



(12) 发明专利申请

(10) 申请公布号 CN 114929897 A

(43) 申请公布日 2022. 08. 19

(21) 申请号 202080091893.9

(22) 申请日 2020.10.28

(30) 优先权数据

62/926,897 2019.10.28 US

(85) PCT国际申请进入国家阶段日

2022.06.28

(86) PCT国际申请的申请数据

PCT/US2020/057638 2020.10.28

(87) PCT国际申请的公布数据

W02021/086901 EN 2021.05.06

(71) 申请人 宽腾矽公司

地址 美国康涅狄格州

(72) 发明人 布莱恩·瑞德 马修·戴尔

(74) 专利代理机构 北京品源专利代理有限公司

11332

专利代理师 刘明海 宁涛

(51) Int.Cl.

C12Q 1/6869 (2006.01)

C12Q 1/6874 (2006.01)

C12N 15/10 (2006.01)

C12Q 1/6804 (2006.01)

G01N 33/543 (2006.01)

G01N 33/68 (2006.01)

权利要求书5页 说明书69页

序列表48页 附图6页

(54) 发明名称

制备用于多肽测序的富集样品的方法

(57) 摘要

本申请的方面提供了制备用于多肽测序的富集样品的方法,以及用于该方法的组合物、试剂盒和装置。



1. 一种方法,所述方法包括:

(i) 使用多个富集分子从多个多肽中选择多肽子集,从而产生包含所述多肽子集的富集样品;和

(ii) 对所述富集样品中的多肽进行平行测序。

2. 一种方法,所述方法包括:

(i) 使多个多肽与多个富集分子接触以产生包含多个多肽中的多肽子集的富集样品;和

(ii) 对所述富集样品的多肽进行平行测序。

3. 根据权利要求1或2所述的方法,其中(i)包括:

(a) 使多个多肽与多个富集分子接触,其中所述多个富集分子中的至少富集分子的子集与所述多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集;和

(b) 分离结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。

4. 根据权利要求1或2所述的方法,其中(i)包括:

(a) 使多个多肽与多个富集分子接触,其中所述多个富集分子中的至少富集分子的子集与所述多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集;和

(b) 分离未结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。

5. 根据权利要求1-4中任一项所述的方法,其中所述多个富集分子中的每个富集分子包括抗体、适体或酶。

6. 根据权利要求1-4中任一项所述的方法,其中所述多个富集分子的子集中的富集分子包括抗体、适体或酶。

7. 根据权利要求1-6中任一项所述的方法,其中所述多个富集分子中的每个富集分子被固定在基质上。

8. 根据权利要求1-6中任一项所述的方法,其中所述多个富集分子的子集中的富集分子被固定在基质上。

9. 根据权利要求7或8所述的方法,其中当包含所述多个多肽的样品接触所述基质时,发生(i)中的所述多个多肽与所述多个富集分子的接触。

10. 根据权利要求6-9中任一项所述的方法,其中所述基质选自由表面、珠粒、颗粒和凝胶组成的组,任选地其中:

所述表面是固体表面;

所述珠粒是磁珠;或

所述颗粒是磁性颗粒。

11. 根据权利要求1-10中任一项所述的方法,其中所述多个富集分子中的每个富集分子与包含不同氨基酸序列的两个或更多个多肽结合。

12. 根据权利要求1-10中任一项所述的方法,其中所述多个富集分子的子集中的富集分子与包含不同氨基酸序列的两个或更多个多肽结合。

13. 根据权利要求1-12中任一项所述的方法,其中所述多个富集分子中的每个富集分子与氨基酸翻译后修饰结合。

14. 根据权利要求1-12中任一项所述的方法,其中所述多个富集分子的子集中的富集分子与氨基酸翻译后修饰结合。

15. 根据权利要求13或14所述的方法, 其中所述翻译后修饰选自乙酰化、ADP-核糖基化、半胱天冬酶裂解、瓜氨酸化、甲酰化、羟基化、甲基化、肉豆蔻酰化、N-连接糖基化、类泛素化、硝化、O-连接糖基化、氧化、棕榈酰化、磷酸化、异戊二烯化、S-亚硝基化、硫酸化、苏素化、泛素化组成的组。

16. 根据权利要求13-15中任一项所述的方法, 其中所述多个富集分子的第一子集与第一翻译后修饰结合, 并且所述多个富集分子的第二富集分子子集与第二翻译后修饰结合。

17. 根据权利要求1-16中任一项所述的方法, 其中 (i) 包括:

(a) 使多个多肽与第一多个富集分子接触, 其中第一多个富集分子中的至少富集分子子集与多个多肽中的多肽子集结合, 从而产生第一结合的多肽子集和第一未结合的多肽子集;

(b) 分离 (a) 的第一结合的多肽子集或第一未结合的多肽子集; 和

(c) 用一个或多个另外的多个富集分子反复重复步骤 (a) 和 (b) 以产生包含多种多肽中的多肽子集的富集样品。

18. 根据权利要求17所述的方法, 其中第一多个中的富集分子与翻译后修饰结合, 并且第二多个中的富集分子与一个或多个目的多肽结合。

19. 根据权利要求17所述的方法, 其中第一多个中的富集分子与一个或多个目的多肽结合, 并且第二多个中的富集分子与翻译后修饰结合。

20. 根据权利要求1-19中任一项所述的方法, 其中通过在 (i) 中的多个多肽与多个富集分子接触之前、同时或之后使多肽与修饰剂接触, 在体外修饰多个多肽中的一种或多种多肽。

21. 根据权利要求20所述的方法, 其中所述修饰剂包括变性剂并且至少一种多肽通过变性而被修饰。

22. 根据权利要求20或21所述的方法, 其中所述修饰剂封闭游离羧酸根基团并且至少一种多肽通过封闭多肽的游离羧酸根基团而被修饰。

23. 根据权利要求20-22中任一项所述的方法, 其中所述修饰剂封闭游离硫醇基团并且至少一种多肽通过封闭多肽的游离硫醇基团而被修饰。

24. 根据权利要求20-23中任一项所述的方法, 其中所述修饰剂包括裂解剂并且至少一种多肽通过裂解而被修饰。

25. 根据权利要求20-24中任一项所述的方法, 其中至少一种多肽通过添加翻译后修饰而被修饰。

26. 根据权利要求25所述的方法, 其中所述翻译后修饰选自乙酰化、ADP-核糖基化、半胱天冬酶裂解、瓜氨酸化、甲酰化、羟基化、甲基化、肉豆蔻酰化、N-连接糖基化、类泛素化、硝化、O-连接糖基化、氧化、棕榈酰化、磷酸化、异戊二烯化、S-亚硝基化、硫酸化、苏素化、泛素化组成的组。

27. 根据权利要求1-26中任一项所述的方法, 其中 (ii) 包括:

(a) 使所述富集样品的单个多肽分子与一或多种末端氨基酸识别分子接触; 和

(b) 检测指示所述一种或多种末端氨基酸识别分子与在单个多肽被降解时暴露在单个多肽末端的连续氨基酸结合的一系列信号脉冲, 从而对所述单个多肽分子进行测序。

28. 根据权利要求1-26中任一项所述的方法, 其中 (ii) 包括:

(a) 使所述富集样品的单个多肽分子与包含一种或多种末端氨基酸识别分子和裂解试剂的组合物接触;和

(b) 在所述裂解试剂存在下检测指示所述一种或多种末端氨基酸识别分子与单个多肽分子末端结合的一系列信号脉冲,其中所述一系列信号脉冲指示由于末端氨基酸被裂解试剂裂解而随时间暴露在末端的一系列氨基酸。

29. 根据权利要求1-26中任一项所述的方法,其中(ii)包括:

(a) 鉴定所述富集样品的单个多肽分子末端的第一个氨基酸;

(b) 去除所述第一个氨基酸以暴露单个多肽分子末端的第二个氨基酸,和

(c) 鉴定单个多肽分子末端的所述第二个氨基酸,

其中(a) - (c)在单一反应混合物中进行。

30. 根据权利要求1-26中任一项所述的方法,其中(ii)包括:

(a) 使所述富集样品的单个多肽分子与一种或多种与所述单个多肽分子结合的氨基酸识别分子接触;

(b) 在多肽降解条件下检测指示所述一种或多种氨基酸识别分子与所述单个多肽分子结合的一系列信号脉冲;和

(c) 基于所述一系列信号脉冲中的第一特征模式鉴定所述单个多肽分子中的第一类型氨基酸。

31. 根据权利要求1-26中任一项所述的方法,其中(ii)包括:

(a) 在多肽降解过程中获得数据;

(b) 分析所述数据以确定对应于在降解过程中在多肽末端依次暴露的氨基酸的数据部分;和

(c) 输出代表所述多肽的氨基酸序列。

32. 根据权利要求1-26中任一项所述的方法,其中(ii)包括:

(a) 使所述富集样品的多肽与一种或多种标记的亲试剂接触,所述亲试剂选择性结合多肽末端的一种或多种类型的末端氨基酸;和

(b) 通过检测所述多肽与所述一种或多种标记的亲试剂的相互作用来鉴定所述多肽末端的末端氨基酸。

33. 根据权利要求1-26中任一项所述的方法,其中(ii)包括:

(a) 使所述富集样品中的多肽与一种或多种标记的亲试剂接触,所述亲试剂选择性结合多肽末端的一种或多种类型的末端氨基酸;

(b) 通过检测所述多肽与所述一种或多种标记的亲试剂的相互作用来鉴定所述多肽末端的末端氨基酸;

(c) 去除所述末端氨基酸;和

(d) 在多肽末端重复(a) - (c)一次或多次以确定所述多肽的氨基酸序列。

34. 根据权利要求33所述的方法,其中所述方法进一步包括:

在(a)之后和(b)之前,去除未选择性结合所述末端氨基酸的一种或多种标记的亲试剂中的任意一种;和/或

在(b)之后和(c)之前,去除选择性结合所述末端氨基酸的一种或多种标记的亲试剂中的任意一种。

35. 根据权利要求33所述的方法, 其中 (c) 包括通过使所述末端氨基酸与异硫氰酸酯接触来修饰所述末端氨基酸, 并且:

使经修饰的末端氨基酸与特异性结合并去除经修饰的末端氨基酸的蛋白酶接触; 或
使经修饰的末端氨基酸经受足以去除经修饰的末端氨基酸的酸性或碱性条件。

36. 根据权利要求33所述的方法, 其中鉴定所述末端氨基酸包括:

将所述末端氨基酸鉴定为与所述一种或多种标记的亲试剂结合的一种或多种类型的末端氨基酸中的一种类型; 或

将所述末端氨基酸鉴定为与所述一种或多种标记的亲试剂结合的一种或多种类型的末端氨基酸之外的类型。

37. 根据权利要求33所述的方法, 其中所述一种或多种标记的亲试剂包括一种或多种标记的适体、一种或多种标记的肽酶、一种或多种标记的抗体、一种或多种标记的降解途径蛋白、一种或多种氨基转移酶、一种或多种tRNA合成酶或其组合。

38. 根据权利要求37所述的方法, 其中所述一种或多种标记的肽酶已被修饰以使裂解活性失活; 或其中所述一种或多种标记的肽酶保留用于去除 (c) 的裂解活性。

39. 一种用于实施根据权利要求1-38中任一项所述的方法的试剂盒, 其中所述试剂盒包含多个富集分子。

40. 根据权利要求39所述的试剂盒, 其中所述多个富集分子中的每个富集分子包括抗体、适体或酶。

41. 根据权利要求39所述的试剂盒, 其中所述多个富集分子的子集中的富集分子包括抗体、适体或酶。

42. 根据权利要求39-41中任一项所述的试剂盒, 还包含修饰剂。

43. 根据权利要求42所述的试剂盒, 其中所述修饰剂介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。

44. 根据权利要求39-43中任一项所述的试剂盒, 还包含标记的亲试剂。

45. 根据权利要求44所述的试剂盒, 其中所述标记的亲试剂包括一种或多种标记的适体、一种或多种标记的肽酶、一种或多种标记的抗体、一种或多种标记的降解途径蛋白、一种或多种氨基转移酶、一种或多种tRNA合成酶或其组合。

46. 一种装置, 所述装置包括:

至少一个硬件处理器; 和

至少一个存储处理器可执行指令的非暂时性计算机可读存储介质, 所述处理器可执行指令在被所述至少一个硬件处理器执行时使所述至少一个硬件处理器执行根据权利要求1-38中任一项所述的方法。

47. 至少一个存储处理器可执行指令的非暂时性计算机可读存储介质, 所述处理器可执行指令在由至少一个硬件处理器执行时使所述至少一个硬件处理器执行根据权利要求1-38中任一项所述的方法。

48. 一种装置, 所述装置包含被设置为与一个或多个盒接合的样品制备模块, 每个盒包括: (a) 一个或多个储存器或反应容器, 其被设置为接收复杂样品; (b) 一种或多种序列样品制备试剂, 其中所述样品制备试剂包含多个富集分子; 和 (c) 基质, 其包含一种或多种固定化的捕获探针。

49. 根据权利要求48所述的装置, 其中所述多个富集分子中的至少富集分子的子集共价连接到固定化的捕获探针。

50. 根据权利要求48或49所述的装置, 其中至少富集分子的子集共价连接到能够被固定化的捕获探针结合的珠粒或颗粒。

51. 根据权利要求48-50中任一项所述的装置, 其中所述多个富集分子中的每个富集分子包括抗体、适体或酶。

52. 根据权利要求48-50中任一项所述的装置, 其中所述多个富集分子的子集中的富集分子包括抗体、适体或酶。

53. 根据权利要求48-52中任一项所述的装置, 其中所述样品制备试剂包括修饰剂。

54. 根据权利要求53所述的装置, 其中所述修饰剂介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。

55. 根据权利要求48-54中任一项所述的装置, 其中所述装置进一步包含测序模块, 所述测序模块包含像素阵列, 其中每个像素被设置为从所述样品制备模块接收测序样品并且包含: (a) 样品孔; 和 (b) 至少一个光检测器。

56. 根据权利要求55所述的装置, 其中所述测序模块进一步包含储存器或反应容器, 所述储存器或反应容器被设置为将测序试剂递送到每个像素的样品孔中。

57. 根据权利要求56所述的装置, 其中所述测序试剂包括标记的亲合试剂。

58. 根据权利要求57所述的装置, 其中所述标记的亲合试剂包括一种或多种标记的适体、一种或多种标记的肽酶、一种或多种标记的抗体、一种或多种标记的降解途径蛋白、一种或多种氨基转移酶、一种或多种tRNA合成酶或其组合。

制备用于多肽测序的富集样品的方法

[0001] 相关申请

[0002] 本申请根据35U.S.C.§119(e)要求2019年10月28日提交的美国临时申请序列号62/926,897的申请日的权益,其全部内容通过引用的方式并入本文。

背景技术

[0003] 蛋白质组学已成为生物学系统研究中基因组学和转录组学的重要且必要的补充。蛋白质组学已成为生物学系统研究中基因组学和转录组学的重要且必要的补充。通常,复杂样品中只有一小部分多肽是感兴趣的(例如,临床相关的)。然而,迄今为止,利用分子以高度多重的方式富集目的多肽一直受到限制。

发明内容

[0004] 本文提供了制备用于多肽测序的富集样品的方法,其利用富集分子(和富集分子的组合)来增加目的多肽的相对丰度。这些方法可以以高度多重的方式进行。本文还提供了用于该方法的组合物、试剂盒和装置。

[0005] 在一些方面,本公开涉及制备用于多肽测序的富集样品的方法。在一些实施方案中,所述方法包括:(i)使用多个富集分子从多个多肽中选择多肽子集,从而产生包含多肽子集的富集样品;和(ii)对富集样品中的多肽进行平行测序。在一些实施方案中,所述方法包括:(i)使多个多肽与多个富集分子接触以产生包含多个多肽中的多肽子集的富集样品;和(ii)对富集样品的多肽进行平行测序。

[0006] 在一些实施方案中,(i)包括:(a)使多个多肽与多个富集分子接触,其中多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集;和(b)分离结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。

[0007] 在一些实施方案中,(i)包括:(a)使多个多肽与多个富集分子接触,其中多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集;和(b)分离未结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。

[0008] 在一些实施方案中,多个富集分子中的每个富集分子包括抗体、适体或酶。在一些实施方案中,多个富集分子的子集中的富集分子包括抗体、适体或酶。

[0009] 在一些实施方案中,多个富集分子中的每个富集分子固定在基质上。在一些实施方案中,多个富集分子的子集中的富集分子固定在基质上。在一些实施方案中,当包含多个多肽的样品接触基质时,发生(i)中的多个多肽与多个富集分子的接触。在一些实施方案中,所述基质选自自由表面、珠粒、颗粒和凝胶组成的组,任选地其中:所述表面是固体表面;所述珠粒是磁珠;或所述颗粒是磁性颗粒。

[0010] 在一些实施方案中,多个富集分子中的每个富集分子与两个或更多个包含不同氨基酸序列的多肽结合。在一些实施方案中,多个富集分子的子集中的富集分子与两个或更

多个包含不同氨基酸序列的多肽结合。

[0011] 在一些实施方案中,多个富集分子中的每个富集分子与氨基酸翻译后修饰结合。在一些实施方案中,多个富集分子的子集中的富集分子与氨基酸翻译后修饰结合。在一些实施方案中,翻译后修饰选自乙酰化、ADP-核糖基化、半胱天冬酶裂解、瓜氨酸化、甲酰化、羟基化、甲基化、肉豆蔻酰化、N-连接糖基化、类泛素化、硝化、O-连接糖基化、氧化、棕榈酰化、磷酸化、异戊二烯化、S-亚硝基化、硫酸化、苏素化、泛素化组成的组。在一些实施方案中,多个富集分子的第一子集与第一翻译后修饰结合,并且多个富集分子的第二富集分子子集与第二翻译后修饰结合。

[0012] 在一些实施方案中,(i)包括:(a)使多个多肽与第一多个富集分子接触,其中第一多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生第一结合的多肽子集和第一未结合的多肽子集;(b)分离(a)的第一结合的多肽子集或第一未结合的多肽子集;和(c)用一种或多种另外的多个富集分子反复重复步骤(a)和(b)以产生包含多个多肽中的多肽子集的富集样品。在一些实施方案中,第一多个中的富集分子与翻译后修饰结合,第二多个中的富集分子与一个或多个目的多肽结合。在一些实施方案中,第一多个中的富集分子与一个或多个目的多肽结合,第二多个中的富集分子与翻译后修饰结合。

[0013] 在一些实施方案中,通过在(i)中的多个多肽与多个富集分子接触之前、同时或之后使多肽与修饰剂接触,在体外修饰多个多肽中的一种或多种多肽。在一些实施方案中,修饰剂包括变性剂并且至少一种多肽通过变性而被修饰。在一些实施方案中,修饰剂封闭游离羧酸根基团并且至少一种多肽通过封闭多肽的游离羧酸根基团而被修饰。在一些实施方案中,修饰剂封闭游离硫醇基团并且至少一种多肽通过封闭多肽的游离硫醇基团而被修饰。在一些实施方案中,修饰剂包括裂解剂并且至少一种多肽通过裂解而被修饰。在一些实施方案中,至少一种多肽通过添加翻译后修饰而被修饰。在一些实施方案中,翻译后修饰选自乙酰化、ADP-核糖基化、半胱天冬酶裂解、瓜氨酸化、甲酰化、羟基化、甲基化、肉豆蔻酰化、N-连接糖基化、类泛素化、硝化、O-连接糖基化、氧化、棕榈酰化、磷酸化、异戊二烯化、S-亚硝基化、硫酸化、苏素化、泛素化组成的组。

[0014] 在一些实施方案中,(ii)包括:(a)使富集样品的单个多肽分子与一种或多种末端氨基酸识别分子接触;和(b)检测指示一种或多种末端氨基酸识别分子与在单个多肽被降解时暴露在单个多肽末端的连续氨基酸结合的一系列信号脉冲,从而对所述单个多肽分子进行测序。

[0015] 在一些实施方案中,(ii)包括:(a)使富集样品的单个多肽分子与包含一种或多种末端氨基酸识别分子和裂解试剂的组合物接触;和(b)在裂解试剂存在下检测指示一种或多种末端氨基酸识别分子与单个多肽分子末端结合的一系列信号脉冲,其中所述一系列信号脉冲指示因末端氨基酸被裂解试剂裂解而随时间暴露在末端的一系列氨基酸。

[0016] 在一些实施方案中,(ii)包括:(a)鉴定富集样品的单个多肽分子末端的第一个氨基酸;(b)去除第一个氨基酸以暴露单个多肽分子末端的第二个氨基酸,和(c)鉴定单个多肽分子末端的第二个氨基酸,其中(a)-(c)在单一反应混合物中进行。

[0017] 在一些实施方案中,(ii)包括:(a)使富集样品的单个多肽分子与一种或多种与单个多肽分子结合的氨基酸识别分子接触;(b)在多肽降解条件下检测指示一种或多种氨基酸识别分子与单个多肽分子结合的一系列信号脉冲;和(c)基于一系列信号脉冲中的第一

特征模式鉴定单个多肽分子中的第一类型的氨基酸。

[0018] 在一些实施方案中, (ii) 包括: (a) 在多肽降解过程中获得数据; (b) 分析数据以确定对应于在降解过程中在多肽末端依次暴露的氨基酸的数据部分; 和 (c) 输出代表所述多肽的氨基酸序列。

[0019] 在一些实施方案中, (ii) 包括: (a) 使富集样品的多肽与一种或多种标记的亲试剂接触, 所述亲试剂选择性结合多肽末端的一种或多种类型的末端氨基酸; 和 (b) 通过检测多肽与一种或多种标记的亲试剂的相互作用来鉴定多肽末端的末端氨基酸。

[0020] 在一些实施方案中, (ii) 包括: (a) 使富集样品中的多肽与一种或多种标记的亲试剂接触, 所述亲试剂选择性结合多肽末端的一种或多种类型的末端氨基酸; (b) 通过检测多肽与所述一种或多种标记的亲试剂的相互作用来鉴定多肽末端的末端氨基酸; (c) 去除末端氨基酸; 和 (d) 在多肽末端重复 (a) - (c) 一次或多次以确定多肽的氨基酸序列。在一些实施方案中, 所述方法进一步包括: 在 (a) 之后和 (b) 之前, 去除未选择性结合末端氨基酸的一种或多种标记的亲试剂中的任意一种; 和/或在 (b) 之后和 (c) 之前, 去除选择性结合末端氨基酸的一种或多种标记的亲试剂中的任意一种。在一些实施方案中, (c) 包括通过使末端氨基酸与异硫氰酸酯接触来修饰末端氨基酸, 和: 使经修饰的末端氨基酸与特异性结合并去除经修饰的末端氨基酸的蛋白酶接触; 或使经修饰的末端氨基酸经受足以去除经修饰的末端氨基酸的酸性或碱性条件。在一些实施方案中, 鉴定末端氨基酸包括: 将末端氨基酸鉴定为与一种或多种标记的亲试剂结合的一种或多种类型的末端氨基酸中的一种类型; 或将末端氨基酸鉴定为与一种或多种标记的亲试剂结合的一种或多种类型的末端氨基酸之外的类型。

