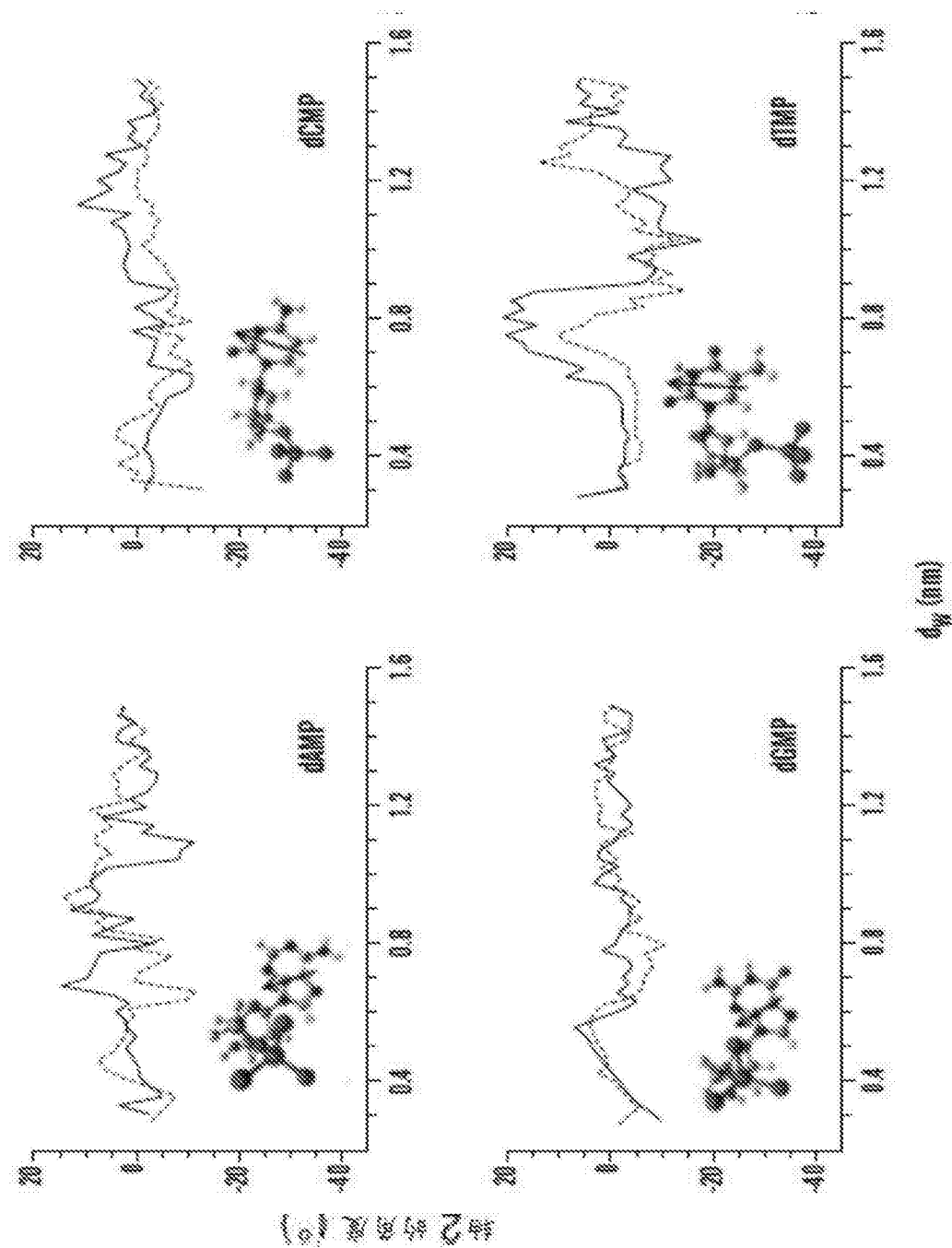


图42

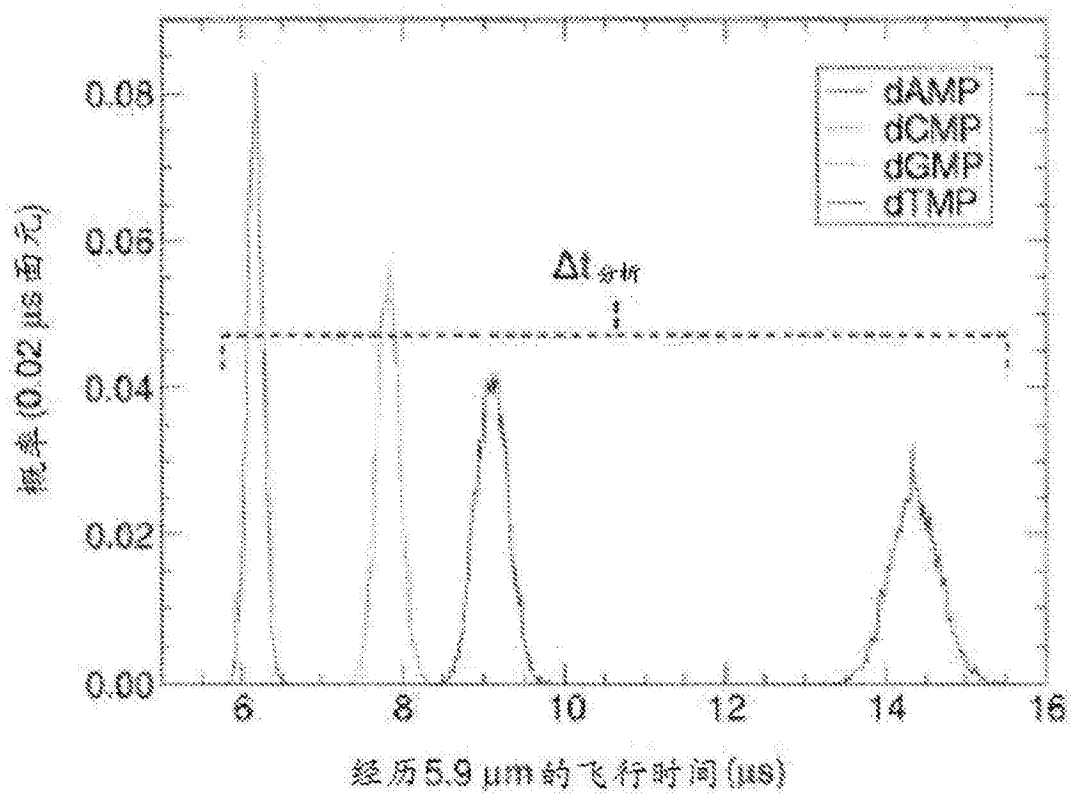


图43

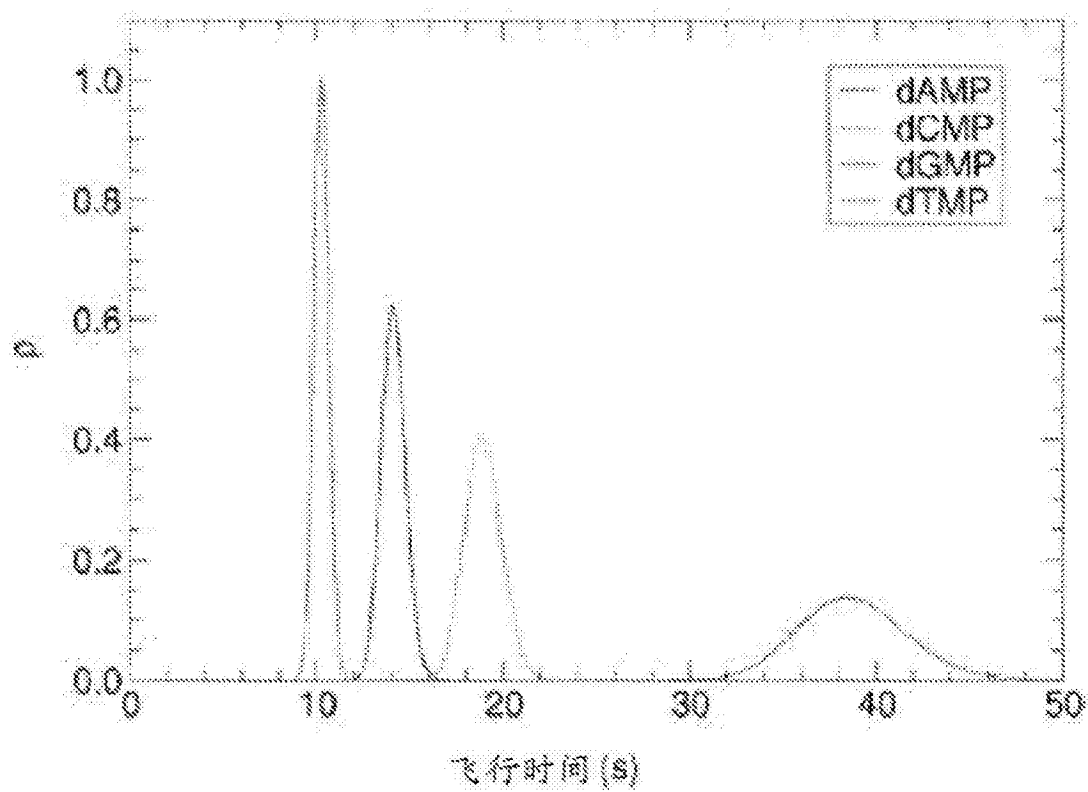


图44

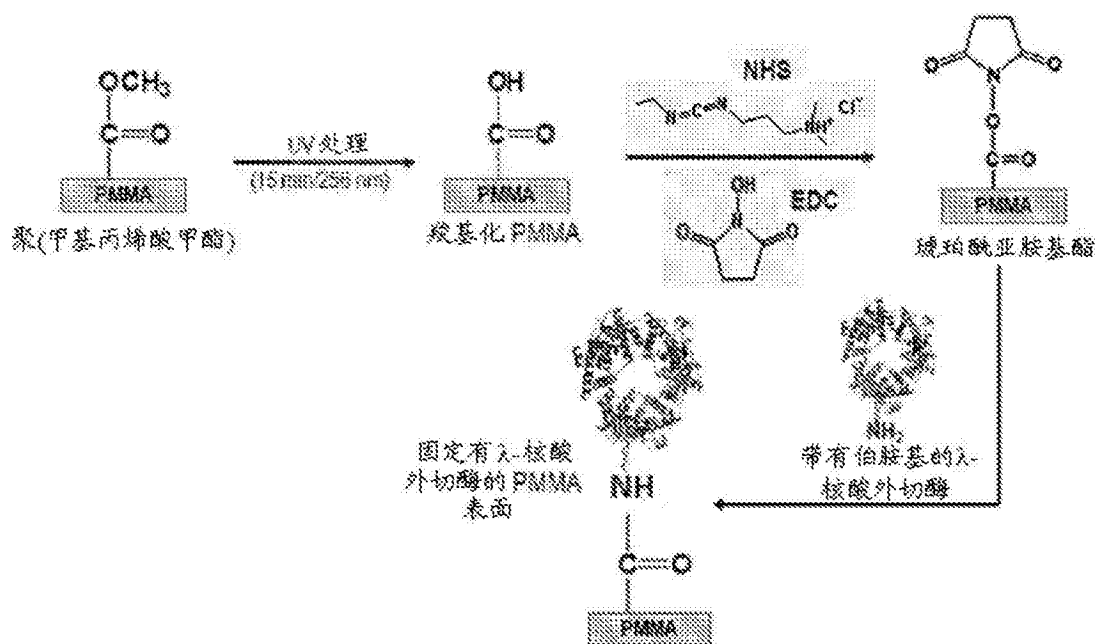


图45

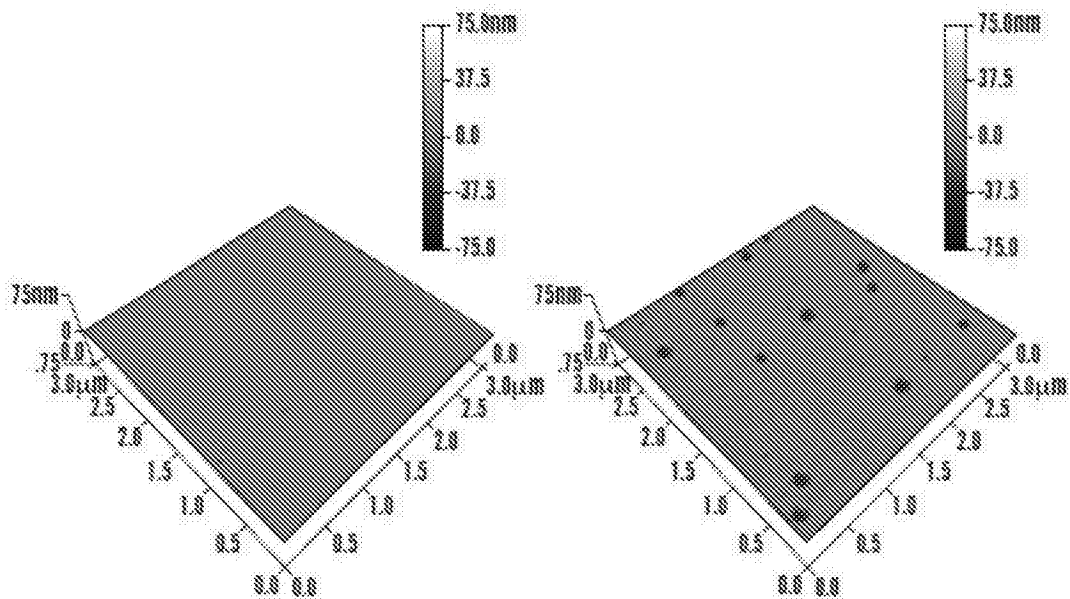


图 46A

图 46B

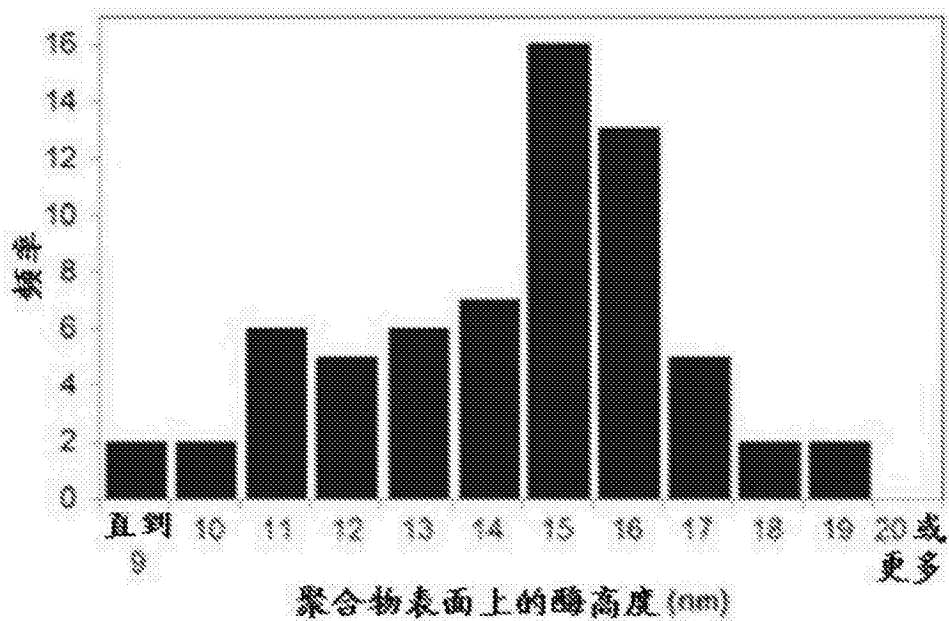