[0021] 在一些实施方案中, 所述一种或多种标记的亲试剂包括一种或多种标记的适体、一种或多种标记的肽酶、一种或多种标记的抗体、一种或多种标记的降解途径蛋白、一种或多种氨基转移酶、一种或多种 tRNA 合成酶或其组合。在一些实施方案中, 所述一种或多种标记的肽酶已被修饰以使裂解活性失活; 或其中所述一种或多种标记的肽酶保留用于去除 (c) 的裂解活性。

[0022] 在一些方面, 本公开涉及在执行本文所述的方法中使用的试剂盒。在一些实施方案中, 所述试剂盒包含多个富集分子。在一些实施方案中, 多个富集分子中的每个富集分子包括抗体、适体或酶。在一些实施方案中, 多个富集分子的子集中的富集分子包括抗体、适体或酶。

[0023] 在一些实施方案中, 所述试剂盒包含修饰剂。在一些实施方案中, 修饰剂介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。

[0024] 在一些实施方案中, 所述试剂盒包含标记的亲试剂。在一些实施方案中, 标记的亲试剂包括一种或多种标记的适体、一种或多种标记的肽酶、一种或多种标记的抗体、一种或多种标记的降解途径蛋白、一种或多种氨基转移酶、一种或多种 tRNA 合成酶或其组合。

[0025] 在一些方面, 本公开涉及一种存储处理器可执行指令的非暂时性计算机可读存储介质, 所述处理器可执行指令在由所述至少一个硬件处理器执行时使所述至少一个硬件处理器执行本文所述的富集方法。

[0026] 在一些方面, 本公开涉及用于执行本文所述的方法的装置。在一些实施方案中, 一种装置包含: 至少一个硬件处理器; 和至少一个存储处理器可执行指令的非暂时性计算机

可读存储介质,所述处理器可执行指令在由所述至少一个硬件处理器执行时使所述至少一个硬件处理器执行富集方法。

[0027] 在一些实施方案中,所述装置包含:(i)样品制备模块,其被设置为与一个或多个盒(cartridge)接合(interface),每个盒包含:(a)一个或多个储存器或反应容器,其被设置为接收复杂样品;(b)一种或多种序列样品制备试剂,其中样品制备试剂包含多个富集分子;和(c)基质,其包含一种或多种固定化的捕获探针;(ii)包含像素(pixel)阵列的测序模块,其中每个像素被设置为从样品制备模块接收测序样品,并且包含:(a)样品孔;和(b)至少一个光检测器。

[0028] 在一些实施方案中,多个富集分子中的至少富集分子的子集被固定在(例如,共价连接至)固定化的捕获探针。在一些实施方案中,至少富集分子的子集被固定在(例如,共价连接至)能够被固定化的捕获探针结合的珠粒或颗粒。

[0029] 在一些实施方案中,多个富集分子中的每个富集分子包括抗体、适体或酶。在一些实施方案中,多个富集分子的子集中的富集分子包括抗体、适体或酶。

[0030] 在一些实施方案中,所述样品制备试剂包括修饰剂。在一些实施方案中,修饰剂介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。

[0031] 在一些实施方案中,所述测序模块进一步包含储存器或反应容器,所述储存器或反应容器被设置为将测序试剂递送到每个像素的样品孔中。

[0032] 在一些实施方案中,所述测序试剂包括标记的亲试剂。在一些实施方案中,标记的亲试剂包括一种或多种标记的适体、一种或多种标记的肽酶、一种或多种标记的抗体、一种或多种标记的降解途径蛋白、一种或多种氨基转移酶、一种或多种tRNA合成酶或其组合。

附图说明

[0033] 本领域技术人员将理解,本文所述的附图仅用于说明目的。应当理解,在一些情况下,本发明的各个方面可能被变大或放大以有助于理解本发明。在附图中,相似的参考符号在各个附图中通常指代相似的特征、功能上相似和/或结构上相似的元件。附图不一定按比例绘制,而是强调说明教导的原理。附图无意以任何方式限制本教导的范围。

[0034] 本发明的特征和优点将在下面结合附图的详细描述中变得更加明显。

[0035] 当参考附图描述实施方案时,可以使用方向参考(“上方”、“下方”、“顶部”、“底部”、“左侧”、“右侧”、“水平”、“垂直”等)。此类参考仅旨在帮助读者以正常方向查看附图。这些方向参考并非旨在描述具体装置的优选或独特方向。装置可以以其他方向体现。

[0036] 如从详细描述中显而易见的,在整个申请中为了说明的目的而在附图中描绘和进一步描述的实例描述了非限制性实施方案,并且在一些情况下可以为了更清楚的说明的目的而简化某些过程或省略特征或步骤。

[0037] 图1提供了复杂样品的示例性说明。复杂样品由许多不同的多肽组成,其中只有一些可能是目的多肽。A(正方形)、B(矩形)和C(圆形)代表不同的多肽。星号突出显示包含修饰(例如,翻译后修饰)的多肽。

[0038] 图2提供了富集分子的示例性说明。富集分子可以与一种或多种多肽结合(或被其结合)。例如,富集分子1与多肽C(圆形)结合(或被其结合)。相反,富集分子2与多肽A(正方

形)和B(矩形)结合(或被其结合),而富集分子3与多肽修饰(这里在多肽A、B和C上发现)结合(或被其结合)。此外,富集分子可以组合使用。

[0039] 图3提供了富集的示例性实施方案。将富集分子添加到复杂混合物中,所述复杂混合物与靶多肽结合(或被其结合)。然后分离未结合的多肽(这里为目的多肽),其相对于靶多肽被富集。以这种方式,目的多肽的相对丰度增加。

[0040] 图4提供了富集的示例性实施方案。将富集分子添加到复杂混合物中,所述复杂混合物与靶多肽(这里为目的多肽)结合(或被其结合)。然后分离结合的多肽,其相对于非靶多肽被富集。以这种方式,目的多肽的相对丰度增加。

[0041] 图5提供了描绘制备富集样品的示例性工作流程的图示。

[0042] 图6提供了描绘制备富集样品的示例性工作流程的图示。

[0043] 图7提供了描绘制备富集样品的示例性工作流程的图示。

[0044] 图8提供了描绘用于执行复杂样品的富集的示例性装置的图示。

具体实施方式

[0045] 如本文所述,发明人已经认识到并理解不同的结合相互作用可以为多肽测序中的常规标记策略提供另外的或替代的方法。常规的多肽测序可以涉及用独特可识别的标记来标记每种类型的氨基酸。这个过程可能费力并且容易出错,因为有至少二十种不同类型的天然存在的氨基酸,以及其多种翻译后变体。在一些方面,本公开涉及使用氨基酸识别分子的技术的发现,所述氨基酸识别分子与不同类型的氨基酸有区别地结合以产生指示多肽的氨基酸序列的可检测特征。

[0046] 在一些方面,本公开涉及可以仅使用单一反应混合物(例如,不需要通过反应容器的反复试剂循环)来实时监测多肽测序反应的发现。常规的多肽测序反应可以涉及将多肽暴露于不同的试剂混合物以在氨基酸检测和氨基酸裂解步骤之间循环。因此,在一些方面,本公开涉及下一代测序的进步,其允许通过氨基酸检测在整个进行中的降解反应中实时分析多肽。

[0047] 对单个生物体的蛋白质组学分析可以提供对细胞过程和反应模式的洞察,从而改进诊断和治疗策略。然而,复杂样品对蛋白质组学分析提出了各种挑战。这里特别相关的是,复杂样品包含大量不同的多肽,并且通常只有一小部分多肽是目的多肽(例如,临床相关的)。例如,在来自血液的复杂样品中,绝大多数蛋白质含量(例如,99%)由“管家蛋白质(housing-keeping protein)”组成,例如白蛋白。因此,通过复杂样品的蛋白质组学分析产生的大部分数据没有什么价值。如果不将生成的数据的内容集中目的方面,重要的洞察可能会被低估和/或完全错过。如本文所述,发明人已经认识到并理解富集目的多肽的能力将提高复杂样品的蛋白质组学分析的效率。

[0048] 因此,在一些方面,本公开涉及制备用于多肽测序(例如,通过本文公开的方法)的富集样品的方法,其利用富集分子(和富集分子的组合)来增加目的多肽的相对丰度。这些方法可以以高度多重的方式进行,从而提高复杂样品的蛋白质组学分析的效率并降低与其相关的成本。本文还提供了对其有用的组合物、试剂盒和装置。

[0049] I. 制备复杂多肽样品的方法

[0050] 在一些方面,本公开涉及制备复杂样品(例如,复杂多肽样品)的方法。如本文所

用,术语“复杂样品”是指包含多种分子(例如,多肽、多核酸、代谢物等)的样品,所述分子的至少两种是化学上独特的。在一些实施方案中,复杂样品包含多个多肽,其中所述多个多肽包含至少两个包含不同氨基酸序列的多肽。

[0051] 通常,复杂样品来源于细胞群(例如,由细胞群产生)。在一些实施方案中,细胞群由单个细胞组成。在其他实施方案中,细胞群包含两个或更多个细胞。

[0052] 例如,在一些实施方案中,细胞群包含至少5个、至少10个、至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少150个、至少200个、至少250个、至少300个、至少350个、至少400个、至少450个、至少500个、至少600个、至少700个、至少800个、至少900个、至少 1×10^3 个、至少 1×10^4 个、至少 1×10^5 个、至少 1×10^6 个、至少 1×10^7 个、至少 1×10^8 个、至少 1×10^9 个或至少 1×10^{10} 个细胞。

[0053] 在一些实施方案中,所述群包含1-5、1-10、1-20、1-30、1-50、1-60、1-70、1-80、1-90、1-100、1-150、1-200、1-250、1-300、1-350、1-400、1-450、1-500、1-600、1-700、1-800、1-900、 $1-1 \times 10^3$ 、 $1-1 \times 10^4$ 、 $1-1 \times 10^5$ 、 $1-1 \times 10^6$ 、 $1-1 \times 10^7$ 、 $1-1 \times 10^8$ 、 $1-1 \times 10^9$ 、 $1-1 \times 10^{10}$ 、100-150、100-200、100-250、100-300、100-350、100-400、100-450、100-500、100-600、100-700、100-800、100-900、 $100-1 \times 10^3$ 、 $100-1 \times 10^4$ 、 $100-1 \times 10^5$ 、 $100-1 \times 10^6$ 、 $100-1 \times 10^7$ 、 $100-1 \times 10^8$ 、 $100-1 \times 10^9$ 、 $100-1 \times 10^{10}$ 、 $1 \times 10^3-1 \times 10^4$ 、 $1 \times 10^3-1 \times 10^5$ 、 $1 \times 10^3-1 \times 10^6$ 、 $1 \times 10^3-1 \times 10^7$ 、 $1 \times 10^3-1 \times 10^8$ 、 $1 \times 10^3-1 \times 10^9$ 、 $1 \times 10^3-1 \times 10^{10}$ 、 $1 \times 10^4-1 \times 10^5$ 、 $1 \times 10^4-1 \times 10^6$ 、 $1 \times 10^4-1 \times 10^7$ 、 $1 \times 10^4-1 \times 10^8$ 、 $1 \times 10^4-1 \times 10^9$ 、 $1 \times 10^4-1 \times 10^{10}$ 、 $1 \times 10^5-1 \times 10^6$ 、 $1 \times 10^5-1 \times 10^7$ 、 $1 \times 10^5-1 \times 10^8$ 、 $1 \times 10^5-1 \times 10^9$ 或 $1 \times 10^5-1 \times 10^{10}$ 个细胞。

[0054] 细胞群可以包含原核细胞和/或真核细胞。细胞群可以包含多个同质细胞。替代地,细胞群可以包含多个异质细胞。

[0055] 可以从受试者(例如,多细胞或共生生物)中分离细胞群。在一些实施方案中,所述受试者是小鼠、大鼠、兔、豚鼠、仓鼠、猪、羊、狗、灵长类动物、猫或人。

[0056] 分离细胞群的方法是本领域技术人员已知的。例如,制备复杂样品的方法可以包括活检、解剖(例如显微解剖,例如激光捕获)、有限稀释、显微操作、免疫磁性细胞分离、荧光激活细胞分选、密度梯度离心、免疫密度细胞分离、微流体细胞分选、沉降、粘附或其组合。

[0057] 在一些实施方案中,制备复杂样品的方法包括裂解细胞群,从而产生包含多种分子(例如,多肽、多核酸、代谢物等)的裂解样品。裂解细胞群的方法是本领域普通技术人员已知的。在一些实施方案中,使用任何一种已知的物理或化学方法裂解包含细胞的样品以从所述细胞释放靶分子。在一些实施方案中,可以使用电解法、酶促法、基于去污剂的方法和/或机械均化来裂解样品。在一些实施方案中,如果样品不包含细胞或组织(例如,包含纯化的多肽的样品),则可以省略裂解步骤。

[0058] 替代地或另外地,制备复杂样品的方法可以包括亚细胞分级分离(即,分离一个或多个细胞区室,例如内体、突触体、细胞质、核质、染色质、线粒体、过氧化物酶体、溶酶体、黑色素体、外来体、高尔基体、内质网、中心体、伪足或其组合)。

[0059] 来源于相同细胞群的分子在本文中被描述为具有相同的“来源”。

[0060] II. 制备富集样品的方法

[0061] 在一些方面,本公开涉及使样品富集一种或多种目的分子(例如,一种或多种目的

多肽)的方法。特别地,在一些方面,本公开涉及多肽富集的方法。如本文所用,术语“多肽富集”是指其中一种或多种目的多肽的丰度相对于一种或多种参考多肽(例如,复杂样品中的非目的多肽)的丰度增加的过程。如本文所用,术语“目的多肽”是指人们寻求富集的多肽。目的多肽可以包含特定的氨基酸序列。替代地或另外地,目的多肽可以包含特定的多肽修饰(例如,翻译后修饰)。这些方法有助于复杂样品的蛋白质组学分析,这些样品由许多不同的多肽组成,其中只有一些可能是感兴趣的(图1)。

[0062] 在一些实施方案中,用于多肽富集的方法包括使用多个富集分子从多个多肽中选择多肽子集,从而产生包含多肽子集的富集样品。在一些实施方案中,所述方法包括使多个多肽与多个富集分子接触以产生包含多个多肽中的多肽子集的富集样品。

[0063] 在一些实施方案中,用于多肽富集的方法包括:(a)使多个多肽与多个富集分子接触,其中多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集;和(b)分离结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。图3中说明的多肽富集方法学提供了该实施方案的实例。

[0064] 在一些实施方案中,用于多肽富集的方法包括:(a)使多个多肽与多个富集分子接触,其中多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集;和(b)分离未结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。图4中说明的多肽富集方法学提供了该实施方案的实例。

[0065] 在前述段落中描述的实施方案中,应理解富集分子与多肽的结合等同于多肽与富集分子的结合。因此,上述实施方案中的步骤(a)可以等效地描述为:(a)使多个多肽与多个富集分子接触,其中多个富集分子中的至少富集分子的子集被多个多肽中的多肽子集结合,从而产生结合的多肽子集和未结合的多肽子集。

[0066] 还应理解,上述实施方案的步骤(a)和(b)可以使用另外的多个富集分子重复一次或多次以产生进一步富集的样品。例如,在一些实施方案中,所述方法包括:(a)使多个多肽与第一多个富集分子接触,其中第一多个富集分子中的至少富集分子的子集与多个多肽中的多肽的子集结合,从而产生第一结合的多肽子集和第一未结合的多肽子集;(b)分离(a)的第一结合的多肽子集或第一未结合的多肽子集;和(c)用一个或多个另外的多个富集分子反复重复步骤(a)和(b)以产生包含多个多肽中的多肽子集的富集样品。在一些实施方案中,使用第二、第三、第四、第五、第六、第七、第八、第九、第十或任何数量的另外的多个富集分子重复步骤(a)和(b)。

[0067] 例如,在一些实施方案中,所述方法包括:(a)使多个多肽与第一多个富集分子接触,其中第一多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生第一结合的多肽子集和第一未结合的多肽子集;(b)分离(a)的第一结合的多肽子集或第一未结合的多肽子集;(c)使(b)的分离的多肽与第二多个富集分子接触,其中第二多个富集分子中的至少富集分子的子集与(b)中分离的多肽子集结合,从而产生第二结合的多肽子集和第二未结合的多肽子集;(d)分离(c)的第二结合的多肽子集或第二未结合的多肽子集以产生包含多个多肽中的多肽子集的富集样品。

[0068] 替代地或另外地,富集方法可以包括色谱法(例如,尺寸排阻、离子交换等)、等电聚焦、膜过滤、分子筛过滤、浓缩、沉淀(例如,冷沉淀)、干燥、透析或其组合。

[0069] 在一些实施方案中,所述方法包括使复杂样品与本文所述的试剂盒或装置接触。

参见“用于样品制备的试剂盒”和“用于样品制备和样品测序的装置”。

[0070] 在一些实施方案中,富集样品中的多肽是相同的(即,含有相同的氨基酸序列)。在一些实施方案中,富集样品包含至少两种独特的多肽(即,具有不同的氨基酸序列)。例如,在一些实施方案中,富集样品包含至少2种、至少3种、至少4种、至少5种、至少6种、至少7种、至少8种、至少9种、至少10种、至少11种、至少12种、至少13种、至少14种、至少15种、至少16种、至少17种、至少18种、至少19种、至少20种、至少25种、至少30种、至少40种、至少50种、至少60种、至少70种、至少80种、至少90种或至少100种独特的多肽。在一些实施方案中,富集样品包含1-2、1-5、1-10、1-15、1-20、1-30、1-40、1-50、1-60、1-70、1-80、1-90、1-100、2-5、2-10、2-15、2-20、2-30、2-40、2-50、2-60、2-70、2-80、2-90、2-100、5-10、5-15、5-20、5-30、5-40、5-50、5-60、5-70、5-80、5-90、5-100、10-15、10-20、10-30、10-40、10-50、10-60、10-70、10-80、10-90、10-100、15-20、20-30、20-40、20-50、20-60、20-70、20-80、20-90、20-100、20-30、20-40、20-50、20-60、20-70、20-80、20-90、20-100、30-40、30-50、30-60、30-70、30-80、30-90、30-100、40-50、40-60、40-70、40-80、40-90、40-100、50-60、50-70、50-80、50-90或50-100种独特的多肽。

[0071] 在一些实施方案中,富集样品包含具有至少50%、60%、70%、80%、90%、95%或99%序列同一性的多肽。在一些实施方案中,富集样品包含具有一种或多种多肽修饰(例如,翻译后修饰)的多肽。翻译后修饰的实例是本领域技术人员已知的,并且包括但不限于乙酰化、腺苷酰化、ADP-核糖基化、烷基化(例如甲基化)、酰胺化、精氨酸酰化、生物素化、丁酰化、氨甲酰化、羰基化、羧基化、瓜氨酸化、脱酰胺、消除、甲酰化、糖基化(例如,N-连接糖基化、O-连接糖基化)、glypyatyon、糖化、羟基化、碘化、ISG化、异戊二烯化、脂化、丙二酰化、肉豆蔻酰化、类泛素化、硝化、氧化、棕榈酰化聚乙二醇化、磷酸化、磷酸泛酰巯基乙炔化、聚乙二醇化、聚谷氨酰化、异戊二烯化、丙酰化、pupylation、S-谷胱甘肽化、S-亚硝基化、S-亚磺酰化、S-亚磺酰化、S-磺酰化、琥珀酰化、硫酸化、SUMO化和泛素化。

[0072] A. 富集分子

[0073] 如本文所用,术语“富集分子”是指表现出优先与(或被)一种或多种靶多肽结合分子。富集分子可以通过与靶多肽的氨基酸序列的直接相互作用而与(或被)靶多肽结合。替代地或另外地,富集分子可以通过与靶多肽的修饰(例如,翻译后修饰)进行相互作用而与(或被)靶多肽结合。富集分子与(或被)靶多肽的结合可以通过静电相互作用、疏水相互作用、互补形状或其组合来介导。

[0074] 在一些实施方案中,靶多肽是目的多肽。在其他实施方案中,靶多肽不是目的多肽。

[0075] 优先与一种或多种靶多肽(或靶多肽变体)结合的示例性富集分子包括免疫球蛋白、anticalin、脂质运载蛋白(lipocalin)、DARPin、适体、酶、凝集素和肽相互作用结构域。

[0076] 如本文所用,术语“免疫球蛋白”是指特征在于具有免疫球蛋白折叠并且起抗体作用并与一种或多种底物(例如,靶多肽)结合的多肽。因此,术语“免疫球蛋白”涵盖常规免疫球蛋白(即IgA、IgD、IgE、IgG和IgM)、单链可变片段(scFv)、抗原结合片段(Fab)、亲和体(affibody)和单域抗体(sdAb),例如纳米抗体、VHH和VNAR。

[0077] 如本文所用,术语“适体”是指优先与一种或多种靶分子(例如,靶多肽)结合的多

核酸(例如,DNA或RNA)或多肽。尽管在自然界中发现了一些实例,但适体通常通过反复几轮的体外选择来进行工程化。

[0078] 如本文所用,术语“酶”是指在结合一种或多种底物(例如,靶多肽)时加速化学反应的大分子生物催化剂。通常,酶会在化学反应完成后释放其底物。因此,在其中富集分子包含酶的一些实施方案中,酶被催化失活以增加酶保持与底物结合的可能性。催化失活可以通过一种或多种酶促辅因子(即对于酶作为催化剂的活性所需的非蛋白质化合物或金属离子)的变异和/或消耗来进行。

[0079] 如本文所用,术语“肽相互作用结构域”是指与一种或多种多肽(例如,靶多肽)相互作用的多肽(或多肽的一部分)。例如,肽相互作用结构域可以是支架蛋白、多蛋白复合物的多肽或其部分。

[0080] 在一些实施方案中,富集分子包含免疫球蛋白、适体、酶和/或肽相互作用结构域。

[0081] 优先被一种或多种靶多肽结合的示例性富集分子包括寡核苷酸(例如,双链DNA、单链DNA、双链RNA、单链RNA等)、寡糖(或多糖)、脂质、糖蛋白、受体配体、受体激动剂、受体拮抗剂、酶底物和酶辅因子。