图46C

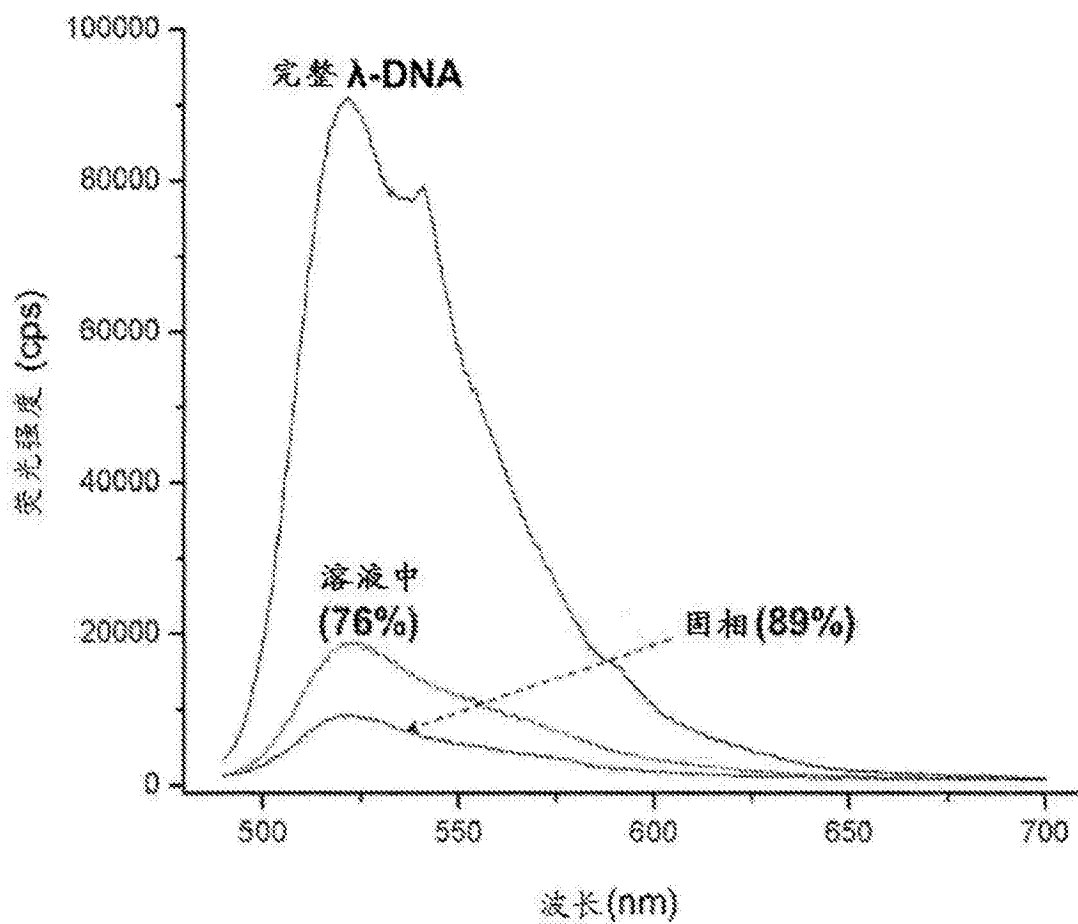


图47

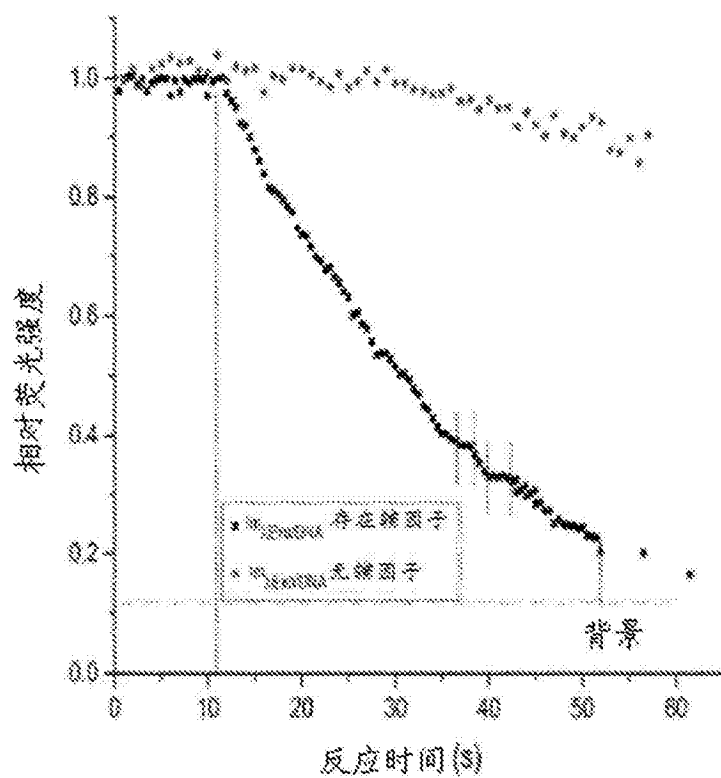


图48

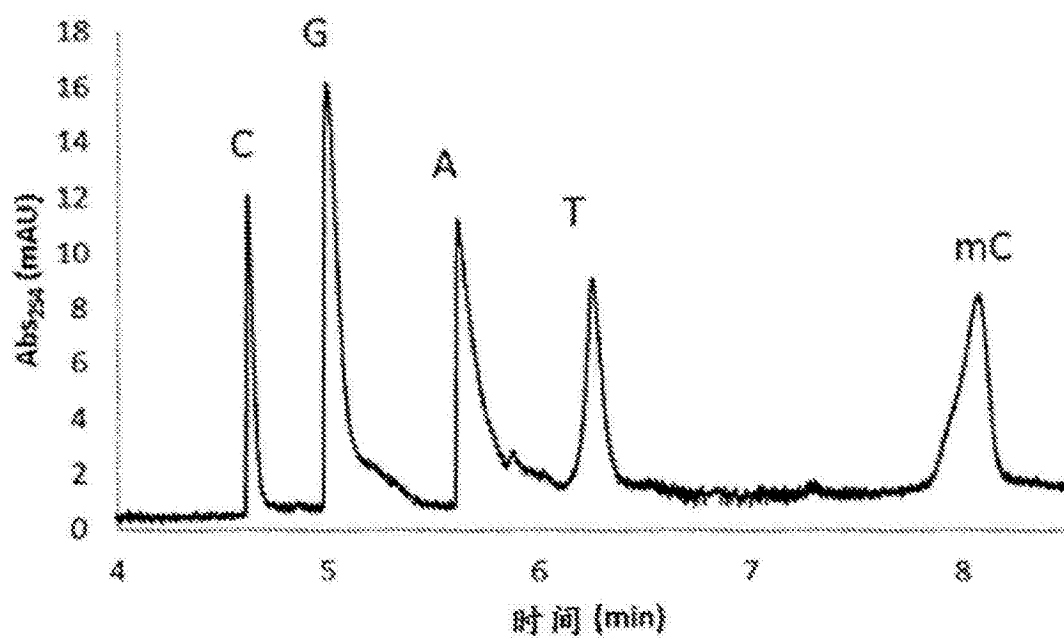


图49