[0082] 在一些实施方案中,富集分子包含寡核苷酸(例如,双链DNA、单链DNA、双链RNA、单链RNA等)、寡糖、脂质、受体配体、受体激动剂、受体拮抗剂、酶底物和/或酶辅因子。

[0083] 本文使用优先结合来表征富集分子以强调:(i)富集分子不需要表现出高特异性(即,仅与单个靶多肽结合(或被其结合)到可观的水平);(ii)富集分子可能表现出某种程度的脱靶结合(即,与脱靶分子结合(或被其结合)到可检测的水平);和(iii)富集分子不需要以100%的效率与靶多肽结合(即,即使在存在过量富集分子的情况下,也并非必然需要复杂样品中的所有靶多肽都被结合)。

[0084] 在一些实施方案中,富集分子优先与单个靶多肽结合(或优先被其结合)(例如图2的富集分子1)。然而,在其他实施方案中,富集分子优先与两个或更多个靶多肽结合(或优先被其结合)(例如图2的富集分子2和3)。

[0085] 在一些实施方案中,富集分子表现出优先与至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少25个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、或至少100个、至少200个、至少300个、至少400个、至少500个、至少600个、至少700个、至少800个、至少900个、至少1000个、至少2000个、至少3000个、至少4000个、至少5000个或至少10,000个靶多肽结合(或优先被其结合)。

[0086] 在一些实施方案中,富集分子表现出优先与二、三、四、五、六、七、八、九、十、十一、十二、十三、十四或十五个靶多肽结合(或优先被其结合)。

[0087] 在一些实施方案中,富集分子表现出优先与1-2、1-5、1-10、1-15、1-20、1-30、1-40、1-50、1-60、1-70、1-80、1-90、1-100、2-5、2-10、2-15、2-20、2-30、2-40、2-50、2-60、2-70、2-80、2-90、2-100、5-10、5-15、5-20、5-30、5-40、5-50、5-60、5-70、5-80、5-90、5-100、10-15、10-20、10-30、10-40、10-50、10-60、10-70、10-80、10-90、10-100、15-20、20-30、20-40、20-50、20-60、20-70、20-80、20-90、20-100、20-30、20-40、20-50、20-60、20-70、20-80、20-90、20-100、30-40、30-50、30-60、30-70、30-80、30-90、30-100、40-50、40-60、40-70、40-

80、40-90、40-100、50-60、50-70、50-80、50-90、或50-100、100-200、100-300、100-400、100-500、100-600、100-700、100-800、100-900、100-1000、100-5000、100-10,000、500-600、500-700、500-800、500-900、500-1000、500-5000、500-10,000、1000-5000或1000-10,000个靶多肽结合(或优先被其结合)。

[0088] 在一些实施方案中,富集分子表现出优先与多个相关靶多肽(例如,2、3、4、5、6、7、8、9、10、20、30、40、50或更多个相关多肽)结合(或优先被其结合),所述多肽具有至少50%、60%、70%、80%、90%、95%或99%的序列同源性。

[0089] 在一些实施方案中,富集分子表现出优先与翻译后修饰结合(或优先被其结合),翻译后修饰例如乙酰化、腺苷酰化、ADP-核糖基化、烷基化(例如甲基化)、酰胺化、精氨酸化、生物素化、丁酰化、氨甲酰化、羰基化、羧基化、瓜氨酸化、脱酰胺、消除、甲酰化、糖基化(例如,N-连接糖基化、O-连接糖基化)、glypyatyon、糖化、羟基化、碘化、ISG化、异戊二烯化、脂化、丙二酰化、肉豆蔻酰化、类泛素化、硝化、氧化、棕榈酰化聚乙二醇化、磷酸化、磷酸泛酰巯基乙炔化、聚乙二醇化、聚谷氨酰化、异戊二烯化、丙酰化、pupylation、S-谷胱甘肽化、S-亚硝基化、S-亚磺酰化、S-亚磺酰化、S-磺酰化、琥珀酰化、硫酸化、SUMO化和泛素化。

[0090] 富集分子可以被固定在(例如,共价连接到)基质(例如,如“用于样品制备和样品测序的装置”中描述的捕获探针)上。基质可以是表面(例如固体表面)、珠粒(例如磁珠)、颗粒(例如磁性颗粒)或凝胶。

[0091] (i) 多个富集分子

[0092] 通常,本文所述的富集方法利用多个富集分子。多个富集分子可以在化学上相同(即,多个具有一种富集分子“类型”)。替代地,多个富集分子可以包含不同富集分子的组合(即,具有两种或更多种富集分子“类型”)。

[0093] 在一些实施方案中,多个富集分子包含单一富集分子类型。在其他实施方案中,多个富集分子包含两种或更多种、三种或更多种、四种或更多种、五种或更多种、六种或更多种、七种或更多种、八种或更多种、九种或更多种、十种或更多种、十一种或更多种、十二种或更多种、十三种或更多种、十四种或更多种、或十五种或更多种富集分子类型的组合。在一些实施方案中,多个富集分子包含至少2种、至少3种、至少4种、至少5种、至少6种、至少7种、至少8种、至少9种、至少10种、至少11种、至少12种、至少13种、至少14种、至少15种、至少16种、至少17种、至少18种、至少19种、至少20种、至少25种、至少30种、至少40种、至少50种、至少60种、至少70种、至少80种、至少90种、或至少100种、至少200种、至少300种、至少400种、至少500种富集分子类型。

[0094] 在一些实施方案中,多个富集分子包含两种、三种、四种、五种、六种、七种、八种、九种、十种、十一种、十二种、十三种、十四种或十五种富集分子类型的组合。

[0095] 在一些实施方案中,多个富集分子包含1-2、1-5、1-10、1-15、1-20、1-30、1-40、1-50、1-60、1-70、1-80、1-90、1-100、2-5、2-10、2-15、2-20、2-30、2-40、2-50、2-60、2-70、2-80、2-90、2-100、5-10、5-15、5-20、5-30、5-40、5-50、5-60、5-70、5-80、5-90、5-100、10-15、10-20、10-30、10-40、10-50、10-60、10-70、10-80、10-90、10-100、15-20、20-30、20-40、20-50、20-60、20-70、20-80、20-90、20-100、20-30、20-40、20-50、20-60、20-70、20-80、20-90、20-100、30-40、30-50、30-60、30-70、30-80、30-90、30-100、40-50、40-60、40-70、40-80、40-90、40-100、50-60、50-70、50-80、50-90、或50-100、100-200、100-300、100-400或100-500种

富集分子类型的组合。

[0096] 在一些实施方案中,多个富集分子中的每个富集分子优先与单个靶多肽结合(或优先被其结合)。在其他实施方案中,多个富集分子中的一个或多个(例如,子集)富集分子表现出优先与两个或更多个靶多肽结合(或优先被其结合)。在其他实施方案中,多个富集分子中的每个富集分子表现出优先与两个或更多个靶多肽结合(或优先被其结合)。

[0097] 在一些实施方案中,多个富集分子中的一种或多种(例如,子集)富集分子与翻译后多肽修饰结合。在其他实施方案中,多个富集分子中的每个富集分子表现出优先与两个或更多个翻译后多肽修饰结合。

[0098] 在一些实施方案中,多个富集分子中的每个富集分子被固定在(例如共价连接到)基质(例如,如“用于样品制备和样品测序的装置”中描述的捕获探针),例如表面(例如固体表面)、珠粒(例如磁珠)、颗粒(例如磁性颗粒,或凝胶)。在一些实施方案中,多个富集分子中的一个或多个(例如,子集)被固定在(例如,共价连接到)基质。因此,在一些实施方案中,当包含多个多肽的样品接触基质时,会发生多个多肽与多个富集分子的接触。

[0099] 例如,在一些实施方案中,富集分子被固定在(例如,共价连接或交联到)凝胶上并且样品被拉过凝胶。在一些实施方案中,富集分子被固定在(例如,共价连接到)珠粒(例如,磁珠)上,然后被拉下。

[0100] (ii) 多重富集分子

[0101] 如上所述,在一些实施方案中,所述方法包括:(a)使多个多肽与第一多个富集分子接触,其中第一多个富集分子中的至少富集分子的子集与多个多肽中的多肽子集结合,从而产生第一结合的多肽子集和第一未结合的多肽子集;(b)分离(a)的第一结合的多肽子集或第一未结合的多肽子集;和(c)用一种或多种另外的多个富集分子反复重复步骤(a)和(b)以产生包含多个多肽中的多肽子集的富集样品。在一些实施方案中,使用第二、第三、第四、第五、第六、第七、第八、第九、第十或任何数量的另外的多个富集分子重复步骤(a)和(b)。

[0102] 在一些实施方案中,多肽富集方法中使用的每个多个富集分子中是独特的(即,每个多个富集分子包含不同的多个富集分子)。在其他实施方案中,两个或更多个多个富集分子是相同的。在一些实施方案中,至少一个多个富集分子中靶向翻译后多肽修饰并且至少一个多个富集分子不靶向翻译后修饰。

[0103] 例如,第一富集步骤(利用第一多个富集分子)可以富集特定的翻译后多肽修饰,并且第二富集步骤(利用第二多个富集分子)可以富集特定的多肽(和该多肽的变体)。替代地,第一富集步骤(利用第一多个富集分子)可以富集特定的多肽(和该多肽的变体),并且第二富集步骤(利用第二多个富集分子)可以富集特定的翻译后修饰。

[0104] B. 多肽修饰

[0105] 复杂样品的一个或多个多肽可以在上述多肽富集之前、同时和/或之后进行体外修饰。例如,在一些实施方案中,在进行多肽富集之前、同时和/或之后将复杂样品与修饰剂接触。其中,修饰剂可以介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。

[0106] 在一些实施方案中,复杂样品的一个或多个多肽通过片段化进行修饰。在一些实施方案中,片段化包括酶促消化。在一些实施方案中,通过在消化条件下使多肽与内肽酶

(例如胰蛋白酶)接触来进行消化。在一些实施方案中,片段化包括化学消化。用于化学和酶促消化的合适试剂的实例是本领域已知的,并且包括但不限于胰蛋白酶、化学胰蛋白酶、Lys-C、Arg-C、Asp-N、Lys-N、BNPS-粪臭素、CNBr、半胱天冬酶、甲酸、谷氨酰内肽酶、羟胺、碘代苯甲酸、中性粒细胞弹性蛋白酶、胃蛋白酶、脯氨酸-内肽酶、蛋白酶K、葡萄球菌肽酶I、嗜热菌蛋白酶和凝血酶。

[0107] 在一些实施方案中,复杂样品的一个或多个多肽通过变性(例如,通过热和/或化学方式)进行修饰。

[0108] 在一些实施方案中,复杂样品的一个或多个多肽通过体外翻译后修饰进行修饰,例如通过乙酰化、腺苷酰化、ADP-核糖基化、烷基化(例如甲基化)、酰胺化、精氨酰化、生物素化、丁酰化、氨甲酰化、羰基化、羧基化、瓜氨酸化、脱酰胺、消除、甲酰化、糖基化(例如,N-连接糖基化、O-连接糖基化)、glypyatyon、糖化、羟基化、碘化、ISG化、异戊二烯化、脂化、丙二酰化、肉豆蔻酰化、类泛素化、硝化、氧化、棕榈酰化聚乙二醇化、磷酸化、磷酸泛酰巯基乙炔化、聚乙二醇化、聚谷氨酰化、异戊二烯化、丙酰化、pupylation、S-谷胱甘肽化、S-亚硝基化、S-亚磺酰化、S-亚磺酰化、S-磺酰化、琥珀酰化、硫酸化、SUMO化或泛素化。

[0109] 在一些实施方案中,复杂样品的一个或多个多肽通过封闭一种或多种官能团(例如,游离羧酸根基团和/或硫醇基团)进行修饰。

[0110] 在一些实施方案中,封闭游离羧酸根基团是指对这些基团改变了相对于未经修饰的羧酸根的化学反应性的化学修饰。合适的羧酸根封闭方法是本领域已知的并且应该将侧链羧酸根基团修饰为在化学上不同于待官能化的多肽的羧基末端羧酸根基团。在一些实施方案中,封闭游离羧酸根基团包括多肽的游离羧酸根基团的酯化或酰胺化。在一些实施方案中,封闭游离羧酸根基团包括多肽的游离羧酸根基团的甲酯化,例如,通过使多肽与甲醇HCl反应。可用于封闭游离羧酸根基团的试剂和技术的另外实例包括但不限于4-磺基-2,3,5,6-四氟苯酚(STP)和/或碳二亚胺例如N-(3-二甲基氨基丙基)-N'-乙基碳二亚胺盐酸盐(EDAC)、脲试剂、重氮甲烷、用于Fischer酯化的醇和酸,使用N-羟基琥珀酰亚胺(NHS)形成NHS酯(可能作为后续酯或胺形成的中间体),或与羰基二咪唑(CDI)反应或形成混合酸酐,或任何其他修饰或封闭羧酸的方法,可能通过形成酯或酰胺。

[0111] 在一些实施方案中,封闭游离硫醇基团是指对这些基团改变了相对于未经修饰的硫醇的化学反应性的化学修饰。在一些实施方案中,封闭游离硫醇基团包括对多肽的游离硫醇基团进行还原和烷基化。在一些实施方案中,通过使多肽与二硫苏糖醇(DTT)以及碘乙酰胺和碘乙酸中的一种或两种接触来进行还原和烷基化。可以使用的另外的和替代的半胱氨酸还原试剂的实例是众所周知的,并且包括但不限于2-巯基乙醇、三(2-羧乙基)膦盐酸盐(TCEP)、三丁基膦、二硫丁胺(DTBA)或任何能够还原硫醇基团的试剂。可以使用的另外的和替代的半胱氨酸封闭(例如,半胱氨酸烷基化)试剂的实例是众所周知的,并且包括但不限于丙烯酰胺、4-乙烯基吡啶、N-乙基马来酰亚胺(NEM)、N-ε-马来酰亚胺基己酸(EMC),或任何修饰半胱氨酸以防止二硫键形成的试剂。

[0112] 在一些实施方案中,多肽的N末端氨基酸或C末端氨基酸被修饰。

[0113] 在一些实施方案中,多肽的羧基末端在包括以下的方法中被修饰:(i)封闭多肽的游离羧酸根基团;(ii)使多肽变性(例如,通过热和/或化学方式);(iii)封闭多肽的游离硫醇基团;(iv)消化多肽以产生至少一个包含游离的C末端羧酸根基团的多肽片段;和(v)将

功能部分缀合(例如,化学地)至游离的C末端羧酸根基团。在一些实施方案中,所述方法进一步包括,在(i)之后和(ii)之前,对包含多肽的样品进行透析。

[0114] 在一些实施方案中,以包括以下的方法修饰多肽的羧基末端:(i)使多肽变性(例如,通过热和/或化学方式);(ii)封闭多肽的游离硫醇基团;(iii)消化多肽以产生至少一个包含游离的C末端羧酸根基团的多肽片段;(iv)封闭游离C末端羧酸根基团以产生至少一个包含封闭的C末端羧酸根基团的多肽片段;和(v)将功能部分缀合(例如,酶促地)至封闭的C末端羧酸根基团。在一些实施方案中,所述方法进一步包括,在(iv)之后和(v)之前,对包含多肽的样品进行透析。

[0115] 在一些实施方案中,复杂样品在富集之前与修饰剂接触以介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。替代地或另外地,在一些实施方案中,复杂样品在富集的同时与修饰剂接触以介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。替代地或另外地,在一些实施方案中,复杂样品(或源自其的样品,包含一个或多个目的多肽)在富集之后与修饰剂接触以介导多肽片段化、多肽变性、翻译后修饰的添加和/或一种或多种官能团的封闭。

[0116] III. 多肽测序方法学

[0117] 在一些实施方案中,在富集之后对富集样品的多肽进行测序和/或进行鉴定。因此,在一些方面,本公开涉及多肽测序和鉴定的方法。对多肽分子进行测序的各种方法是本领域普通技术人员已知的,并且包括质谱法(例如,肽质量指纹图谱和串联质谱法)和Edman降解。另外,本文描述了对多肽进行测序的方法。

[0118] 如本文所用,关于多肽的“测序”、“序列确定”、“确定序列”和类似术语包括确定多肽的部分氨基酸序列信息以及完整氨基酸序列信息。也就是说,该术语包括序列比较、指纹识别和关于靶分子的类似信息水平,以及靶分子在目的区域内的每个氨基酸的明确鉴定和排序。该术语包括鉴定多肽的单个氨基酸(或单个氨基酸的概率)。在一些实施方案中,鉴定了多肽的多于一个氨基酸(或多于一个氨基酸的概率)。因此,在一些实施方案中,如本文所用的术语“氨基酸序列”和“多肽序列”可以指多肽材料本身并且不限于在生物化学上表征特定多肽的特定序列信息(例如,表示氨基酸顺序的从一个末端到另一个末端的字母串)。

[0119] 在一些实施方案中,确定多肽内特定位置处氨基酸的概率并在概率阵列中进行说明。例如,对于由两个氨基酸组成的多肽,术语“测序”、“序列确定”、“确定序列”等术语可能涉及确定位置1和/或位置2处的氨基酸的概率,例如[[0.80,0.12,0.05,0.01,0.01,0.01,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00],[0.00,0.10,0.90,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00,0.00],[0.00,0.00,0.00,0.00]],其中阵列中的概率分别对应于A、R、N、D、C、Q、E、G、H、I、L、K、M、F、P、S、T、W、Y和V。本领域普通技术人员将理解,该实例(和示例性概率阵列)可以扩大到适应另外的氨基酸身份(例如,修饰的氨基酸)的分析,例如本文所述的那些。

[0120] 在一些实施方案中,多肽分子的测序包括鉴定多肽分子中至少两个(例如,至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少25个、至少30个、至少35个、至少40个、至少45个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个或更多个)氨基酸(或氨基酸概率)。在一些实施方案中,所述至少

两个氨基酸是连续的氨基酸。在一些实施方案中,所述至少两个氨基酸是非连续的氨基酸。

[0121] 在一些实施方案中,多肽分子的测序包括鉴定多肽分子中所有氨基酸的少于100% (例如,少于99%、少于95%、少于90%、少于85%、少于80%、少于75%、少于70%、少于65%、少于60%、少于55%、少于50%、少于45%、少于40%、少于35%、少于30%、少于25%、少于20%、少于15%、少于10%、少于5%、少于1%或更少)。例如,在一些实施方案中,多肽分子的测序包括鉴定多肽分子中一种类型的氨基酸的少于100% (例如,鉴定多肽分子中的一种类型的所有氨基酸的一部分)。在一些实施方案中,多肽分子的测序包括鉴定多肽分子中每种类型的氨基酸的少于100%。

[0122] 在一些实施方案中,多肽分子的测序包括鉴定多肽中至少1、至少5、至少10、至少15、至少20、至少25、至少30、至少35、至少40、至少45、至少50、至少55、至少60、至少65、至少70、至少75、至少80、至少85、至少90、至少95、至少100或更多种类型的氨基酸。

[0123] 在一些实施方案中,本申请提供了用于通过随时间鉴定存在于多肽末端的一系列氨基酸 (例如,通过末端氨基酸的迭代检测和裂解) 来对多肽进行测序的组合物和方法。在其他实施方案中,本申请提供了用于通过鉴定多肽的标记的氨基含量并与参考序列数据库进行比较来对多肽进行测序的组合物和方法。

[0124] 在一些实施方案中,本申请提供了用于通过对多肽的多个片段进行测序来对多肽进行测序的组合物和方法。在一些实施方案中,对多肽进行测序包括组合多个多肽片段的序列信息以鉴定和/或确定多肽的序列。在一些实施方案中,组合序列信息可以由计算机硬件和软件来执行。参见“用于样品制备和样品测序的装置”。本文所述的方法可以允许对一组相关的多肽,例如生物体的整个蛋白质组进行测序。在一些实施方案中,根据本申请的方面,多个单分子测序反应并行进行 (例如,在单个芯片上)。例如,在一些实施方案中,多个单分子测序反应各自在单个芯片或阵列上的单独样品孔中进行。

[0125] 在一些实施方案中,本文提供的方法可用于对包含多肽的复杂混合物或富集混合物的样品中的单个多肽进行测序和鉴定。在一些实施方案中,本申请提供了独特鉴定多肽的复杂混合物或富集混合物中的单个多肽的方法。在一些实施方案中,通过确定多肽的部分氨基酸序列来检测混合样品中的单个多肽。在一些实施方案中,多肽的部分氨基酸序列在大约5到50个氨基酸的连续区段内。

[0126] 不希望囿于任何特定理论的,认为大多数人类蛋白质可以参考蛋白质组数据库使用不完整的序列信息来鉴定。例如,人类蛋白质组的简单建模表明,大约98%的蛋白质可以通过仅检测6到40个氨基酸的区段中的四种类型的氨基酸来独特鉴定 (参见例如 Swaminathan 等人, PLoS Comput Biol. 2015, 11 (2) : e1004080; 和 Yao 等人, Phys. Biol. 2015, 12 (5) : 055003)。因此,多肽的复杂混合物或富集混合物可以被降解 (例如,化学降解、酶促降解) 成大约6到40个氨基酸的短多肽片段,并且对该多肽文库的测序将揭示存在于原始复杂混合物或富集混合物中的每种多肽的身份和丰度。用于通过确定部分序列信息来选择性标记氨基酸和鉴定多肽的组合物和方法在2015年9月15日提交的标题为“SINGLE MOLECULE PEPTIDE SEQUENCING”的美国专利申请号15/510,962中进行了详细描述,其全文通过引用的方式并入本文。

[0127] 实施方案能够以高准确度,例如以至少约50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、99.9%、99.99%、99.999%或99.9999%的准确度对单

个多肽分子进行测序。在一些实施方案中,单分子测序中使用的靶分子是被固定在固体支持物表面(例如样品孔的底部表面或侧壁表面)上的多肽。根据本申请,样品孔还可以包含测序反应所需的其他试剂,例如一种或多种合适的缓冲液、辅因子、标记的亲试剂和酶(例如有催化活性或无催化活性的外肽酶,其可以被发光标记或不被发光标记)。

[0128] 在一些方面,根据本申请的测序可以涉及将多肽固定在基质(例如,固相支持物,例如芯片,例如本文所述的集成装置)的表面上。在一些实施方案中,可以将多肽固定在基质上的样品孔的表面上(例如,样品孔的底部表面上)。在一些实施方案中,多肽的N末端氨基酸是固定化的(例如,连接到表面)。在一些实施方案中,多肽的C末端氨基酸是固定化的(例如,连接到表面)。在一些实施方案中,一种或多种非末端氨基酸是固定化的(例如,连接到表面)。可以使用任何合适的共价或非共价键连接固定化的氨基酸,例如如本申请中所述。在一些实施方案中,多个多肽被连接到多个样品孔(例如,一个多肽连接到每个样品孔的表面,例如底部表面),例如在基质上的样品孔的阵列中。

[0129] 在一些方面,根据本申请的测序可以使用允许单分子分析的系统进行。所述系统可以包括测序装置和被设置为与测序装置接合的仪器。参见“用于样品制备和样品测序的装置”。

[0130] A. 标记的亲试剂和使用方法

[0131] 在一些实施方案中,本文提供的方法包括使多肽与选择性结合一种类型的末端氨基酸的标记的亲试剂(本文也称为氨基酸识别分子,其可以包含或不包含标记)接触。如本文所用,在一些实施方案中,末端氨基酸可以指多肽的氨基末端氨基酸或多肽的羧基末端氨基酸。在一些实施方案中,标记的亲试剂选择性结合一种类型的末端氨基酸而不是其他类型的末端氨基酸。在一些实施方案中,标记的亲试剂选择性结合一种类型的末端氨基酸而不是相同类型的内部氨基酸。在其他实施方案中,标记的亲试剂在多肽的任何位置选择性结合一种类型的氨基酸,例如,与末端氨基酸和内部氨基酸相同类型的氨基酸。

[0132] 如本文所用,在一些实施方案中,一种类型的氨基酸是指二十种天然存在的氨基酸中的一种或其类型的子集。在一些实施方案中,一种类型的氨基酸是指二十种天然存在的氨基酸之一的经修饰的变体或其未经修饰的和/或经修饰的变体的子集。经修饰的氨基酸变体的实例包括但不限于经翻译后修饰(例如乙酰化、ADP-核糖基化、半胱天冬酶裂解、瓜氨酸化、甲酰化、N-连接糖基化、O-连接糖基化、羟基化、甲基化、肉豆蔻酰化、类泛素化、硝化、氧化、棕榈酰化、磷酸化、异戊二烯化、S-亚硝基化、硫酸化、苏素化和泛素化)的变体、经化学修饰的变体、非天然氨基酸和蛋白原氨基酸(例如硒代半胱氨酸和吡咯赖氨酸)。在一些实施方案中,氨基酸类型的子集包括多于一种且少于二十种氨基酸,其具有一种或多种相似的生化特性。例如,在一些实施方案中,一种类型的氨基酸是指选自以下的一种类型:具有带电荷侧链(例如,带正电荷和/或带负电荷的侧链)的氨基酸、具有极性侧链(例如,极性不带电荷的侧链)的氨基酸、具有非极性侧链(例如,非极性脂肪族和/或芳香族侧链)的氨基酸和具有疏水性侧链的氨基酸。

[0133] 在一些实施方案中,本文提供的方法包括使多肽与选择性结合一种或多种类型的末端氨基酸的一种或多种标记的亲试剂接触。作为说明性和非限制性实例,当在本申请的方法中使用四种标记的亲试剂时,任何一种试剂选择性结合一种类型的末端氨基酸,所述末端氨基酸不同于其他三种氨基酸中的任何一种选择性结合的另一类型的氨基酸

(例如,第一试剂结合第一类型,第二试剂结合第二类型,第三试剂结合第三类型,第四试剂结合末端氨基酸的第四类型)。出于此讨论的目的,在本文所述的方法的上下文中的一种或多种标记的亲试剂可以替代地称为一组标记的亲试剂。

[0134] 在一些实施方案中,一组标记的亲试剂包括至少一种和多达六种标记的亲试剂。例如,在一些实施方案中,一组标记的亲试剂包括一种、两种、三种、四种、五种或六种标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括十种或更少的标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括八种或更少的标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括六种或更少的标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括四种或更少的标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括三种或更少的标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括两种或更少的标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括四种标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括至少两种和多达二十种(例如,至少两种和多达十种、至少两种和多达八种、至少四种和多达二十种、至少四种和达到十种)标记的亲试剂。在一些实施方案中,一组标记的亲试剂包括超过二十种(例如,20至25、20至30)亲试剂。然而,应当理解,可以根据本申请的方法使用任意数量的亲试剂以适应所需用途。

[0135] 根据本申请,在一些实施方案中,通过检测标记的亲试剂(例如,包含发光标记的氨基酸识别分子)的发光来鉴定一种或多种类型的氨基酸。在一些实施方案中,标记的亲试剂包括选择性结合一种类型的氨基酸的亲试剂和具有与该亲试剂相关的发光的发光标记。以这种方式,发光(例如,发光寿命、发光强度和本文别处描述的其他发光特性)可以与亲试剂的选择性结合相关以鉴定多肽的氨基酸。在一些实施方案中,多种类型的标记的亲试剂可用于根据本申请的方法中,其中每种类型包括具有可从多种中独特识别的发光的发光标记。合适的发光标记可以包括发光分子,例如荧光团染料,并且在本文别处进行了描述。

[0136] 在一些实施方案中,通过检测标记的亲试剂的一种或多种电特性来鉴定一种或多种类型的氨基酸。在一些实施方案中,标记的亲试剂包括选择性结合一种类型的氨基酸的亲试剂和与该亲试剂相关的电导标记。以这种方式,一种或多种电特性(例如,电荷、电流振荡颜色(current oscillation color)和其他电特性)可以与亲试剂的选择性结合相关以鉴定多肽的氨基酸。在一些实施方案中,多种类型的标记的亲试剂可用于根据本申请的方法中,其中每种类型包含电导标记,其产生电信号的变化(例如,电导的变化,例如特征模式的电导率和电导率转变的幅度),其可从多个中独特地鉴定。在一些实施方案中,多种类型的标记的亲试剂各自包含具有不同数量的带电荷基团(例如,不同数量的带负电荷和/或带正电荷的基团)的电导标记。因此,在一些实施方案中,电导标记是电荷标记。电荷标记的实例包括树枝状大分子、纳米颗粒、核酸和其他具有多个带电荷基团的聚合物。在一些实施方案中,电导标记可通过其净电荷(例如,净正电荷或净负电荷)、通过其电荷密度和/或通过其带电荷基团的数量来独特地鉴定。

[0137] 在一些实施方案中,亲试剂(例如,氨基酸识别分子)可以由本领域技术人员使用常规已知技术进行工程化。在一些实施方案中,期望的性质可以包括仅当一种类型的氨基酸位于多肽的末端(例如,N末端或C末端)时才以高亲和力选择性结合一种类型的氨基酸

的能力。在其他实施方案中,期望的性质可以包括当一种类型的氨基酸位于多肽的末端(例如,N末端或C末端)时以及当其位于多肽的内部位置时以高亲和力选择性结合一种类型的氨基酸的能力。

[0138] 如本文所用,在一些实施方案中,术语“选择性的”和“特异性的”(及其变形,例如,选择性地、特异性地、选择性、特异性)是指优先结合相互作用。例如,在一些实施方案中,选择性结合一种类型的氨基酸的标记的亲试剂优先结合一种类型而不是另一种类型的氨基酸。选择性结合相互作用将区分一种类型的氨基酸(例如,一种类型的末端氨基酸)和其他类型的氨基酸(例如,其他类型的末端氨基酸),通常超过约10至100倍或更多(例如,超过约1,000或10,000倍)。因此,应当理解,选择性结合相互作用可以指相比于其他类型的氨基酸可以与一种类型的氨基酸独特地识别的任何结合相互作用。例如,在一些方面,本申请通过获得指示一种或多种氨基酸识别分子与多肽分子的关联的数据来提供多肽测序的方法。在一些实施方案中,所述数据包括对应于与多肽分子的氨基酸的一系列可逆氨基酸识别分子结合相互作用的一系列信号脉冲,并且所述数据可用于确定氨基酸的身份。因此,在一些实施方案中,“选择性”或“特异性”结合相互作用是指检测到的区分一种类型的氨基酸和其他类型的氨基酸的结合相互作用。

[0139] 在一些实施方案中,标记的亲试剂(例如,氨基酸识别分子)以小于约 10^{-6} M(例如,小于约 10^{-7} M、小于约 10^{-8} M、小于约 10^{-9} M、小于约 10^{-10} M、小于约 10^{-11} M、小于约 10^{-12} M、至低至 10^{-16} M)的解离常数(K_D)选择性结合一种类型的氨基酸,而不与其他类型的氨基酸显著结合。在一些实施方案中,标记的亲试剂以小于约100nM、小于约50nM、小于约25nM、小于约10nM或小于约1nM的 K_D 选择性结合一种类型的氨基酸(例如,一种类型的末端氨基酸)。在一些实施方案中,标记的亲试剂以约50nM至约50 μ M(例如,约50nM至约500nM、约50nM至约5 μ M、约500nM至约50 μ M、约5 μ M至约50 μ M或约10 μ M至约50 μ M)的 K_D 选择性结合一种类型的氨基酸。在一些实施方案中,氨基酸识别分子以约50nM的 K_D 结合一种类型的氨基酸。

[0140] 在一些实施方案中,标记的亲试剂(例如,氨基酸识别分子)以小于约 10^{-6} M(例如,小于约 10^{-7} M、小于约 10^{-8} M、小于约 10^{-9} M、小于约 10^{-10} M、小于约 10^{-11} M、小于约 10^{-12} M、至低至 10^{-16} M)的 K_D 结合两种或更多种类型的氨基酸。在一些实施方案中,氨基酸识别分子以小于约100nM、小于约50nM、小于约25nM、小于约10nM或小于约1nM的 K_D 结合两种或更多种类型的氨基酸。在一些实施方案中,氨基酸识别分子以约50nM至约50 μ M(例如,约50nM至约500nM、约50nM至约5 μ M、约500nM至约50 μ M、约5 μ M至约50 μ M或约10 μ M至约50 μ M)的 K_D 结合两种或更多种类型的氨基酸。在一些实施方案中,氨基酸识别分子以约50nM的 K_D 结合两种或更多种类型的氨基酸。

[0141] 在一些实施方案中,标记的亲试剂(例如氨基酸识别分子)以至少 $0.1s^{-1}$ 的解离速率(k_{off})结合至少一种类型的氨基酸。在一些实施方案中,解离速率在约 $0.1s^{-1}$ 和约1,000 s^{-1} 之间(例如,在约 $0.5s^{-1}$ 和约500 s^{-1} 之间、在约 $0.1s^{-1}$ 和约100 s^{-1} 之间、在约 $1s^{-1}$ 和约100 s^{-1} 之间或在约 $0.5s^{-1}$ 和约50 s^{-1} 之间)。在一些实施方案中,解离速率在约 $0.5s^{-1}$ 和约20 s^{-1} 之间。在一些实施方案中,解离速率在约2 s^{-1} 和约20 s^{-1} 之间。在一些实施方案中,解离速率在约 $0.5s^{-1}$ 和约2 s^{-1} 之间。

[0142] 在一些实施方案中, K_D 或 k_{off} 的值可以是已知的文献值,或者该值可以根据经验确定。例如, K_D 或 k_{off} 的值可以在单分子测定或整体测定中测量。在一些实施方案中, k_{off}

的值可以基于在如本文别处所述的单分子测定中获得的信号脉冲信息根据经验确定。例如, koff的值可以近似为平均脉冲持续时间的倒数。在一些实施方案中, 氨基酸识别分子结合两种或更多种类型的氨基酸, 这两种或更多种类型中的每一种具有不同的KD或koff。在一些实施方案中, 第一类型氨基酸的第一KD或koff与第二类型氨基酸的第二KD或koff相差至少10% (例如, 至少25%、至少50%、至少100%或更多)。在一些实施方案中, KD或koff的第一和第二值相差约10-25%、25-50%、50-75%、75-100%或大于100%, 例如相差约2倍、3倍、4倍、5倍或更多。

[0143] 在一些实施方案中, 标记的亲试剂包含发光标记 (例如, 标记) 和选择性结合多肽的一种或多种类型的末端氨基酸的亲试剂。在一些实施方案中, 亲试剂对末端位置或末端和内部位置处的一种类型的氨基酸或氨基酸类型的子集 (例如, 少于二十种常见类型的氨基酸) 具有选择性。

[0144] 如本文所述, 亲试剂 (也称为“识别分子”) 可以是能够选择性或特异性结合一个分子而不是另一个分子 (例如, 一种类型的氨基酸而不是另一种类型的氨基酸, 如具有本文所指的“氨基酸识别分子”) 的任何生物分子。亲试剂 (例如识别分子) 包括例如蛋白质和核酸, 其可以是合成的或重组的。在一些实施方案中, 亲试剂或识别分子可以是抗体或抗体的抗原结合部分, 或酶促生物分子, 例如肽酶、氨基转移酶、核酶、适体酶或tRNA合成酶, 包括氨酰基-tRNA合成酶和描述于2016年9月2日提交的标题为“MOLECULES AND METHODS FOR ITERATIVE POLYPEPTIDE ANALYSIS AND PROCESSING”的美国专利申请号15/255,433中的相关分子。

[0145] 在一些实施方案中, 本申请的亲试剂或识别分子是降解途径蛋白。适合用作识别分子的降解途径蛋白的实例包括但不限于N端规则途径蛋白, 例如Arg/N端规则途径蛋白、Ac/N端规则途径蛋白和Pro/N端规则途径蛋白。在一些实施方案中, 识别分子是选自Gid4蛋白、Ubr1 UBR盒蛋白和ClpS蛋白 (例如, ClpS2) 的N端规则途径蛋白。

[0146] 肽酶, 也称为蛋白酶, 是催化肽键水解的酶。肽酶将多肽消化成较短的片段, 通常可分为内肽酶和外肽酶, 它们分别在内部和末端裂解多肽链。在一些实施方案中, 标记的亲试剂包含已被修饰以使外肽酶或内肽酶活性失活的肽酶。以这种方式, 标记的亲试剂选择性结合而不会裂解多肽中的氨基酸。在其他实施方案中, 可以使用未经修饰以使外肽酶或内肽酶活性失活的肽酶。例如, 在一些实施方案中, 标记的亲试剂包含标记的外肽酶。

[0147] 根据本申请的某些实施方案, 多肽测序方法可以包括在多肽末端的迭代检测和裂解。在一些实施方案中, 标记的外肽酶可以用作执行氨基酸检测和裂解这两个步骤的单一试剂。如一般性描述的, 在一些实施方案中, 标记的外肽酶具有氨肽酶或羧肽酶活性, 从而其分别选择性结合和裂解多肽的N末端或C末端氨基酸。应当理解, 在某些实施方案中, 标记的外肽酶可以由本领域技术人员催化失活, 使得标记的外肽酶保留选择性结合特性以用作非裂解标记的亲试剂, 如本文所述。

[0148] 外肽酶通常需要多肽底物在其氨基末端包含游离氨基或在其羧基末端包含游离羧基中的至少一个。在一些实施方案中, 根据本申请的外肽酶水解多肽末端处或附近的键。在一些实施方案中, 外肽酶水解距多肽末端不超过三个残基的键。例如, 在一些实施方案中, 由外肽酶催化的单个水解反应从多肽末端裂解单个氨基酸、二肽或三肽。

[0149] 在一些实施方案中,根据本申请的外肽酶是氨肽酶或羧肽酶,其分别从氨基末端或羧基末端裂解单个氨基酸。在一些实施方案中,根据本申请的外肽酶是二肽基-肽酶或肽基-二肽酶,其分别从氨基末端或羧基末端裂解二肽。在其他实施方案中,根据本申请的外肽酶是三肽基-肽酶,其从氨基末端裂解三肽。每个类别或其亚类的肽酶分类和活性是众所周知的,并在文献中进行了描述(参见,例如,Gurupriya,V.S.&Roy,S.C. *Proteases and Protease Inhibitors in Male Reproduction. Proteases in Physiology and Pathology* 195-216 (2017); 和Brix,K.&Stöcker,W. *Proteases: Structure and Function*. Chapter 1)。

[0150] 可以基于测序反应的方向性对根据本申请的外肽酶进行选择或工程化。例如,在从多肽的氨基末端到羧基末端测序的实施方案中,外肽酶包含氨肽酶活性。相反,在从多肽的羧基末端到氨基末端测序的实施方案中,外肽酶包含羧肽酶活性。识别特定羧基末端氨基酸的羧肽酶的实例,其可用作标记的外肽酶或被灭活以用作本文所述的非裂解标记的亲试剂,已在文献中进行了描述(参见,例如,Garcia-Guerrero,M.C.等人,(2018) PNAS 115 (17))。

[0151] 用作裂解试剂和/或亲和试剂(例如识别分子)的合适肽酶包括选择性结合一种或多种类型的氨基酸的氨肽酶。在一些实施方案中,氨肽酶识别分子被修饰以使氨肽酶活性失活。在一些实施方案中,氨肽酶裂解试剂是非特异性的,因此它从多肽的末端裂解大多数或所有类型的氨基酸。在一些实施方案中,与多肽末端的其他类型的氨基酸相比,氨肽酶裂解试剂在裂解多肽末端的一种或多种类型的氨基酸方面更有效。例如,根据本申请的氨肽酶特异性裂解丙氨酸、精氨酸、天冬酰胺、天冬氨酸、半胱氨酸、谷氨酰胺、谷氨酸、甘氨酸、组氨酸、异亮氨酸、亮氨酸、赖氨酸、甲硫氨酸、苯丙氨酸、脯氨酸、硒代半胱氨酸、丝氨酸、苏氨酸、色氨酸、酪氨酸和/或缬氨酸。在一些实施方案中,氨肽酶是脯氨酸氨肽酶。在一些实施方案中,氨肽酶是脯氨酸亚氨基肽酶。在一些实施方案中,氨肽酶是谷氨酸/天冬氨酸特异性氨肽酶。在一些实施方案中,氨肽酶是甲硫氨酸特异性氨肽酶。在一些实施方案中,氨肽酶是表1中列出的氨肽酶。在一些实施方案中,氨肽酶裂解试剂裂解表1中列出的肽底物。

[0152] 在一些实施方案中,氨肽酶是非特异性氨肽酶。在一些实施方案中,非特异性氨肽酶是锌金属蛋白酶。在一些实施方案中,非特异性氨肽酶是表2中列出的氨肽酶。在一些实施方案中,非特异性氨肽酶裂解表2中列出的肽底物。

[0153] 因此,在一些实施方案中,本申请提供了具有选自表1或表2的氨基酸序列(或具有与选自表1或表2的氨基酸序列具有至少50%、至少60%、至少70%、至少80%、80-90%、90-95%、95-99%或更高的氨基酸序列同一性的氨基酸序列)的氨肽酶(例如,氨肽酶识别分子、氨肽酶裂解试剂)。在一些实施方案中,氨肽酶与表1或表2中列出的氨肽酶具有25-50%、50-60%、60-70%、70-80%、80-90%、90-95%或95-99%或更高的氨基酸序列同一性。在一些实施方案中,氨肽酶是经修饰的氨肽酶,并且相对于表1或表2中列出的序列包括一个或多个氨基酸突变。

[0154] 表1. 氨肽酶的非限制性实例。

[0155]

名称	SEQ ID NO:	序列
嗜肺军团菌 (<i>L. pneumophila</i>) M1 氨肽酶 (Glu/Asp 特异性)	1	MGSSHHHHHSSGLVPRGSHMMVKQGVFMKTDQSKVKKLSDYKSLDYF VIHVDLQIDLSKKPVESKARLTVVPNLNVDSHSNDLVLDGENMTLVSLQM NDNLLKENEYELTKDSLIIKNIPQNTPTIEMTSLLGENTDLFGLYETEGVAL VKAESEGLRRVFYLPDRPDNLATYKTTIHANQEDYPVLLSNGVLIEKKELPL GLHSVTWLDDVPKPSYLFALVAGNLQRSVTYYQTKSGRELPIEFYVPPSAT SKCDFAKEVLKEAMAWDERTFNLECALRQH MVAGVDKYASGASEPTGLN LFNTENLFASPETKTDLGILRVLEVVAHEFFHYWSGDRVITIRDWFLPLKE GLTTFRAAMFREELFGTDLIRLLDGKNLDERAPRQSAYTAVRSLYTAAAYE KSADIFRMMMLFIGKEPFIEAVAKFFKDNDGGAVTLEDFIESISNSSGKDLR SFLSWFTESGIPELIVTDELNPDTKQYFLKIKTVNGRNRPIPLMGLLDSSGA EIVADKLLIVDQEEIEFQFENIQTRPIPSLLRSFSAPVHMKYEYSYQDLLLLM QFDTNLYNRCEAAKQLISALINDFCIGKKIELSPQFFAVYKALLSDNSLNEW MLAELITLPSLEELIENQDKPDFEKLNEGRQLIQNALANELKTD FYNLLFRI QISGDDDKQKLKGFDLQAGLRRLKSVCFSYLLNVDFEKTKEKLILQFED ALGKNMTETALALSMLCEINCEEADVALEDYYHYWKNDPGAVNNWFSIQ ALAHSPDVIERVKKLMRHGDFDLSNPNKVYALLGSFIKNPF GFH SVT GEG YQLVADAIFDLDKINPTLAANLTEKFTYWDKYDVNRQAMMISTLKHYSNA TSSDVRTMAKKGLDKVKEDLPLPIHLTFHGGSTMQDRTAQLIADGNKENA YQLH
大肠杆菌甲硫氨酸氨肽酶 (Met 特异性)	2	MAHHHHHHMGT AISIKTPEDIEKMRVAGRLAAEVLEMI EPYVKPGVSTGE LDRICNDYIVNEQHAVSACLG YHGYPK SVCISINEVVCHGIPDDAKLLKDG DIVNIDVTVIKDGFGHGDTSKMFIVGKPTIMGERLCRITQESLYLALRMVKP GINLREIGAAIQKFVEAEGFSVVREYCGHGIGRGFHEEPQVLHYDSRETNV VLKPGMTFTIEPMVNAGKKEIRTMKDGWTVKTKDRSLSAQYEHTIVVTD NGCEILTLRKDDTIPAIISHD

[0156]