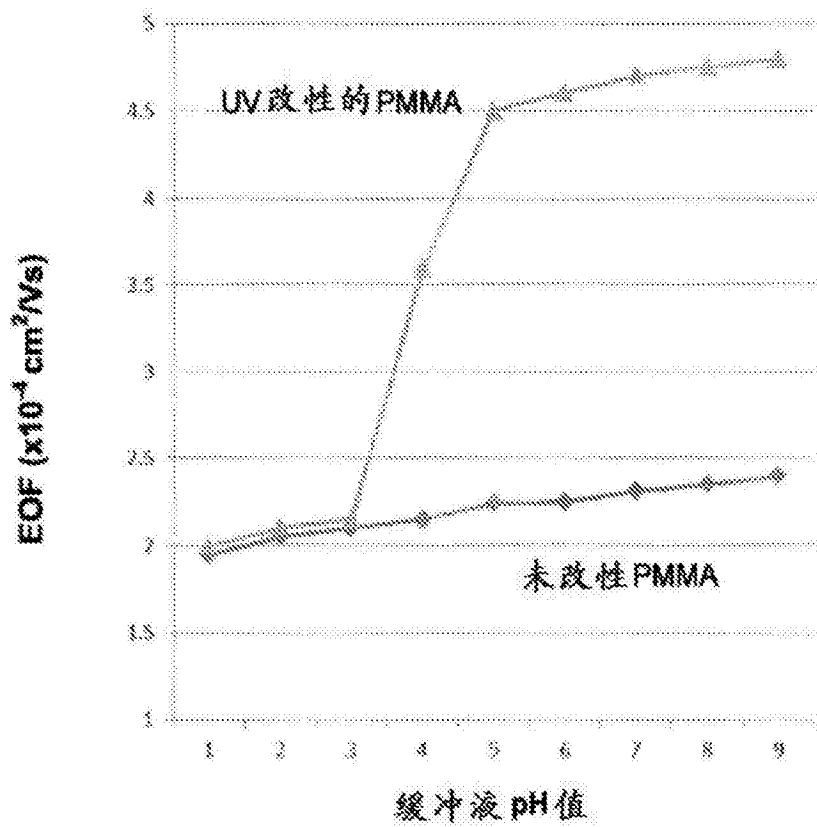


图50A

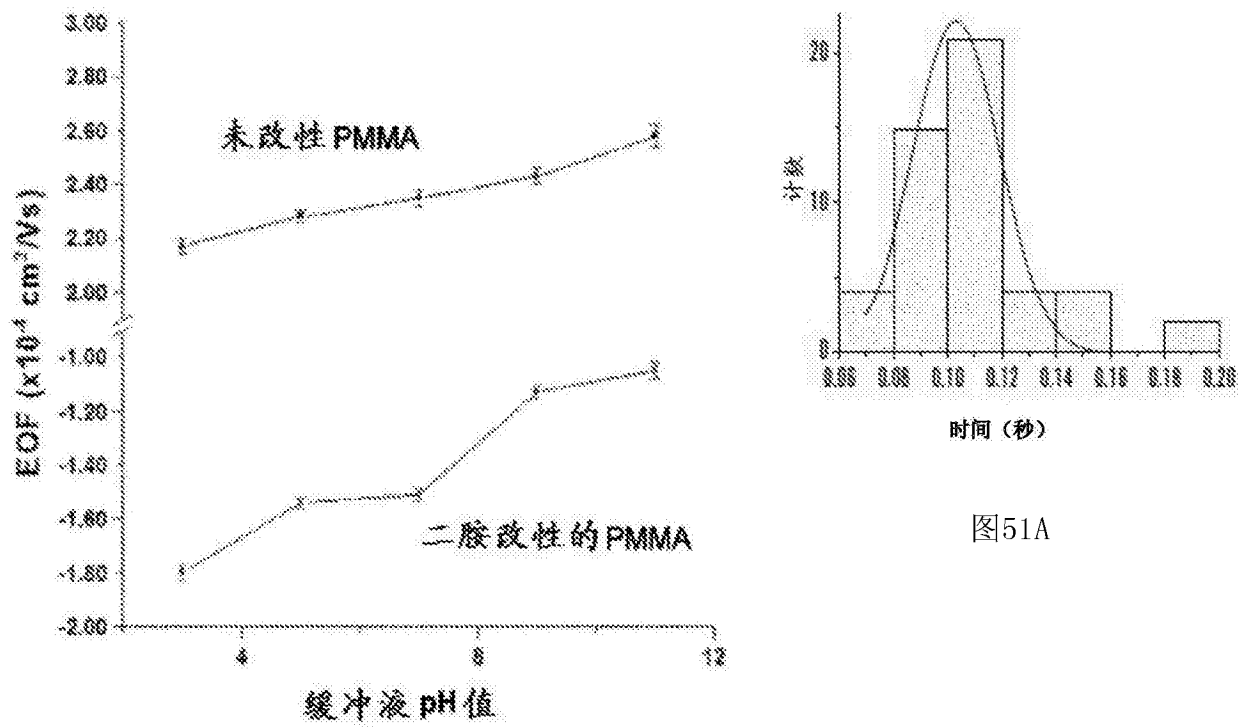


图51A

图50B

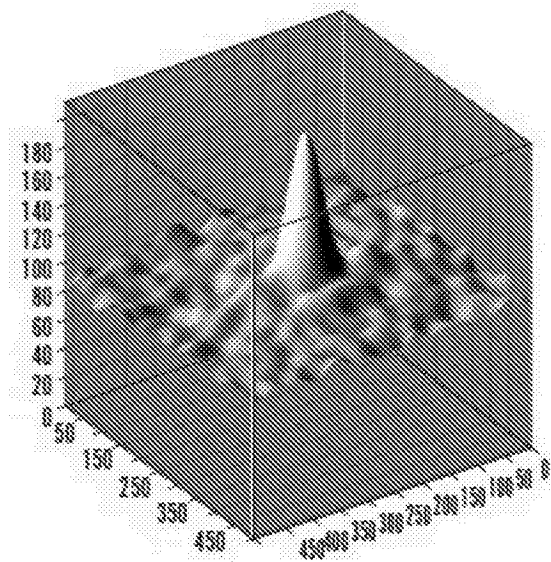