耻垢分枝杆菌 (<i>M. smegmatis</i>) 脯氨酸亚氨基酸酶 (Pro 特异性)	3	MAHHHHHHMGTLEANTNGPGSMLSRMPVSSRTVPFGDHETWVQVTTPE NAQPHALPLIVLHGGPGMAHNYVANIAALADETGRTVIHYDQVGCNST HLPDAPADFWTPQLFVDEFHAVCTALGIERYHVLGQSWGGM LGAEIAVRQ PSGLVSLAICNSPASMRLWSEAAGDLRAQLPAETRAALDRHEAAGTITHPD YLQAAAEFYRRHVCRVVPTPQDFADSVAQMEAEPTVYHTMNGPNEFHVV GTLGDWSVIDRLPDVTAPVLVIAGEHDEATPKTWQPFVDHIPDVRSRVFPG TSHCTHLEKPEEFRAVVAQFLHQHDLAADARV
鼠疫耶尔森菌 (<i>Y. pestis</i>) 脯氨酸亚氨基酸酶 (Pro 特异性)	4	MTQQEYQNRQALLAKMAPGSAAIIFAAPEATRSADSEYPYRQNSDFSYL TGFNEPEAVLILVKSDETHNHSVLFNRIRDLTAEIWFGRRLGQEAAPTCLAV DRALPFDEINEQLYLLLNRLDVIYHAQQQYAYADNIVFAALEKLRHGFRCN LRAPATLTDWRPWLHEMRLFKSAEEIAVLRRAGEISALAHTRAMEKCRPG MFEYQLEGEILHEFTRHGARYPAYNTIVGGGENG CILHYTENECELRDGDL VLIDAGCEYRGYAGDITRTFPVNGKFTPAQRAVYDIVLAAINKSLTLFRPGT SIREVTEEVVRIMVVGVLVELGILKGDIEQLIAEQAHRPFFMHGLSHWLGM D VHDVG DYGSSDRGRILEPGMVLTVEPGLYIAPDADVPPQYRGIGIRIEDDIV ITATGNENLTASVVKDPDDIEALMALNHAGENLYFQEH HHHHHH
嗜热古菌 (<i>P. furiosus</i>) 甲硫氨酸氨肽酶	5	MDTEKLMKAGEIAKKVREKAIKLARPGMLLLELAESIEKMIMELGGKPAF PVNLSINEIAAHYTPYKGDTTVLKEGDY LKIDVGVHIDGFIADTAVTVRVG MEEDELMEAAKEALNAAISVARAGVEIKELGKAIENEIRKRGFKPIVNL SG HKIERYKLHAGISIPNIYRPHDNYVLKEGDVFAIEPFATIGAGQVIEVPPTLI YMYVRDVPVRVAQARFLAKIKREYGTLPFAYRWLQNDMPEGQLKLALK TLEKAGAIYGY PVLKEIRNGIVAQFEHTIIVEKDSVIVTQDMINKSTLE
温和气单胞菌 (<i>Aeromonas sobria</i>) 脯氨酸氨肽酶	6	HMSSPLHYVLDGIHCEPHFFT VPLDHQQPDDEETITLFGRTLCKDRLDDE LPWLLYLQGGPGFGAPRPSANGGWIKRALQEFRVLLDQRTGHSTPIHA ELLAHLNPRQQADYLSHFRADSIVRDAELIREQLSPDHPWSLLGQSFGGFC SLTYLSLFPDSLHEVYLTGGVAPIGRSADEVYRATYQRVADKNRAFFARFP HAQAIANRLATHLQRHDVRLPNGQRLTVEQLQQQGLDLGASGA FEELY LLEDAFIGEKLNPAFLYQVQAMQPFNTNPVFAILHELIYCEGAASHWAAER

[0157]

		VRGEFPALAWAQGKDFAF TGEMIFPWFMEQFRELIPK EAAHLLAEKADW GPLYDPVQLARNKVPVACAVYAEDMYVEFDYSRETLKGLSNSRAWITNEY EHNGLRVDGEQILDRLIRLNRDCLE
嗜 热 古 菌 (<i>Pyrococcus</i> <i>furiosus</i>) 脯 氨 酸 氨 肽 酶 (X-/-Pro)	7	MKERLEKLVKFM DENSIDRVFI AKPVNVYYFSGT SPLGGGYIIVDGDEATL YVPELEYEMAKEESKLPVVKFKKFDEIYEILKNTETLGIEGTLSSYSMVENF KEKSNVKEFKKIDDVIKDLRIIKTKEEIEIEKACEIADKAVMAAIEEITEGK REREVA AKVEYLMKMNGAEKPAFDTHIASGHR SALPHGVASDKRIERGDL VVIDLGALYNHYN SDITRTIVV GSPNEKQREIYEIVLEAQKRAVEAAKPGM TAKELDSIAREHKEYGYGDYFIHSLGHGVGLEIHEWPRISQYDET V LKEG MVITIEPGIYIPKLGGVRIEDTVLITENGAKRLTKTERELL
脑膜败血伊丽莎白 菌 (<i>Elizabethkingia</i> <i>meningoseptica</i>) 脯氨酸氨肽酶	8	MIPITTPVGNFKVWTKRFGTNP KIKVLLLHGGPAMTHEYMECFETFFQRE GFEFYEYDQLGSYYSDQPTDEKLWNIDRFVDEVEQVRKAIHADKENFYVL GNSWGGILAMEYALKYQQNLKGLIVANMMASAP EYVKYAEVLSKQMKP EVLAEVRAIEAKKDYANPRYTELLFPNYYAQHICRLKEWPDALNRSLKHV NSTVYTLMQGPSELGMSSDARLAKWDIKNRLHEIATPTLMIGARYDTMDP KAMEEQSKLVQKGRYLYCPNGSHLAMWDDQKV FMDGVIKFIKD VDTKS FN
温 和 气 单 胞 菌 (<i>Aeromonas sobria</i>) 脯氨酸氨肽酶	9	HMSSPLHYVLDGIHCEPHFFT VPLDHQQPDDEETITLFGRTLCKRDRLDDE LPWLLYLQGGPGFGAPRPSANGGWIKRALQEFRVLLLDQRG TGHTPIHA ELLAHLNPRQQADYLSHFRADSIVRDAELIREQLSPDHPWSLLGQSFGGFC SLTYLSLFPDSLHEVYLTGGVAPIGRSADEVYRATYQ RVADKNRAFFARFP HAQAIANRLATHLQRHDVRLPNGQRLTVEQLQQQGLDLGASGA FEELY LLEDAFIGEKLNP AFLYQVQAMQPFNTNPVFAILHELIYCEGAASHWAAER VRGEFPALAWAQGKDFAF TGEMIFPWFMEQFRELIPK EAAHLLAEKADW GPLYDPVQLARNKVPVACAVYAEDMYVEFDYSRETLKGLSNSRAWITNEY EHNGLRVDGEQILDRLIRLNRDCLE
淋病奈瑟菌 (<i>N.</i> <i>gonorrhoeae</i>)	10	MYEIKQPFHSGYLQVSEIHQIYWEESGNPDGVPVIFLHGGPGAGASPECRG FFNPDVFRIVIIDQRCGRSHPYACAEDNTTWDLVADIEKVREMLGIGKWL

[0158]

脯氨酸亚氨基酸酶		VFGGSWGSTLSLAYAQTHPERVKGLVLRGIFLCRPSETAWLNEAGGVSRİY PEQWQKFVAPIAENRRNRLIEAYHGLLFHQDEEVCLSAAKAWADWESYLI RFEPEGVDEDAYASLAIARLENHYFVNGGWLQGDKAILNNIGKIRHIPTVI VQGRYDLCTPMQSAWELSKAFPEAELRVVQAGHCAFDPPPLADALVQAVE DILPRLL
----------	--	---

[0159] 表2. 非特异性氨肽酶的非限制性实例

[0160]

名称	SEQ ID NO:	序列
大肠杆菌氨肽酶 N (锌金属蛋白酶)*	11	MGSSHHHHHHSSGENLYFQGHMTQQPQAKYRHDYRAPDYQITDIDLTDFD LDAQKTVVTAVSQAVRHGASDAPLRLNGEDLKLVSVINDEPWTAWKE EEGALVISNLPERFTLKIINEISPAANTALEGLYQSGDALCTQCEAEGFRHIT YYLDRPDVLARFTTKIADKIKYPFLLSNGNRVAQGELENGRHVVQWQD PFPKPCYLFALVAGDFDVLRTFTTRSGREVALELYVDRGNLDRAPWAM TSLKNSMKWDEERFGLEYDLDIYMIVAVDFFNMGAMENKGLNIFNSKYV LARTDTATDKDYLDIERVIGHEYFHNWTGNRVTCRDWFQLSLKEGLTVF RDQEFSSDLGSRVNRINNVRTMRGLQFAEDASPMAPHIRPDMVIEMNNF YTLTVYEKGAEVIRMIHTLLGEENFQKGMQLYFERHDGSAATCDDFVQA MEDASNVDLSHFRRWYSQSGTPIVTVKDDYNPETEQYTLTISQRTPATPD QAEKQPLHIPFAIELYDNEGKVIPLQKGGHPVNSVLNVTQAEQTFVFDNV YFQVPVALLCEFSAPVKLEYKWSQQQLTFLMRHARNDFSRWDAAQSLLA TYIKLNVARHQGGQPLSLPVHVADAFRAVLLDEKIDPALAAEILTLPSVNE MAELFDIIDPIAIAEVREALTRTLATELADELLAIYNANYQSEYRVEHEDIA KRTLNRNACLRFLAFGETHLADVLVSKQFHEANNMTDALAALSAVAAQL PCRDALMQEYDDKWHQNGLVMDKWFILQATSPAANVLETVRGLLQHRS FTMSNPNRIRSLIGAFAGSNPAAFHAEDGSGYLFLVEMLTDLNSRNPQVAS RLIEPLIRLKRYDAKRQEKMRAALEQLKGLNLSGDLYEKITKALA
恶性疟原虫(<i>P.</i>	12	PKIHYRKDYKPSGFIINQVTLNINIHDQETIVRSVLDMDISKHNVGEDLVFD

[0161]

<i>falciparum</i>) M1 氨 肽酶**		GVGLKINEISINNKKLVEGEEYTYDNEFLTIFSKFVPSKSKFAFSSEVIIHPET NYALTGLYKSKNIIVSQCEATGFRRTFFIDRPDMMAKYDVTVTADKEYY PVLLSNGDKVNEFEIPGGRHGARFNDPPLKPCYLFAVVAGDLKHLSATYI TKYTKKKVELYVFSEEKYVSKLQWALECLKKSMAFDEDEDYFGLEYDLSRL NLVAVSDFNVGAMENKGLNIFNANSLASKKNSIDFSYARILTVVGHEYF HQYTGNRVTLRDWFQLTLKEGLTVHRENLFSEEMTKTVTTRLSHVDLLR SVQFLEDSSPLSHPIRPESYVSMENFYTTTVYDKGSEVMRMYLTILGEEYY KKGFDIYIKKNDGNTATCEDFNYAMEQAYKMKKADNSANLNQYLLWFS QSGTPHVSFKYNYDAEKKQYSIHVNQYTKPDENQKEKKPLFIPISVGLINP ENGKEMISQTTLELTKESDTFVFNNIAVKPIPSLFRGFSAPVYIEDQLTDEE RILLKYDSDAFVRYSCTNIYMKQILMNYNEFLKAKNEKLESFQLTPVN AQFIDAIKYLLEDPHADAGFKSYIVSLPQDRYIINFVSNLDTDVLADTKEYI YKQIGDKLNDVYYKMFKSLEAKADDLTYNDESHVDFDQMNMRTLRNT LLSLLSKAQYPNILEIIEHSKSPYPSNWLTSLSVSAYFDKYFELYDKTYKL SKDDELLQEWLKTVSRSRDKDIYEILKKLENEVLKDSKNPNDIRAVYLPF TNNLRRFHDISGKGKYLIAEVITKTDKFNPMVATQLCEPFKLWNKLDTKR QELMLNEMNTMLQEPQISNNLKEYLLRLTNK
NPEPPS	13	MGSSHHHHHHSSGMWLAAPSLARRLLFLGPPPPPLLLLVSRSRRRL HSLGLAAMPEKRPFERLPADVSPINYSCLKPDLLDFTFEGKLEAAQVR QATNQIVMNCADIDIITASYAPEGDEEIHATGFNYQNEDEKVTLSPSTLQ TGTGTLKIDFVGELNDKMKGFYRSKYTPSGEVRYAAVTQFEATDARRA FPCWDEPAIKATFDISLVVPKDRVALSNMNVDRKPYPDENLVEVKFAR TPVMSTYLVAFFVGEYDFVETRSDGVCVRVYTPVGKAEQKFALEVA AKTLPFYKDYFNVPYPLPKIDLIAIADFAAGAMENWGLVITYRETALLIDPK NSCSSSRQWVALVVGHELAHQWFGNLVTMEWWTHLWLNEGFASWIEY LCVDHCFPEYDIWTQFVSADYTRAQELDALDNSHPIEVSVGHPSEVDEIFD AISYSKGASVIRMLHDYIGDKDFKKGMNMYLTKFQQKNAATEDLWESLE NASGKPIAAVMNTWTKQMGFPLIYVEAEQVEDDRLLRLSQKKFCAGGSY

[0162]

		VGEDCPQWMVPITISTSEDPNQAKLKILMDKPEMNVVLKNVKPDQWVKL NLGTVGIFYRTQYSSAMLESLLPGIRDLSLPPVDRLGLQNDLFLSARAGIIST VEVLKVMEAFVNEPNYTVWSDLSCNLGILSTLLSHTDFYEEIQEFVKDVFS PIGERLGWDPKPGEGHLDALLRGLVLGKLGKAGHKATLEEARRRFKDHV EGKQILSADLRSPVYLTVLKHGDGTTLDIMLKLHKQADMQEEKNRIERVL GATLLPDLIQKVLTFALSEEVRPQDTVSVIGGVAGGSKHGRKAAWKFIKD NWEELYNRYQGGFLISRLIKLSVEGFAVDKMAGEVKAFFESHAPSARTI QQCCENILLNAAWLKRDAESIHQYLLQRKASPPTV
NPEPPS E366V	14	MGSSHHHHHHSSGMWLAAPSLARRLLFLGPPPPPLLLLVSRRSSRRRL HSLGLAAMPEKRPFERLPADVSPINYSCLKPDLLDFTFEGKLEAAAQVR QATNQIVMNCADIDIITASYAPEGDEEIHATGFNYQNEDEKVTLSFPSTLQ TGTGTLKIDFVGELNDKMKGFYRSKYTTPSGEVRYAAVTQFEATDARRA FPCWDEPAIKATFDISLVVPKDRVALSNMNVDRKPYPDENLVEVKFAR TPVMSTYLVAFFVGEYDFVETRSGDGVCRVYTPVGKAEQKGFALEVA AKTLFPYKDYFNVPYPLPKIDLIAIADFAAGAMENWGLVITYRETALLIDPK NSCSSSRQWVALVVGHVLAHQWFGNLVTMEWWTHLWLNFGFASWIEY LCVDHCFPEYDIWTQFVSADYTRAQELDALDNHPIEVSVGHPSEVDEIFD AISYSGGASVIRMLHDYIGDKDFKKGMNMYLTKFQQKNAATEDLWESLE NASGKPIAAVMNTWTQKMGFPLIYVEAEQVEDRLLRLSQKKFCAGGSY VGEDCPQWMVPITISTSEDPNQAKLKILMDKPEMNVVLKNVKPDQWVKL NLGTVGIFYRTQYSSAMLESLLPGIRDLSLPPVDRLGLQNDLFLSARAGIIST VEVLKVMEAFVNEPNYTVWSDLSCNLGILSTLLSHTDFYEEIQEFVKDVFS PIGERLGWDPKPGEGHLDALLRGLVLGKLGKAGHKATLEEARRRFKDHV EGKQILSADLRSPVYLTVLKHGDGTTLDIMLKLHKQADMQEEKNRIERVL GATLLPDLIQKVLTFALSEEVRPQDTVSVIGGVAGGSKHGRKAAWKFIKD NWEELYNRYQGGFLISRLIKLSVEGFAVDKMAGEVKAFFESHAPSARTI QQCCENILLNAAWLKRDAESIHQYLLQRKASPPTV
土拉弗朗西斯菌	15	MIYEFVMTDPKIKYLKDYKPSNYLIDETHLIFELDESKTRVTANLYIVANR

[0163]

<i>(Francisella tularensis)</i> 氨肽酶 N		ENRENN TLVLDGVELKLLSIKLNKHLSPA EFAVNENQLIINN VPEKFVLQ TVVEINPSANTSLEGLYKSGDV FSTQCEATGFRKITYYLD RPDVMAAFTV KIIADKKKYPIILSNGDKIDSGDISDNQHFA VWKDPFKKPCYLFALVAGDL ASIKDTYITKSQRKVSLEIYAFKQDIDKCHYAMQAVKDSMKWDEDRFGL EYDLDTFMIVAVPDFNAGAMENKGLNIFNTKYIMASNKTATDKDFELVQ SVVGHEYFHNWTGDRVTCRDWFQLSLKEGLTVFRDQEFTSDLSNRDVKR IDDVRIIRSAQFAEDASPM SHPIRPESYIEMNNFYTVTVYNKGAEIIRMIHTL LGEEGFQKGMKLYFERHDGQAVTCDDFVNAMADANNRDFSLFKRWYA QSGTPNIKVSENYDASSQTYSLTLEQTTLPTADQKEKQALHIPVKMGLINP EGKNIAEQVIELKEQKQTYTFENIAAKPVASLFRDFSAPVKVEHKRSEKDL LHIVKYDNNAFNRWDSLQQIATNIILNNADLNDEFNFAFKSILHDKDLDK ALISNALLIPIESTIAEAMRVIMVDDIVLSRKNVVNQLADKLDKDDWLAVY QQCNDNKPYSLSAEQIAKRKLKGVCLSYLMNASDQKVGTDLAQQLFDN ADNMTDQQTAFTELLKSNDKQVRDNAIN EFYNRWRHEDLVVNKWLLSQ AQISHESALDIVKGLVNHPAYNPKNPNKVYSLIGGFGANFLQYHCKDGLG YAFMADTVLALDKFNHQVAARMARNLMSWKRYDSDRQAMMKNALEKI KASNPSKNVFEIVSKSLES
超嗜热古菌 <i>(Pyrococcus horikoshii)</i> TET 氨肽酶	16	MGSSHHHHHSSGMEVRNMVDYELLKKVVEAPGVSGYEFLGIRDVVIIE IKDYVDEVKVDKLG NVIAHKKGEGPKVMIAAHMDQIGLMVTHIEKNGFL RVAPIGGVDPKTLIAQRFKVWIDKGKFIYGV GASVPPHIQKPEDRKKAPD WDQIFIDIGAESKEEAEDMGVKIGTVITWDGR LERLGKHRFVSIAFDDRIA VYTILEVAKQLKDAKADVVFVATVQEEVGLRGARTSAFGIEPDYGF AIDV TIAADIPGTPEHKQVTHLGKGTAIKIMDRSVICHPTIVRWLEELAKKHEIPY QLEILLGGGTDAGAIHLTKAGVPTGALSVPARYIHSNTEVVDERDV DATV ELMTKALENIHELKI
嗜热水生菌(<i>T. aquaticus</i>) 氨肽酶 T	17	MDAFTENLNKLAELAIRVGLNLEEGQEIVATAPIEA VDFVRLLAEKAYEN GASLFTVLYGDNLIARKRLALVPEAHLDRAPAWLYEGMAKAFHEGAARL AVSGNDPKALEGLPPERVGRAQQAQSRAYRPTLSAITEFVTNWTIVPFAH

[0164]

		PGWAKAVFPGLPEEEAVQRLWQAIFQATRVDQEDPVA AWEAHNRVLHA KVAFLNEKRFBALHFQGPGLTDLTVGLAEGHLWQGGATPTKKGRLCNPNL PTEEVFTAPHRERVEGVVRASRPLALSGQLVEGLWARFEGGVAVEVGAE KGEEVLKKLLDTDEGARRLGEVALVPADNPIAKTGLVFFDTLFDENAASH IAFGQAYAENLEGRPSGEEFRRRGGNESMVHVDWMIGSEEVVDVGLLED GTRVPLMRRGRWVI
嗜热脂肪芽孢杆菌 (<i>Bacillus stearothermophilus</i>) 肽酶 M28	18	MAKLDETLTMLKALTDAGKVPGNEREARDVMKTYIAPYADEVTTDGLG SLIAKKEGKSGGPKVMIAGHLDEVGFMVTQIDDKGFIRFQTLGGWWSQV MLAQRVTIVTKGDITGVIGSKPPHILPSEARKKPVEIKDMFIDIGATSREE AMEWGVPRPGDMIVPYFEFTVLNNEKMLLAKAWDNRIGCAVAIDVLKQL KGVDPNTVYGVGTVQEEVGLRGARTAAQFIQPDIAFAVDVGIAGDTPG VSEKEAMGKLGAGPHIVLYDATMVSHRGLREFVIEVAEELNIPHHFDAMP GVGTDAGAIHLTGIGVPSLTIAIPTRYIHSHAAILHRDDYENTVKLLVEVIK RLDADKVKQLTFDE
霍乱弧菌(<i>Vibrio cholera</i>)氨肽酶	19	MEDKVVISMGADAVGSLNPALSESLLPHSFASGSQVWIGEVAIDELAELS HTMHEQHNRCGGYMVHTSAQGAMAALMPESIANFTIPAPSQQDLVNA WLPQVSADQITNTIRALSSFNNRFYTTTSGAQASDWLANEWRSLISSLPGS RIEQIKHSGYNQKSVVLTIQGSEKPDEWVIVGGHLDSTLGSHTNEQSIAPG ADDDASGIASLSEIIRVLRDNNFRPKRSVALMAYAAEEVGLRGSQDLANQ YKAQGKKVSVLQLDMTNYRGS AEDIVFITDYTDSNLTQFLTTLIDEYLP ELTYGYDRCGYACSDHASWHKAGFSAAMPFESKFKDYNPKIHTSQDTLA NSDPTGNHAVKFTKLGLAYVIEMANAGSSQVPDDSVLQDGTAKINLSGA RGTQKRFTFELS QSKPLTIQTYGGSGD VDLVVKYGSAPSKSNWDCRPYQN GNRETCSFNNAQPGIYHVMLDGYTNYNDVALKASTQH HHHHHH
<i>Photobacterium halotolerans</i> 氨肽酶 j	20	MEDKVVISIGSDASQTVKSVMQSNARSLLPESLASNGPVWVGQVDYSQL AELSHMHEDHQRCGGYMVHSPESAIAASNMPQSLVAFSIPEISQQDTV NAWLPQVNSQAITGTITSLTSFINRFYTTTSGAQASDWLANEWRSL SASLP NASVRQVSHFGYNQKSVVLTITGSEKPDEWIVLGGHLDSTIGSHTNEQSV