图51B

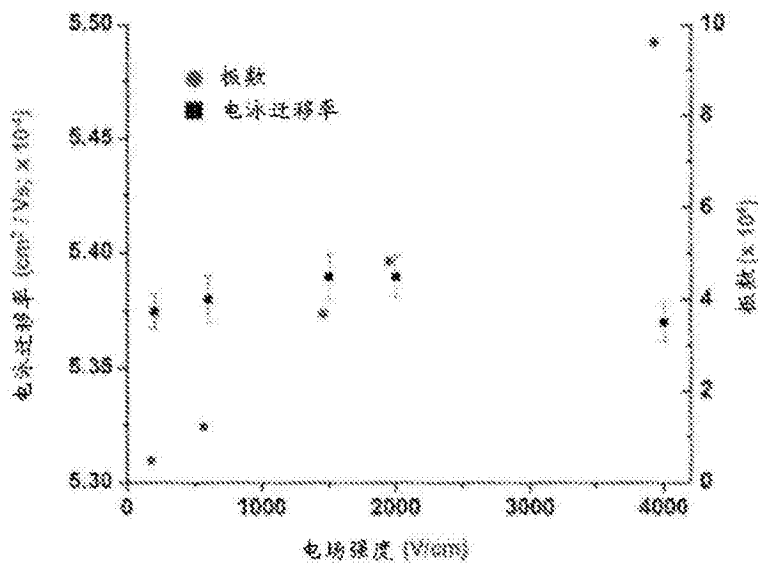


图51C

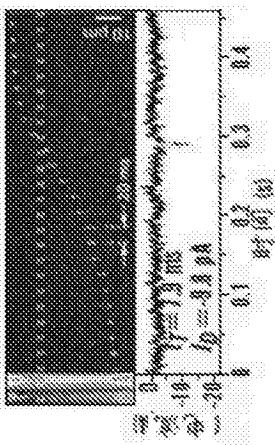


图52A

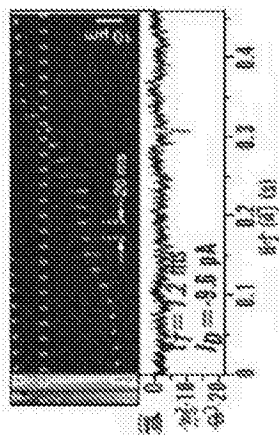


图52B

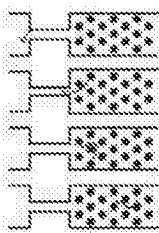


图52C

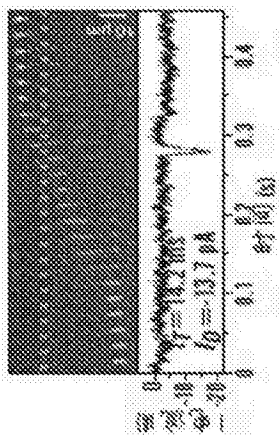


图52D

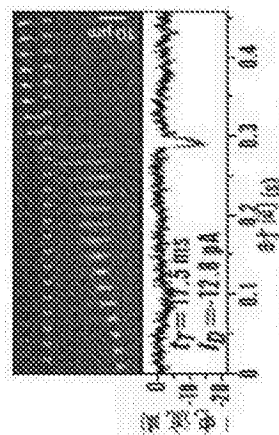


图52E

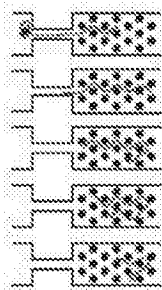


图52F

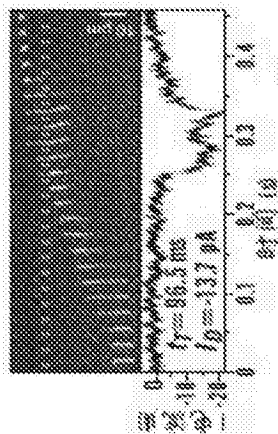


图52G

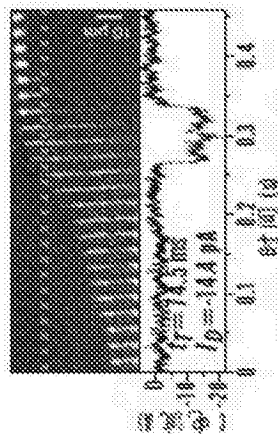