[0165]

		APGADDDASGIASVTEIIRVLENNFQPKRSIAFMAYAAEEVGLRGSQDLA NQYKAEGKQVISALQLDMTNYKGSVEDIVFITDYTDSNLTTFLSQLVDEY LPSLTYGFDTCGYACSDHASWHKAGFSAAMPFEAKFNDYNPMIHTPNDT LQNSDPTASHAVKFTKLGLAYAIEMASTTGGTPPPTGNVLKDGVPVNGLS GATGSQVHYSFELPAQKNLQISTAGGSGDVLDLYVSFGSEATKQNWDCRP YRNGNNEVCTFAGATPGTYSIMLDGYRQFSGVTLKASTQHHHHHH
鼠疫耶尔森菌 (<i>Yersinia pestis</i>) 氨肽酶 N	21	MTQQPQAKYRHDYRAPDYTITDIDLDFALDAQKTTVTAVSKVKRQGTDV TPLILNGEDLTLSVSVSDGQAWPHYRQQDNTLVIEQLPADFTLTIVNDIHPA TNSALEGLYLSGEALCTQCEAEGFRHITYYLDPRPDVLARFTTRIVADKSRY PYLLSNGNRVGGQELDDGRHWVKWEDPFPKPSYLFALVAGDFDVLQDK FITRSGREVALEIFVDRGNLDRADWAMTSLKNSMKWDETRFGLEYDLDI YMIVAVDFFNMGAMENKGLNVFNSKYVLAKAETATDKDYLNIEAVIGHE YFHNWTGNRVTCRDWFQLSLKEGLTVFRDQEFSSDLGSRSVNRIENVRV MRAAQFAEDASPMMAHAIRPDKVIEMNNFYTLTVYEKGSEVIRMMHTLLG EQQFQAGMRLYFERHDGSAATCDDFVQAMEDVSNVDLSLFRRWYSQSG TPLLTVHDDYDVEKQQYHLFVSQKTLPTADQPEKLPLHIPLDIELYDSKGN VIPLQHNGLPVHHVLNVTEAEQTFTFDNVAQKPIPSLLREFSAPVKLDYPY SDQQLTFLMQHARNEFSRWDAQAQSLATYIKLNVAKYQQQQPLSLPAHV ADAFRAILLDEHLDPALAAQILTLSENEMAELFTTIDPQAISTVHEAITRC LAQELSDELLAVYVANMTPVYRIEHGDIKRALRNTCLNYLAFGDEEFAN KLVSLQYHQADNMTDSLAAALAAVAAQLPCRDELLAAFDVVRWNHDGL VMDKWFALQATSPAANVLVQVRTLLKHPAFSLSNPNRTRSLIGSFASGNP AAFHAADGSGYQFLVEILSDLNTRNPQVAARLIEPLIRLKRYDAGRQALM RKALEQLKTLDNLSGDLYEKITKALAAHHHHHH
鳃弧菌(<i>Vibrio anguillarum</i>)氨肽酶	22	MEEKVWISIGDATQTALRSGAQSLLENLINQTSVWVGQVPVSELATLS HEMHENHQRCGGYMVHPSAQSAMSVSAMPLNLNAFSAPEITQTTVNA WLPSVSAQQITSTITTLTQFKNRFYTTSTGAQASNWIADHWRSLSASLPAS KVEQITHSGYNQKSVMLTITGSEKPDEWVVIGGHLDSLGSRTNESSIAPG

[0166]

		ADDDASGIAGVTEIIRLLSEQNFRPKRSIAFMAYAAEEVGLRGSQDLANRF KAEGKKVMSVMQLDMTNYQGSREDIVFITDYTDSNFTQYLTQLLDEYLP SLTYGFDTCGYACSDHASWHAVGYPAAMPFESKFNDYNPNIHSPQDTLQ NSDPTGFHAVKFTKLGLAYVVEMGNASTPPTPSNQLKNGVPVNGLSASR NSKTWYQFELQEAGNLSIVLSGGSGDADLYVKYQTDADLQQYDCRPYRS GNNETCQFSNAQPGRYSILLHGYNNYSNASLVANAQHHHHHH
盐弧菌 <i>sp</i> YCSC6 (<i>Salinivibrio</i> <i>sp</i> YCSC6) 氨肽酶	23	MEDKKVWISIGADAQQTALSSGAQPLLAQSVAHNGQAWIGEVSESELA LSHEMHENHHRCGGYIVHSSAQSAAMAASN MPLSRASFIAPISQQALVTP WISQIDSALIVNTIDRLTDFPNRFYTTTSGAQASDWIKQRWQSLAGLAGA SVTQISHSGYNQASVMLTIEGSESPDEWVVVGGHLDSTIGSRTNEQSIAPG ADDDASGIAAVTEVIRVLAQNNFQPKRSIAFVAYAAEEVGLRGSQDVAN QFKQAGKDVRGVLQLDMTNYQGS AEDIVFITDYTDNQLTQYLTQLLDEY LPTLNYGFDTCGYACSDHASWHQVGYPAMPFEAKFNDYNPNIHPTQDT LANSDEGAHA AKFTKLGLAYTVELANADSSPNPGNELKLGEPIGLSGA RGNEKYFNYRLDQSGELVIRTYGGSGD VDLVYKANGDVSTGNWDCRPY RSGNDEVCRFDNATPGNYAVMLRGYRTYDNVSLIVEHHHHHH
解蛋白弧菌(<i>Vibrio</i> <i>proteolyticus</i>) 氨肽 酶 I	24	<u>G</u> MPPITQQATVTAWLPQVDASQITGTISSLESFTNRFYTTTSGAQASDWIA SEWQ <u>F</u> LSASLPNASVKQVSHSGYNQKSVMTITGSEAPDEWIVIGGHLD TIGSHTNEQSVAPGADDDASGIAAVTEVIRVLSENNFPKRSIAFMAYAAE EVGLRGSQDLANQYKSEGNVVSALQLDMTNYKGS AQDVVFITDYTDS NFTQYLTQLMDEYLP SLTYGFDTCGYACSDHASWHNAGYPAMPFESKF NDYNPRIHTTQDTLANS DPTGSHAKKFTQLGLAYA IEMGSATGDTPTPGN QLEHHHHHH
嗜热古菌(<i>P.</i> <i>furius</i>) 氨肽酶 I	25	MVDWELMKKIIESPGVSGYEHLGIRDLVVDILKDVADEVKIDKLG NVIAH FKGSAPKVMVA AHMDKIGLMVNHIDKDG YLRVVP IGGVLPETLIAQKIRF FTEKGERYGVVGVLPPHLRREAKDQGGKIDWDSIIVDVGASSREEAEEMG FRIGTIGEFAPNFTRLSEHRFATPYLDDRICLYAMIEAARQLGEHEAD IYIV

[0167]

ASVQEEIGLRGARVASFAIDPEVGIAMDVTFQKQPNKKGKIVPELGKGPV
MDVGPNNPKLRQFADEVAKKYEIPLQVEPSRPTGTDANVMQINREGVA
TAVLSIPRYMHSQVELADARDVDNTIKLAKALLEELKPMDFTPLEHHHH
HH

[0168] *裂解效率(从最高到最低):精氨酸>赖氨酸>疏水残基(包括丙氨酸、亮氨酸、甲硫氨酸和苯丙氨酸)>脯氨酸(参见,例如,Matthews Biochemistry 47,2008,5303-5311)。

[0169] **裂解效率(从最高到最低):亮氨酸>丙氨酸>精氨酸>苯丙氨酸>脯氨酸;在谷氨酸和天冬氨酸后不裂解。

[0170] 出于比较两个或更多个氨基酸序列的目的,第一氨基酸序列和第二氨基酸序列之间的“序列同一性”(本文也称为“氨基酸同一性”)百分比可以通过[第一氨基酸序列中与第二氨基酸序列中相应位置的氨基酸残基相同的氨基酸残基的数量]除以[第一氨基酸序列中氨基酸残基的总数]并乘以[100]来计算,其中与第一氨基酸序列相比,第二氨基酸序列中氨基酸残基的每个缺失、插入、取代或添加被认为是单个氨基酸残基(位置)的差异。替代地,可以使用已知的计算机算法(例如,通过Smith和Waterman(1970) Adv. Appl. Math. 2: 482c的局部同源算法,通过Needleman和Wunsch, J. Mol. Biol. (1970) 48:443的同源比对算法,通过Pearson和Lipman. Proc. Natl. Acad. Sci. USA (1998) 85:2444的相似性搜索方法,或通过可作为Blast、Clustal Omega或其他序列比对算法的计算机化实现算法),例如使用标准设置,来计算两个氨基酸序列之间的序列同一性程度。通常,出于根据上文概述的计算方法确定两个氨基酸序列之间的“序列同一性”百分比的目的,将氨基酸残基数最多的氨基酸序列作为“第一”氨基酸序列,另一个氨基酸序列将作为“第二”氨基酸序列。

[0171] 另外地或替代地,可以评估两个或更多个序列的序列之间的同一性。在两个或更多个核酸或氨基酸序列的上下文中,术语“同一的”或百分比“同一性”是指两个或更多个相同的序列或子序列。当在比较窗口或使用上述序列比较算法之一或通过人工比对和目视检查测量的指定区域上进行比较和比对时,如果两个序列在指定区域或整个序列上具有指定百分比的相同氨基酸残基或核苷酸(例如,至少80%、85%、90%、95%、96%、97%、98%、99%、99.5%、99.6%、99.7%、99.8%或99.9%同一的),则两个序列是“基本上同一的”。任选地,同一性存在于长度为至少约25、50、75或100个氨基酸的区域上,或存在于长度为100至150、150至200、100至200、或200或更多个氨基酸的区域上。

[0172] 另外地或替代地,可以评估两个或更多个序列的序列之间的比对。在两个或更多个核酸或氨基酸序列的上下文中,术语“比对”或百分比“比对”是指两个或更多个相同的序列或子序列。当在比较窗口或使用上述序列比较算法之一或通过人工比对和目视检查测量的指定区域上进行比较和比对时,如果两个序列在指定区域或整个序列上具有指定百分比的相同氨基酸残基或核苷酸(例如,至少80%、85%、90%、95%、96%、97%、98%、99%、99.5%、99.6%、99.7%、99.8%或99.9%同一的),则两个序列是“基本上比对上的”。任选地,比对存在于长度为至少约25、50、75或100个氨基酸的区域上,或存在于长度为100至150、150至200、100至200、或200或更多个氨基酸的区域上。

[0173] 除了多肽分子之外,核酸分子还具有多种有利的特性,可根据本申请用作亲和试

剂(例如氨基酸识别分子)。

[0174] 核酸适体是经过工程化以高亲和力和选择性结合所需靶标的核酸分子。因此,可以使用本领域已知的选择和/或富集技术对核酸适体进行工程化以选择性结合所需类型的氨基酸。因此,在一些实施方案中,亲和试剂包含核酸适体(例如,DNA适体、RNA适体)。在一些实施方案中,标记的亲和试剂是选择性结合一种类型的末端氨基酸的标记的适体。例如,在一些实施方案中,如本文所述,标记的适体在多肽末端选择性结合一种类型的氨基酸(例如,单一类型的氨基酸或氨基酸类型的子集)。尽管未显示,但应当理解,标记的适体可以被工程化以根据本申请的方法在多肽的任何位置(例如,在多肽的末端位置或末端和内部位置)选择性结合一种类型的氨基酸。

[0175] 在一些实施方案中,标记的亲和试剂包含具有结合诱导的发光的标记。例如,在一些实施方案中,标记的适体包含供体标记和受体标记以及功能。在其他实施方案中,标记的适体包含淬灭部分并且功能类似于分子信标,其中标记的适体的发光在内部被淬灭为游离分子并恢复为选择性结合的分子(参见例如Hamaguchi等人,(2001) *Analytical Biochemistry* 294,126-131)。不希望囿于理论的,认为用于结合诱导的发光的这些和其他类型的机制可以有利地减少或消除背景发光以提高本文所述的方法的总体灵敏度和准确度。

[0176] 除了鉴定多肽末端氨基酸的方法之外,本申请还提供了使用标记的亲和试剂对多肽进行测序的方法。在一些实施方案中,测序方法可以涉及使多肽末端经受末端氨基酸检测和末端氨基酸裂解的重复循环。例如,在一些实施方案中,本申请提供了一种确定多肽的氨基酸序列的方法,所述方法包括使多肽与本文所述的一种或多种标记的亲和试剂接触并使多肽经受Edman降解。

[0177] 常规Edman降解涉及修饰和裂解多肽末端氨基酸的重复循环,其中鉴定每个连续裂解的氨基酸以确定多肽的氨基酸序列。作为常规Edman降解的说明性实例,多肽的N端氨基酸使用异硫氰酸苯酯(PITC)修饰以形成PITC衍生的N末端氨基酸。然后使用酸性条件、碱性条件和/或高温裂解PITC衍生的N末端氨基酸。还表明,裂解PITC衍生的N末端氨基酸的步骤可以使用来自原生动物克氏锥虫(*Trypanosoma cruzi*)的经修饰的半胱氨酸蛋白酶通过酶促完成,这涉及在中性或接近中性pH下相对温和的裂解条件。有用的酶的非限制性实例描述于2016年9月2日提交的标题为“MOLECULES AND METHODS FOR ITERATIVE POLYPEPTIDE ANALYSIS AND PROCESSING”的美国专利申请号15/255,433中。

[0178] 在一些实施方案中,通过Edman降解进行测序包括提供通过接头固定在固体支持物表面(例如,固定在样品孔的底部或侧壁表面)上的多肽。在一些实施方案中,如本文所述,多肽被固定在一个末端(例如氨基末端氨基酸或羧基末端氨基酸),使得另一末端是游离的,用于检测和裂解末端氨基酸。因此,在一些实施方案中,本文所述的在Edman降解方法中使用的试剂优先与多肽的非固定(例如,游离)末端处的末端氨基酸相互作用。以这种方式,多肽在检测和裂解的重复循环中保持固定。为此,在一些实施方案中,可以根据用于检测和裂解的所需条件组设计接头,例如,以限制多肽在化学裂解条件下从表面脱离。用于将多肽固定在表面的合适的接头组合物和技术在本文别处详细描述。

[0179] 根据本申请,在一些实施方案中,通过Edman降解进行测序的方法包括步骤(i),使多肽与一种或多种选择性结合一种或多种类型的末端氨基酸的标记的亲和试剂接触。在一

些实施方案中,标记的亲试剂通过选择性结合末端氨基酸而与多肽相互作用。在一些实施方案中,步骤(i)进一步包括去除未选择性结合多肽的末端氨基酸(例如,游离末端氨基酸)的一种或多种标记的亲试剂中的任一种。

[0180] 在一些实施方案中,所述方法进一步包括通过检测标记的亲试剂来鉴定多肽的末端氨基酸。在一些实施方案中,检测包括检测来自标记的亲试剂的发光。如本文所述,在一些实施方案中,发光与标记的亲试剂独特地关联,因此发光与标记的亲试剂选择性结合的氨基酸类型相关。因此,在一些实施方案中,通过确定标记的亲试剂的一种或多种发光特性来鉴定氨基酸的类型。

[0181] 在一些实施方案中,通过Edman降解进行测序的方法包括步骤(ii),去除多肽的末端氨基酸。在一些实施方案中,步骤(ii)包括从多肽去除标记的亲试剂(例如,一种或多种选择性结合末端氨基酸的标记的亲试剂中的任一种)。在一些实施方案中,步骤(ii)包括通过使末端氨基酸与异硫氰酸酯(例如,PITC)接触以形成异硫氰酸酯修饰的末端氨基酸来修饰多肽的末端氨基酸(例如,游离末端氨基酸)。在一些实施方案中,异硫氰酸酯修饰的末端氨基酸比未经修饰的末端氨基酸更容易被裂解试剂(例如化学或酶促裂解试剂)去除。

[0182] 在一些实施方案中,步骤(ii)包括通过使多肽与特异性结合和裂解异硫氰酸酯修饰的末端氨基酸的蛋白酶接触来去除末端氨基酸。在一些实施方案中,蛋白酶包括修饰的半胱氨酸蛋白酶。在一些实施方案中,蛋白酶包括修饰的半胱氨酸蛋白酶,例如来自克氏锥虫的半胱氨酸蛋白酶(参见例如,Borgo等人,(2015)Protein Science 24:571-579)。在其他实施方案中,步骤(ii)包括通过使多肽经受足以裂解异硫氰酸酯修饰的末端氨基酸的化学(例如酸性、碱性)条件来去除末端氨基酸。

[0183] 在一些实施方案中,通过Edman降解进行测序的方法包括步骤(iii),在末端氨基酸裂解后洗涤多肽。在一些实施方案中,洗涤包括去除蛋白酶。在一些实施方案中,洗涤包括将多肽恢复至中性pH条件(例如,在通过酸性或碱性条件进行化学裂解之后)。在一些实施方案中,通过Edman降解进行测序的方法包括重复步骤(i)至(iii)多个循环。

[0184] 在一些实施方案中,含有多肽的复杂混合物或富集混合物(例如,多肽混合物)的样品可以使用普通酶降解成约6至40个氨基酸的短多肽片段。在一些实施方案中,根据本申请的方法对该多肽文库进行测序将揭示存在于原始复杂混合物或富集混合物中的每种多肽的身份和丰度。如本文和文献中所述,大小范围为6至40个氨基酸的大多数多肽可以通过确定多肽链中仅四个氨基酸的数量和位置来独特地鉴定。

[0185] 因此,在一些实施方案中,通过Edman降解进行测序的方法可以使用一组标记的适体进行,该组标记的适体包括四种DNA适体类型,每种类型识别不同的N末端氨基酸。每种适体类型可以用不同的发光标记进行标记,从而可以基于一种或多种发光特性来区分不同的适体类型。为了说明性目的,标记的适体的实例组包括:用第一发光标记(“染料1”)标记的半胱氨酸特异性适体;用第二发光标记(“染料2”)标记的赖氨酸特异性适体;用第三发光标记(“染料3”)标记的色氨酸特异性适体;以及用第四发光标记(“染料4”)标记的谷氨酸特异性适体。

[0186] 在一些实施方案中,在步骤(i)之前,将来自多肽文库的单个多肽分子固定在固体支持物的表面,例如样品孔阵列的样品孔的底部或侧壁表面。在一些实施方案中,如本文别处所述,能够实现表面固定的部分(例如,生物素)或提高溶解度的部分(例如,寡核苷酸)可

以化学或酶促连接至多肽的C末端。为了确定每个多肽的序列,在一些实施方案中,固定化的多肽经受N末端氨基酸检测和N末端氨基酸裂解的重复循环。在一些实施方案中,所述方法包括试剂添加和洗涤步骤,这些步骤通过使用自动化流体系统注入到检测表面上方的流通池中进行。在一些实施方案中,步骤(i)至(iv)说明了使用标记的适体进行检测和裂解的一个循环。

[0187] 在一些实施方案中,通过Edman降解进行测序的方法包括步骤(i),流入四种正交标记的DNA适体的混合物中并孵育以使适体与任何固定化的多肽(例如,固定在阵列的样品孔内的多肽)结合,所述在N末端包含四个正确氨基酸中的一个。在一些实施方案中,所述方法进一步包括洗涤固定化的多肽以去除未结合的适体。在一些实施方案中,所述方法进一步包括对固定化的多肽进行成像(“成像步骤(i)”)。在一些实施方案中,所获得的图像包含足够的信息来确定与适体结合的多肽的位置(例如,在样品孔阵列内的位置)以及四个适体中的哪一个在每个位置处被结合。在一些实施方案中,所述方法进一步包括使用合适的缓冲液来洗涤固定化的多肽以从固定化的多肽中去除适体。

[0188] 在一些实施方案中,测序方法包括步骤(ii)在含有特异性修饰N末端胺基的反应性分子(例如,PITC,如图所示)的溶液中流动。在一些实施方案中,诸如PITC的异硫氰酸酯分子将N末端氨基酸修饰成底物,用于通过经修饰的蛋白酶例如来自克氏锥虫(*Trypanosoma Cruzi*)的半胱氨酸蛋白酶cruzain的裂解。

[0189] 在一些实施方案中,测序方法包括步骤(iii),洗涤固定化的多肽,再流入合适的经修饰的蛋白酶,所述经修饰的蛋白酶识别并裂解来自固定化的多肽的修饰的N末端氨基酸。

[0190] 在一些实施方案中,该方法包括在酶促裂解后洗涤固定化多肽的步骤(iv)。在一些实施方案中,步骤(i)至(iv)描绘了Edman降解的一个循环。因此,所示的步骤(i')是下一个反应循环的开始,该反应循环作为如上文对步骤(i)至(iv)所述进行的步骤(i')至(iv')进行。在一些实施方案中,步骤(i)至(iv)重复大约20-40个循环。

[0191] 在一些实施方案中,标记的异硫氰酸酯(例如,染料标记的PITC)可用于监测样品上样。例如,在一些实施方案中,在对多肽样品进行测序方法之前,通过使用染料标记的PITC修饰末端,使多肽样品在末端与发光标记预缀合。以这种方式,可以通过在上述步骤(i)之前检测来自标记的发光来监测多肽样品向样品孔阵列中的上样。在一些实施方案中,发光用于确定阵列中样品孔的单个占据(例如,含有单个多肽分子的样品孔的一部分),这可以有利地增加对于给定样品可靠地获得的信息量。在进行步骤(i)之前,一旦通过发光确定了所需的样品上样状态,就可以如所描述的进行化学或酶促裂解。

[0192] 在一些实施方案中,标记的异硫氰酸酯(例如,染料标记的PITC)可用于监测阵列中多肽样品的反应进程。例如,在一些实施方案中,步骤(ii)包括流入含有染料标记的PITC的溶液,所述PITC特异性地修饰和标记样品中多肽的N末端胺基。在一些实施方案中,可以在步骤(ii)期间或之后检测来自标记的发光以评估样品中多肽的N末端PITC修饰。因此,在一些实施方案中,发光用于确定是否或何时从步骤(ii)进行到步骤(iii)。在一些实施方案中,可以在步骤(iii)期间或之后检测来自标记的发光以评估样品中多肽的N末端氨基酸裂解——例如以确定是否或何时从步骤(iii)进行到步骤(iv)。