图52H

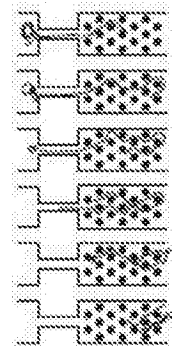


图52I

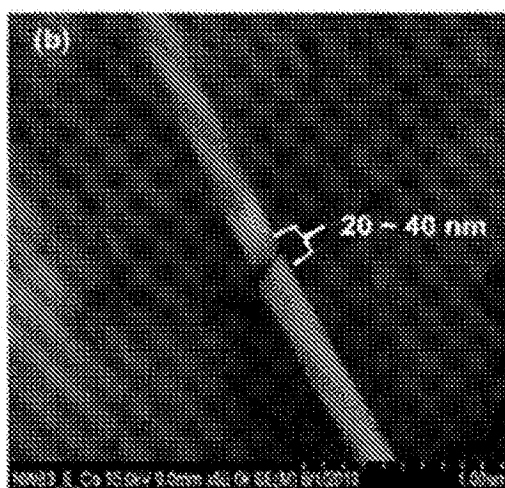
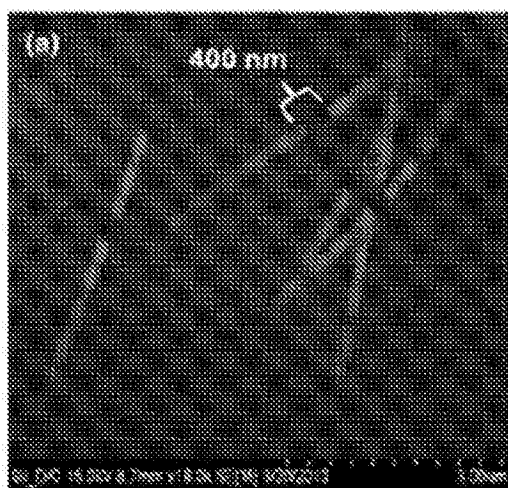


图53A-53B