[0193] 测序方法可以使用单独的试剂来检测和裂解多肽的末端氨基酸。尽管如此,在一

些方面,本申请提供了一种测序方法,其中包含肽酶(例如选择性结合和裂解不同类型的末端氨基酸的标记的外肽酶)的单一试剂可用于检测和裂解多肽的末端氨基酸。

[0194] 标记的外肽酶可以包括包含第一发光标记的赖氨酸特异性的外肽酶、包含第二发光标记的甘氨酸特异性的外肽酶、包含第三发光标记的天冬氨酸特异性的外肽酶和包含第四发光标记的亮氨酸特异性的外肽酶。根据本文所述的某些实施方案,每个标记的外肽酶仅在其相应氨基酸位于多肽的氨基末端或羧基末端时选择性结合和裂解该氨基酸。因此,随着通过这种方法的测序从肽的一个末端向另一个末端进行,标记的外肽酶被工程化或选择,使得该组的所有试剂将具有氨肽酶或羧肽酶活性。

[0195] 在一些方面,本申请通过评估末端氨基酸与标记的氨基酸识别分子(例如,标记的亲试剂)和标记的裂解试剂(例如,标记的非特异性外肽酶)的结合相互作用来提供实时多肽测序的方法。不希望囿于理论的,标记的亲试剂根据由结合率或结合的“开”率(k_{on})和解离率或结合的“关”率(k_{off})定义的结合亲和性(K_D)选择性结合。速率常数 k_{off} 和 k_{on} 分别是脉冲持续时间(例如,对应于可检测结合事件的时间)和脉冲间持续时间(例如,可检测结合事件之间的时间)的关键决定因素。在一些实施方案中,可以设计这些速率以实现给出最佳测序准确度的脉冲持续时间和脉冲频率(例如,信号脉冲的频率)。

[0196] 测序反应混合物可以进一步包含标记的非特异性外肽酶,其包含不同于标记的亲试剂的发光标记。在一些实施方案中,标记的非特异性外肽酶以低于标记的亲试剂的浓度存在于混合物中。在一些实施方案中,标记的非特异性外肽酶显示出广泛的特异性,使得它裂解大多数或所有类型的末端氨基酸。

[0197] 在一些实施方案中,末端氨基酸被标记的非特异性外肽酶裂解产生信号脉冲,并且这些事件以比标记的亲试剂的结合脉冲更低的频率发生。以这种方式,可以在实时测序过程中对多肽的氨基酸进行计数和/或鉴定。在一些实施方案中,可以使用多种标记的亲试剂,每种都具有可用于鉴定相应末端氨基酸的诊断脉冲模式(例如,特征模式)。例如,在一些实施方案中,不同的特征模式对应于多于一种标记的亲试剂与不同类型的末端氨基酸的关联。如本文所述,应当理解,可以根据本申请使用与多于一种类型的氨基酸相关的单一亲试剂。因此,在一些实施方案中,不同的特征模式对应于一种标记的亲试剂与不同类型的末端氨基酸的关联。

[0198] 如上详述,实时测序过程通常可以涉及末端氨基酸识别和末端氨基酸裂解的循环,其中识别和裂解的相对发生可以通过标记的亲试剂和标记的非特异性外肽酶之间的浓度差来控制。在一些实施方案中,可以优化浓度差,使得在识别单个氨基酸期间检测到的信号脉冲的数量为鉴定提供所需的置信区间。例如,如果初始测序反应提供的信号数据在裂解事件之间的信号脉冲太少而无法确定具有所需置信区间的特征模式,则可以使用相对于亲试剂降低浓度的非特异性外肽酶来重复测序反应。发明人已经认识到用于控制实时测序反应的其他技术,这些技术可以与所描述的浓度差方法结合使用,或替代地使用。

[0199] 在一些实施方案中,测序反应涉及温度依赖性末端氨基酸识别和末端氨基酸裂解的循环。测序反应的每个循环可以在两个温度范围内进行:亲试剂活性优于外肽酶活性(例如,以促进末端氨基酸识别)的第一温度范围(T_1),以及外肽酶活性优于亲试剂活性(例如,以促进末端氨基酸裂解)的第二温度范围(T_2)。测序反应可以通过在第一温度范围 T_1 (以启动氨基酸识别)和第二温度范围 T_2 (以启动氨基酸裂解)之间交替反应混合物温

度来进行。因此,温度依赖性测序过程的进展可通过温度控制,并在不同温度范围(例如, T_1 和 T_2 之间)之间交替,这可以通过手动或自动过程进行。在一些实施方案中,与第二温度范围 T_2 相比,第一温度范围 T_1 内的亲和试剂活性(例如,对氨基酸的结合亲和性(K_D))增加至少10倍、至少100倍、至少1,000倍、至少10,000倍、至少100,000倍或更多。在一些实施方案中,与第一温度范围 T_1 相比,第二温度范围 T_2 内的外肽酶活性(例如,底物转化为裂解产物的速率)增加至少2倍、10倍、至少25倍、至少50倍、至少100倍、至少1,000倍或更多。

[0200] 在一些实施方案中,第一温度范围 T_1 低于第二温度范围 T_2 。在一些实施方案中,第一温度范围 T_1 在约15°C和约40°C之间(例如,在约25°C和约35°C之间、在约15°C和约30°C之间、在约20°C和约30°C之间)。在一些实施方案中,第二温度范围 T_2 在约40°C和约100°C之间(例如,在约50°C和约90°C之间、在约60°C和约90°C之间、在约70°C和约90°C之间)。在一些实施方案中,第一温度范围 T_1 在约20°C和约40°C之间(例如,约30°C),并且第二温度范围 T_2 在约60°C和约100°C之间(例如,约80°C)。

[0201] 在一些实施方案中,第一温度范围 T_1 高于第二温度范围 T_2 。在一些实施方案中,第一温度范围 T_1 在约40°C和约100°C之间(例如,在约50°C和约90°C之间、在约60°C和约90°C之间、在约70°C和约90°C之间)。在一些实施方案中,第二温度范围 T_2 在约15°C和约40°C之间(例如,在约25°C和约35°C之间、在约15°C和约30°C之间、在约20°C和约30°C之间)。在一些实施方案中,第一温度范围 T_1 在约60°C和约100°C之间(例如,约80°C),并且第二温度范围 T_2 在约20°C和约40°C之间(例如,约30°C)。

[0202] 在一些实施方案中,本申请提供了使用发光激活试剂的发光依赖性测序过程。在一些实施方案中,发光依赖性测序过程涉及发光依赖性氨基酸识别和裂解的循环。测序反应的每个循环可以通过将测序反应混合物暴露于两种不同的发光条件来进行:亲和试剂活性优于外肽酶活性(例如,以促进氨基酸识别)的第一发光条件,以及外肽酶活性优于亲和试剂活性(例如,以促进氨基酸裂解)的第二发光条件。通过在将反应混合物暴露于第一发光条件(以启动氨基酸识别)和将反应混合物暴露于第二发光条件(以启动氨基酸裂解)之间交替来进行测序反应。作为实例而非限制,在一些实施方案中,两种不同的发光条件包括第一波长和第二波长。

[0203] 在一些方面,本申请通过评估一种或多种标记的亲和试剂与末端和内部氨基酸的结合相互作用以及标记的非特异性外肽酶与末端氨基酸的结合相互作用提供了实时多肽测序的方法。在一些实施方案中,使用标记的亲和试剂,其在末端和内部位置与一种类型的氨基酸选择性结合和解离。选择性结合在信号输出中产生一系列脉冲。然而,在这种方法中,所述一系列脉冲以由整个多肽中氨基酸类型的数量决定的速率发生。因此,在一些实施方案中,对应于结合事件的脉冲频率将诊断当前存在于多肽中的同源氨基酸的数量。

[0204] 标记的非特异性肽酶可以以比标记的亲和试剂相对较低的浓度存在,例如,以在裂解事件之间提供最佳时间窗。此外,在某些实施方案中,标记的非特异性肽酶的独特可识别发光标记将指示裂解事件何时发生。随着多肽经历迭代裂解,每当末端氨基酸被标记的非特异性肽酶裂解时,对应于标记的亲和试剂结合的脉冲频率将逐步下降。因此,在一些实施方案中,可以在这种方法中基于脉冲模式和/或基于在裂解事件之间检测到的模式内发生的脉冲频率来鉴定氨基酸并由此对多肽进行测序。

[0205] B. 通过标记多肽的降解进行测序

[0206] 在一些方面,本申请提供了通过鉴定对应于已知多肽序列的独特氨基酸组合来对多肽进行测序的方法。在一些实施方案中,所述方法包括检测标记多肽的选择性标记的氨基酸。在一些实施方案中,标记多肽包含经选择性修饰的氨基酸,使得不同的氨基酸类型包含不同的发光标记。如本文所用,除非另有说明,否则标记多肽是指包含一个或多个选择性标记的氨基酸侧链的多肽。选择性标记方法和与标记多肽的制备和分析有关的细节是本领域已知的(参见,例如,Swaminathan等人,PLoS Comput Biol.2015,11(2):e1004080)。

[0207] 如本文所述,在一些方面,本申请提供了对多肽进行测序的方法,所述方法通过多肽降解过程中获得数据,并分析数据以确定与在多肽降解过程中顺序暴露于多肽末端的氨基酸相对应的数据部分。在一些实施方案中,数据的部分包括一系列信号脉冲,指示一种或多种氨基酸识别分子与暴露于多肽末端的连续氨基酸(例如,在降解期间)的关联。在一些实施方案中,所述一系列信号脉冲对应于降解过程中在多肽末端的一系列可逆单分子结合相互作用。

[0208] 在一些方面,本文所述的多肽测序技术产生的数据表明当多肽被裂解手段(例如,一种或多种裂解试剂)降解时,多肽如何与结合手段(例如,一种或多种氨基酸识别分子)相互作用。如上所讨论的,数据可以包括对应于在末端的裂解事件之间的多肽末端的关联事件的一系列特征模式。在一些实施方案中,本文所述的测序方法包括使单个多肽分子与结合手段和裂解手段接触,其中结合手段和裂解手段被设置为在裂解事件之前实现至少10个关联事件。在一些实施方案中,所述手段被设置为实现两个裂解事件之间的至少10个关联事件。

[0209] 如本文所述,在一些实施方案中,多个单分子测序反应在样品孔阵列中并行进行。在一些实施方案中,阵列包含约10,000至约1,000,000个样品孔。在一些实施方式中,样品孔的体积可以在约 10^{-21} 升和约 10^{-15} 升之间。因为样品孔的体积小,单分子的检测事件可能是可能的,因为在任何给定时间,样品孔内可能只有大约一个多肽。统计上,一些样品孔可能不包含单分子测序反应,而一些样品孔可能包含不止一个单个多肽分子。然而,可观数量的样品孔可以各自包含单分子反应(例如,在一些实施方案中至少30%),从而可以对大量样品孔并行进行单分子分析。在一些实施方案中,结合手段和裂解手段被设置为在裂解事件之前在至少10%(例如,10-50%、超过50%、25-75%、至少80%或更多)的样品孔中实现至少10个关联事件,其中发生单分子反应。在一些实施方案中,结合手段和裂解手段被设置为在裂解事件之前对单分子反应中多肽的至少50%(例如,超过50%、50-75%、至少80%或更多)的氨基酸实现至少10个关联事件。

[0210] 在一些实施方案中,标记多肽被固定并暴露于激发源。可以检测来自标记多肽的聚集发光,并且在一些实施方案中,随时间暴露于发光可由于发光标记降解(例如,由于光漂白引起的降解)而导致检测信号的损失。在一些实施方案中,标记多肽包含选择性标记的氨基酸的独特组合,其产生初始检测信号。发光标记随时间的降解导致光漂白的标记多肽的检测信号相应降低。在一些实施方案中,可以通过分析一种或多种发光特性(例如,通过发光寿命分析的信号去卷积)对信号进行去卷积。在一些实施方案中,标记多肽的选择性标记的氨基酸的独特组合已经在计算上被预先计算和在经验上被验证——例如,基于蛋白质组的已知多肽序列。在一些实施方案中,将检测到的氨基酸标记的组合与生物体蛋白质组的已知序列数据库进行比较,以鉴定数据库中对应于标记多肽的特定多肽。

[0211] 在一些实施方案中,确定最佳样品浓度以进行在大规模并行分析中最大化采样的测序反应。在一些实施方案中,选择浓度使得阵列中所需分数(例如,30%)的样品孔在任何给定时间被占据。不希望囿于理论的,认为尽管多肽经一段时间被漂白,相同的孔仍可用于进一步分析。通过扩散,阵列中大约30%的样品孔可用于每3分钟进行一次分析。作为说明性实例,在百万个样品孔芯片中,每小时可以采样6,000,000个多肽,或在4小时阶段内采样24,000,000个多肽。

[0212] 在一些方面,本申请提供了一种通过检测经受末端氨基酸修饰和裂解的重复循环的标记多肽的发光来对多肽进行测序的方法。在一些实施方案中,对于通过Edman降解进行测序的其他方法,所述方法通常如本文所述进行。

[0213] 在一些实施方案中,所述方法包括(i)修饰标记多肽的末端氨基酸的步骤。如本文别处所述,在一些实施方案中,修饰包括使末端氨基酸与异硫氰酸酯(例如,PITC)接触以形成异硫氰酸酯修饰的末端氨基酸。在一些实施方案中,异硫氰酸酯修饰将末端氨基酸转化为更易于被裂解试剂(例如,如本文所述的化学或酶促裂解试剂)去除的形式。因此,在一些实施方案中,所述方法包括(ii)使用本文别处详述的用于Edman降解的化学或酶促方法去除经修饰的末端氨基酸的步骤。

[0214] 在一些实施方案中,所述方法包括对多个循环重复步骤(i)至(ii),在此期间检测标记多肽的发光,并且可以检测对应于从末端去除标记氨基酸的裂解事件作为检测信号的减少。在一些实施方案中,步骤(ii)之后的信号没有变化鉴定了未知类型的氨基酸。因此,在一些实施方案中,部分序列信息可以通过评估在每个连续轮次中在步骤(ii)之后检测到的信号来确定,通过基于检测到的信号的变化通过确定的同一性指定氨基酸类型或基于检测到的信号没有变化将氨基酸类型鉴定为未知。

[0215] 在一些方面,根据本申请对多肽进行测序的方法包括通过标记多肽的持续酶促裂解进行测序。在一些实施方案中,使用修饰的进行性外肽酶对标记多肽进行降解,所述修饰的进行性外肽酶从一个末端向另一个末端连续裂解末端氨基酸。外肽酶在本文别处详细描述。在一些实施方案中,标记多肽经受固定化的进行性外肽酶的降解。在一些实施方案中,固定化的标记多肽经受进行性外肽酶的降解。

[0216] 在一些实施方案中,进行性外肽酶的持续合成速率是已知的,使得检测到的信号降低之间的时序可用于计算每个检测事件之间的未标记氨基酸的数量。例如,如果以每秒去除一个氨基酸的方式裂解40个氨基酸的多肽,则具有3个信号的标记多肽最初会显示所有3个信号,然后是2个信号,然后是1个信号,最后没有信号。以这种方式,可以确定标记氨基酸的顺序。因此,这些方法可用于确定部分序列信息,例如用于基于多肽片段测序的蛋白质组学分析。

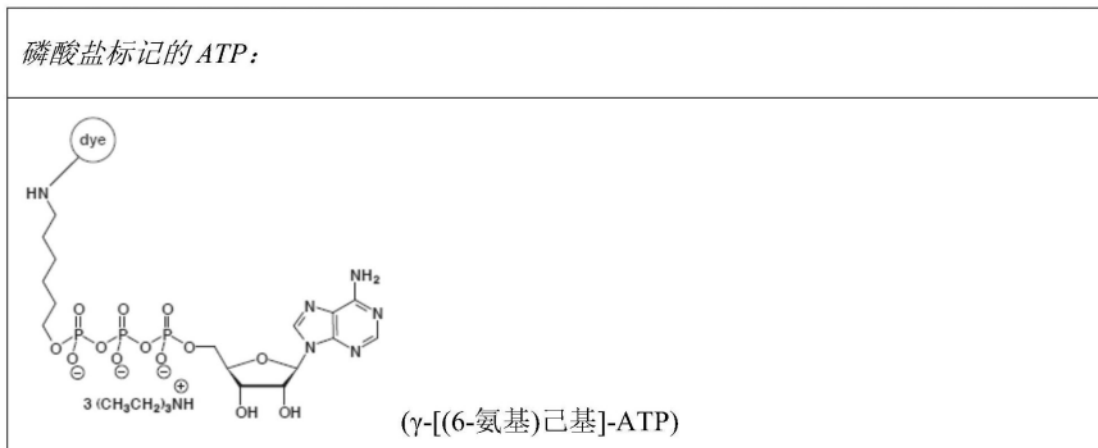
[0217] 在一些实施方案中,单分子多肽测序可以使用基于ATP的福斯特共振能量转移(FRET)方案(例如,使用一种或多种标记的辅因子)来实现。在一些实施方案中,可以使用固定化的ATP依赖性蛋白酶、供体标记的ATP和多肽底物的受体标记的氨基酸进行通过基于辅因子的FRET的测序。在一些实施方案中,可以用受体标记氨基酸,并且可以用供体标记一种或多种辅因子。

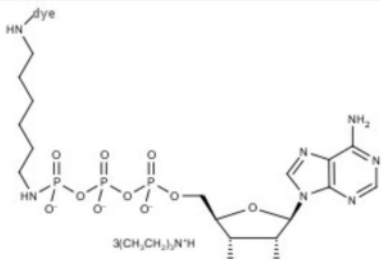
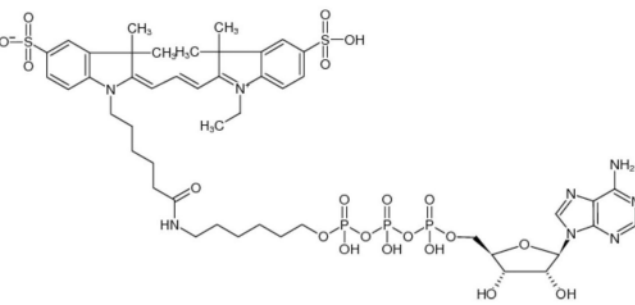
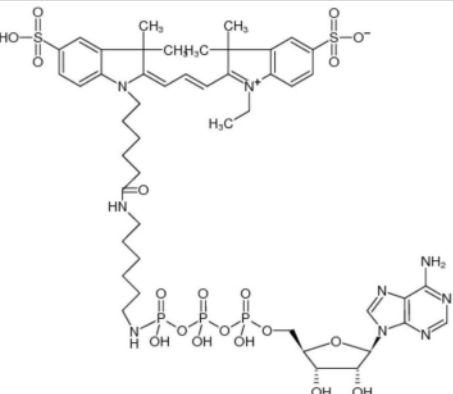
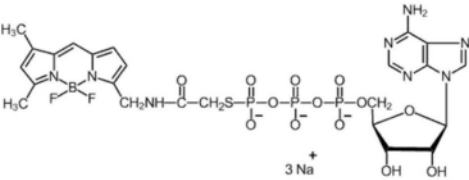
[0218] 例如,在一些实施方案中,使提取的多肽变性,并且用荧光染料标记半胱氨酸和赖氨酸。在一些实施方案中,使用蛋白质转位酶(例如细菌C1pX)的工程化形式与单个底物多

肽结合,将它们展开,并通过其纳米通道将它们转位。在一些实施方案中,转位酶用供体染料标记,并且当基质通过纳米通道时,在转位酶上的供体与基质上的两种或更多种不同的受体染料之间发生FRET。然后可以根据FRET信号确定标记的氨基酸的顺序。在一些实施方案中,可以使用表3中所示的一种或多种以下非限制性标记的ATP类似物。

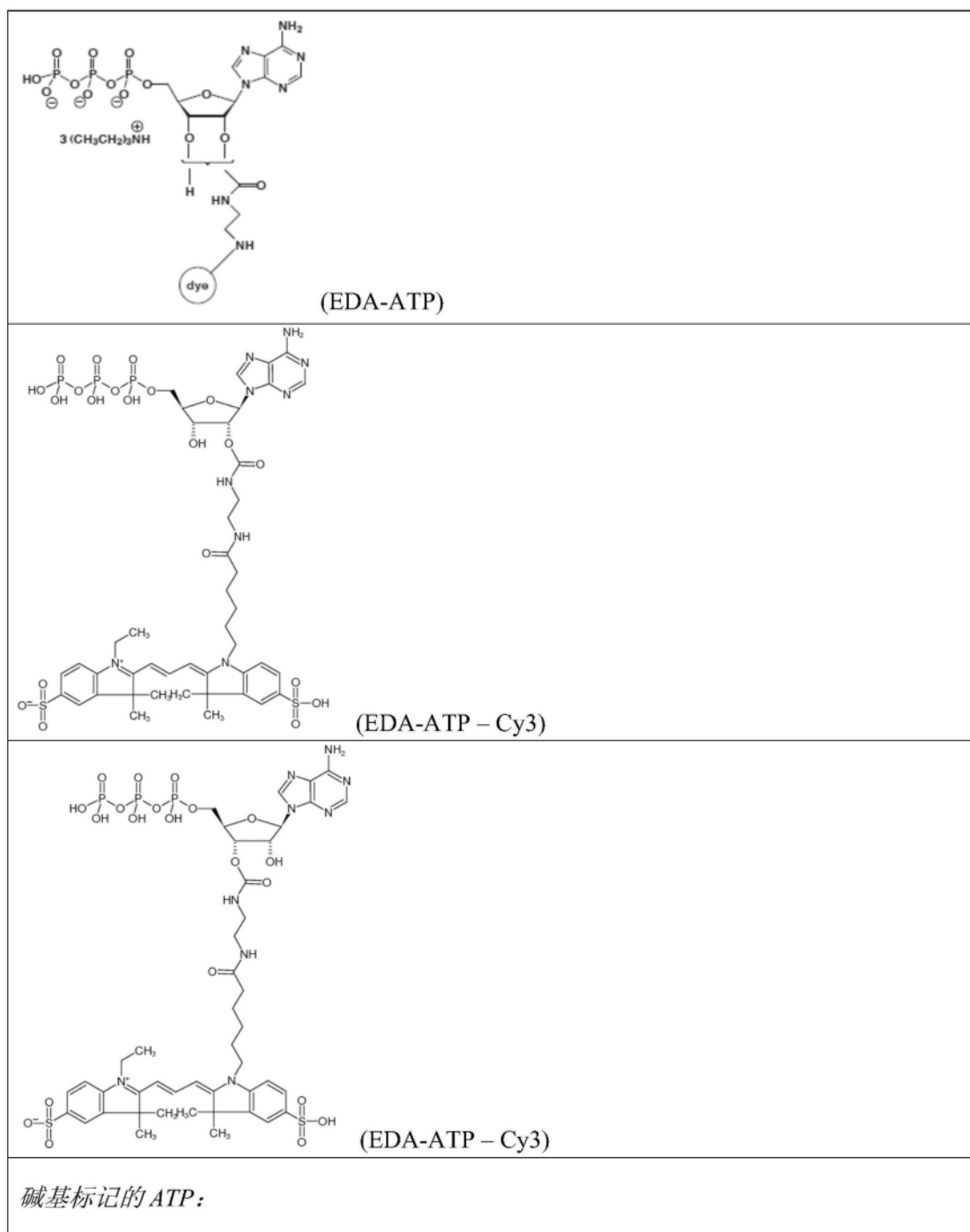
[0219] 表3. 标记的ATP类似物的非限制性实例

[0220]

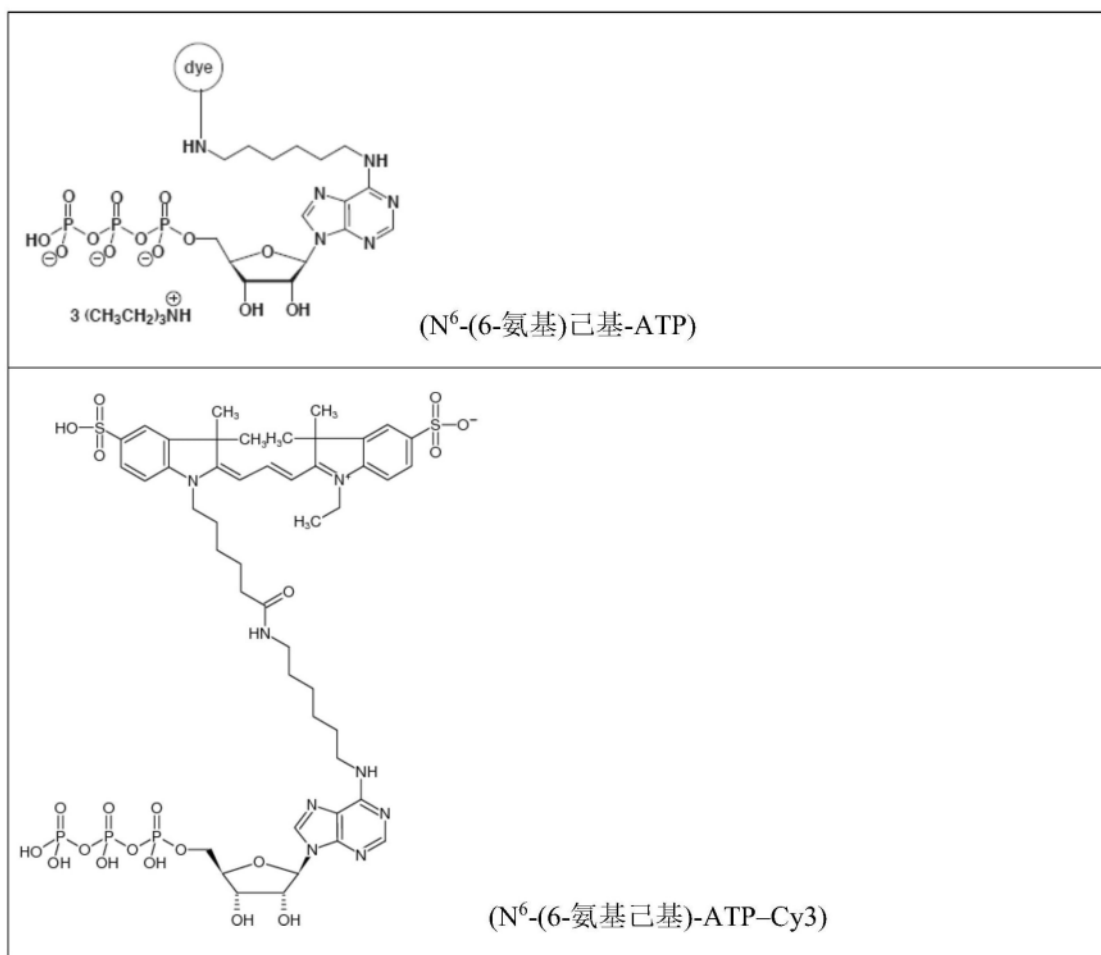


	 (γ-[(6-氨基己基)亚胺基]-ATP)
	 (γ-(6-氨基己基)-ATP-Cy3)
[0221]	 (γ-[(6-氨基己基)亚胺基]-ATP-Cy3)
	 (BODIPY FL ATPγS)
	核糖标记的 ATP:

[0222]



[0223]



[0224] C. 测序样品的制备

[0225] 可以在测序之前修饰多肽样品(例如,富集的多肽样品)。

[0226] 在一些实施方案中,多肽的N末端氨基酸或C末端氨基酸被修饰。在一些实施方案中,用能够固定在表面(例如,用于多肽分析的芯片上的样品孔的表面)上的部分来修饰多肽的末端。在一些实施方案中,此类方法包括根据本申请修饰待分析的标记多肽的末端。在其他实施方案中,此类方法包括根据本申请使多肽底物修饰降解或易位的蛋白质或酶的末端。

[0227] 在一些实施方案中,多肽的羧基末端在包括以下的方法中被修饰:(i)封闭多肽的游离羧酸根基团;(ii)使多肽变性(例如,通过热和/或化学手段);(iii)封闭多肽的游离硫醇基团;(iv)消化多肽以产生至少一个包含游离C末端羧酸根基团的多肽片段;和(v)将功能部分缀合(例如,化学地)到游离的C末端羧酸根基团上。在一些实施方案中,所述方法进一步包括,在(i)之后和(ii)之前,对包含多肽的样品进行透析。

[0228] 在一些实施方案中,多肽的羧基末端以包括以下的方法被修饰:(i)使多肽变性(例如,通过热和/或化学手段);(ii)封闭多肽的游离硫醇基团;(iii)消化多肽以产生至少一个包含游离C末端羧酸根基团的多肽片段;(iv)封闭游离的C末端羧酸根基团以产生至少一个包含封闭的C末端羧酸根基团的多肽片段;和(v)将功能部分缀合(例如,酶促地)到封闭的C末端羧酸根基团上。在一些实施方案中,所述方法进一步包括,在(iv)之后和(v)之前,对包含多肽的样品进行透析。

[0229] 在一些实施方案中,封闭游离羧酸根基团是指这些基团的化学修饰,其相对于未经修饰的羧酸根改变化学反应性。合适的羧酸根封闭方法是本领域已知的,并且应该将侧链羧酸根基团修饰为在化学上不同于待官能化的多肽的羧基末端羧酸根基团。在一些实施方案中,封闭游离羧酸根基团包括多肽的游离羧酸根基团的酯化或酰胺化。在一些实施方案中,封闭游离羧酸根基团包括多肽的游离羧酸根基团的甲酯化,例如,通过使多肽与甲醇HCl反应。可用于封闭游离羧酸根基团的试剂和技术的其他实例包括但不限于4-磺基-2,3,5,6-四氟苯酚(STP)和/或碳二亚胺例如N-(3-二甲基氨基丙基)-N'-乙基碳二亚胺盐酸盐(EDAC)、脲试剂、重氮甲烷、用于Fischer酯化的醇和酸,使用N-羟基琥珀酰亚胺(NHS)形成NHS酯(可能作为后续酯或胺形成的中间体),或与羰基二咪唑(CDI)反应或形成混合酸酐,或任何其他可能通过形成酯或酰胺来修饰或封闭羧酸的方法。

[0230] 在一些实施方案中,封闭游离硫醇基团是指这些基团的化学修饰,其相对于未经修饰的硫醇改变化学反应性。在一些实施方案中,封闭游离硫醇基团包括对多肽的游离硫醇基团进行还原和烷基化。在一些实施方案中,通过使多肽与二硫苏糖醇(DTT)以及碘乙酰胺和碘乙酸中的一种或两种接触来进行还原和烷基化。可以使用的另外的和替代的半胱氨酸还原试剂的实例是众所周知的,并且包括但不限于2-巯基乙醇、三(2-羧乙基)膦盐酸盐(TCEP)、三丁基膦、二硫丁胺(DTBA)或任何能够还原硫醇基团的试剂。可以使用的另外的和替代的半胱氨酸封闭(例如,半胱氨酸烷基化)试剂的实例是众所周知的,并且包括但不限于丙烯酰胺、4-乙烯基吡啶、N-乙基马来酰亚胺(NEM)、N-ε-马来酰亚胺基己酸(EMC),或任何修饰半胱氨酸以防止二硫键形成的试剂。

[0231] 在一些实施方案中,消化包括酶促消化。在一些实施方案中,通过在消化条件下使多肽与内肽酶(例如胰蛋白酶)接触来进行消化。在一些实施方案中,消化包括化学消化。用于化学和酶促消化的合适试剂的实例是本领域已知的,并且包括但不限于胰蛋白酶、化学胰蛋白酶、Lys-C、Arg-C、Asp-N、Lys-N、BNPS-粪臭素、CNBr、半胱天冬酶、甲酸、谷氨酰内肽酶、羟胺、碘代苯甲酸、中性粒细胞弹性蛋白酶、胃蛋白酶、脯氨酸-内肽酶、蛋白酶K、葡萄球菌肽酶I、嗜热菌蛋白酶和凝血酶。

[0232] 在一些实施方案中,功能部分包含生物素分子。在一些实施方案中,功能部分包含反应性化学部分,例如炔基。在一些实施方案中,缀合功能部分包括如本领域已知的通过羧肽酶Y对羧基末端羧甲基酯基团进行生物素化。

[0233] 在一些实施方案中,将增溶部分添加到多肽中。因此,在一些实施方案中,本文提供的方法和组合物可用于用增加其溶解度的部分修饰多肽的末端。在一些实施方案中,增溶部分可用于由片段化(例如,酶促片段化,例如使用胰蛋白酶)产生并且相对不溶的小多肽。例如,在一些实施方案中,多肽库中的短多肽可以通过将聚合物(例如,短寡核苷酸、糖或其他带电荷的聚合物)缀合至多肽来溶解。

[0234] D. 发光标记

[0235] 如本文所用,发光标记是吸收一个或多个光子并且可以随后在一个或多个时段后发射一个或多个光子的分子。在一些实施方案中,该术语可与“标记”或“发光分子”互换使用,这取决于上下文。根据本文所述的某些实施方案的发光标记可以指标记的亲试剂的发光标记、标记的肽酶(例如标记的外肽酶、标记的非特异性外肽酶)的发光标记、标记的肽的发光标记、标记的辅因子的发光标记或本文所述的另一种标记的组合物。在一些实施方

案中,根据本申请的发光标记是指包含一种或多种标记氨基酸的标记多肽的标记氨基酸。

[0236] 在一些实施方案中,发光标记可以包含第一和第二生色团。在一些实施方案中,第一生色团的激发态能够通过能量转移到第二生色团而弛豫。在一些实施方案中,能量转移是福斯特共振能量转移(FRET)。这样的FRET对可用于提供具有使标记更容易从混合物中的多个发光标记中区分的性质的发光标记。在其他实施方案中,FRET对包含第一发光标记的第一生色团和第二发光标记的第二生色团。在某些实施方案中,FRET对可以吸收第一光谱范围内的激发能量并发射第二光谱范围内的发光。

[0237] 在一些实施方案中,发光标记是指荧光团或染料。通常,发光标记包含芳族或杂芳族化合物,并且可以是芘、蒽、萘、萘胺、吡啶、芪、吲哚、苯并吲哚、恶唑、呋唑、噻唑、苯并噻唑、苯并恶唑、菲啶、吩恶嗪、卟啉、喹啉、乙锭、苯甲酰胺、花青、羧花青、水杨酸盐、邻氨基苯甲酸盐、香豆素、荧光素、罗丹明、氧杂蒽或其他类似化合物。

[0238] 在一些实施方案中,发光标记包含选自以下中的一种或多种的染料:5/6-羧基罗丹明6G、5-羧基罗丹明6G、6-羧基罗丹明6G、6-TAMRA、Abberior® STAR440SXP、Abberior® STAR 470SXP、Abberior® STAR 488、Abberior® STAR 512、Abberior® STAR 520SXP、Abberior® STAR 580、Abberior® STAR 600、Abberior® STAR 635、Abberior® STAR 635P、Abberior® STAR RED、Alexa Fluor® 350、Alexa Fluor® 405、Alexa Fluor® 430、Alexa Fluor® 480、Alexa Fluor® 488、Alexa Fluor® 514、Alexa Fluor® 532、Alexa Fluor® 546、Alexa Fluor® 555、Alexa Fluor® 568、Alexa Fluor® 594、Alexa Fluor® 610-X、Alexa Fluor® 633、Alexa Fluor® 647、Alexa Fluor® 660、Alexa Fluor® 680、Alexa Fluor® 700、Alexa Fluor® 750、Alexa Fluor® 790、AMCA、ATTO 390、ATTO 425、ATTO 465、ATTO 488、ATTO 495、ATTO 514、ATTO 520、ATTO 532、ATTO 542、ATTO 550、ATTO 565、ATTO 590、ATTO 610、ATTO 620、ATTO 633、ATTO 647、ATTO 647N、ATTO 655、ATTO 665、ATTO 680、ATTO 700、ATTO 725、ATTO 740、ATTO Oxa12、ATTO Rho101、ATTO Rho11、ATTO Rho12、ATTO Rho13、ATTO Rho14、ATTO Rho3B、ATTO Rho6G、ATTO Thiol2、BD Horizon™ V450、BODIPY® 493/501、BODIPY® 530/550、BODIPY® 558/568、BODIPY® 564/570、BODIPY® 576/589、BODIPY® 581/591、BODIPY® 630/650、BODIPY® 650/665、BODIPY® FL、BODIPY® FL-X、BODIPY® R6G、BODIPY® TMR、BODIPY® TR、CAL Fluor® Gold 540、CAL Fluor® Green 510、CAL Fluor® Orange 560、CAL Fluor® Red 590、CAL Fluor® Red 610、CAL Fluor® Red 615、CAL Fluor® Red 635、Cascade® Blue、CF™350、CF™405M、CF™405S、CF™488A、CF™514、CF™532、CF™543、CF™546、CF™555、CF™568、CF™594、CF™620R、CF™633、CF™633-V1、CF™640R、CF™640R-V1、CF™640R-V2、CF™660C、CF™660R、CF™680、CF™680R、CF™680R-V1、CF™750、CF™770、CF™790、Chromeo™642、Chromis 425N、Chromis 500N、Chromis 515N、Chromis 530N、Chromis 550A、Chromis 550C、Chromis 550Z、Chromis 560N、Chromis 570N、Chromis 577N、Chromis 600N、Chromis 630N、Chromis 645A、Chromis 645C、Chromis 645Z、Chromis 678A、Chromis 678C、Chromis 678Z、Chromis 770A、Chromis 770C、Chromis 800A、Chromis 800C、Chromis 830A、Chromis 830C、Cy®3、

Cy®3.5、Cy®3B、Cy®5、Cy®5.5、Cy®7、DyLight® 350、DyLight® 405、DyLight® 415-Co1、DyLight® 425Q、DyLight® 485-LS、DyLight® 488、DyLight® 504Q、DyLight® 510-LS、DyLight® 515-LS、DyLight® 521-LS、DyLight® 530-R2、DyLight® 543Q、DyLight® 550、DyLight® 554-R0、DyLight® 554-R1、DyLight® 590-R2、DyLight® 594、DyLight® 610-B1、DyLight® 615-B2、DyLight® 633、DyLight® 633-B1、DyLight® 633-B2、DyLight® 650、DyLight® 655-B1、DyLight® 655-B2、DyLight® 655-B3、DyLight® 655-B4、DyLight® 662Q、DyLight® 675-B1、DyLight® 675-B2、DyLight® 675-B3、DyLight® 675-B4、DyLight® 679-C5、DyLight® 680、DyLight® 683Q、DyLight® 690-B1、DyLight® 690-B2、DyLight® 696Q、DyLight® 700-B1、DyLight® 700-B1、DyLight® 730-B1、DyLight® 730-B2、DyLight® 730-B3、DyLight® 730-B4、DyLight® 747、DyLight® 747-B1、DyLight® 747-B2、DyLight® 747-B3、DyLight® 747-B4、DyLight® 755、DyLight® 766Q、DyLight® 775-B2、DyLight® 775-B3、DyLight® 775-B4、DyLight® 780-B1、DyLight® 780-B2、DyLight® 780-B3、DyLight® 800、DyLight® 830-B2、Dyomics-350、Dyomics-350XL、Dyomics-360XL、Dyomics-370XL、Dyomics-375XL、Dyomics-380XL、Dyomics-390XL、Dyomics-405、Dyomics-415、Dyomics-430、Dyomics-431、Dyomics-478、Dyomics-480XL、Dyomics-481XL、Dyomics-485XL、Dyomics-490、Dyomics-495、Dyomics-505、Dyomics-510XL、Dyomics-511XL、Dyomics-520XL、Dyomics-521XL、Dyomics-530、Dyomics-547、Dyomics-547P1、Dyomics-548、Dyomics-549、Dyomics-549P1、Dyomics-550、Dyomics-554、Dyomics-555、Dyomics-556、Dyomics-560、Dyomics-590、Dyomics-591、Dyomics-594、Dyomics-601XL、Dyomics-605、Dyomics-610、Dyomics-615、Dyomics-630、Dyomics-631、Dyomics-632、Dyomics-633、Dyomics-634、Dyomics-635、Dyomics-636、Dyomics-647、Dyomics-647P1、Dyomics-648、Dyomics-648P1、Dyomics-649、Dyomics-649P1、Dyomics-650、Dyomics-651、Dyomics-652、Dyomics-654、Dyomics-675、Dyomics-676、Dyomics-677、Dyomics-678、Dyomics-679P1、Dyomics-680、Dyomics-681、Dyomics-682、Dyomics-700、Dyomics-701、Dyomics-703、Dyomics-704、Dyomics-730、Dyomics-731、Dyomics-732、Dyomics-734、Dyomics-749、Dyomics-749P1、Dyomics-750、Dyomics-751、Dyomics-752、Dyomics-754、Dyomics-776、Dyomics-777、Dyomics-778、Dyomics-780、Dyomics-781、Dyomics-782、Dyomics-800、Dyomics-831、eFluor® 450、伊红、FITC、荧光素、HiLyte™ Fluor 405、HiLyte™ Fluor 488、HiLyte™ Fluor 532、HiLyte™ Fluor 555、HiLyte™ Fluor 594、HiLyte™ Fluor 647、HiLyte™ Fluor 680、HiLyte™ Fluor 750、IRDye® 680LT、IRDye® 750、IRDye® 800CW、JOE、LightCycler® 640R、LightCycler® Red 610、LightCycler® Red 640、LightCycler® Red 670、LightCycler® Red 705、丽丝胺罗丹明

B、Naphthofluorescein、Oregon Green® 488、Oregon Green® 514、Pacific Blue™、Pacific Green™、Pacific Orange™、PET、PF350、PF405、PF415、PF488、PF505、PF532、PF546、PF555P、PF568、PF594、PF610、PF633P、PF647P、Quasar® 570、Quasar® 670、Quasar® 705、罗丹明123、罗丹明6G、罗丹明B、罗丹明绿、罗丹明绿-X、罗丹明红、ROX、Seta™ 375、Seta™ 470、Seta™ 555、Seta™ 632、Seta™ 633、Seta™ 650、Seta™ 660、Seta™ 670、Seta™ 680、Seta™ 700、Seta™ 750、Seta™ 780、Seta™ APC-780、Seta™ PerCP-680、Seta™ R-PE-670、Seta™ 646、SeTau 380、SeTau 425、SeTau 647、SeTau 405、Square 635、Square 650、Square 660、Square 672、Square 680、磺酰罗丹明101、TAMRA、TET、Texas Red®、TMR、TRITC、Yakima Yellow™、Zenon®、Zy3、Zy5、Zy5.5和Zy7。

[0239] E. 发光

[0240] 在一些方面，本申请涉及基于发光标记的一种或多种发光特性的多肽测序和/或鉴定。在一些实施方案中，基于发光寿命、发光强度、亮度、吸收光谱、发射光谱、发光量子产率或其两种或更多种的组合来鉴定发光标记。在一些实施方案中，多种类型的发光标记可以基于不同的发光寿命、发光强度、亮度、吸收光谱、发射光谱、发光量子产率或其中两种或更多种的组合而彼此区分。鉴定可以指指定与发光标记相关的一种类型的氨基酸（例如，单一类型或类型的子集）的确切身份和/或数量，并且还可以指指定多肽中氨基酸相对于其他类型的氨基酸的位置。

[0241] 在一些实施方案中，通过将发光标记暴露于一系列单独的光脉冲并评估从标记发射的每个光子的时序或其他特性来检测发光。在一些实施方案中，从标记顺序发射的多个光子的信息被聚集和评估以鉴定标记并由此鉴定相关类型的氨基酸。在一些实施方案中，标记的发光寿命由从标记顺序发射的多个光子确定，并且发光寿命可用于鉴定标记。在一些实施方案中，标记的发光强度由从标记顺序发射的多个光子确定，并且发光强度可用于鉴定标记。在一些实施方案中，标记的发光寿命和发光强度由从标记顺序发射的多个光子确定，并且发光寿命和发光强度可用于鉴定标记。

[0242] 在本申请的一些方面，将单个多肽分子暴露于多个单独的光脉冲，并检测和分析一系列发射光子。在一些实施方案中，所述一系列发射光子提供了关于存在的并且在实验期间在反应样品中不改变的单个多肽分子的信息。然而，在一些实施方案中，所述一系列发射光子提供关于在反应样品中在不同时间（例如，随着反应或过程进行）存在的一系列不同分子的信息。作为实例而非限制，此类信息可用于根据本申请对经受化学或酶促降解的多肽进行测序和/或鉴定。

[0243] 在某些实施方案中，发光标记吸收一个光子并在一个时段后发射一个光子。在一些实施方案中，可以通过测量所述时段来确定或估计标记的发光寿命。在一些实施方案中，可以通过测量多个脉冲事件和发射事件的多个时段来确定或估计标记的发光寿命。在一些实施方案中，可以通过测量时段在多种类型的标记的发光寿命中区分标记的发光寿命。在一些实施方案中，可以通过测量多个脉冲事件和发射事件的多个时段在多种类型的标记的发光寿命中区分标记的发光寿命。在某些实施方案中，通过确定或估计标记的发光寿命在多种类型的标记中鉴定或区分标记。在某些实施方案中，通过在多种类型的标记的多种发光寿命中区分标记的发光寿命，在多种类型的标记中鉴定或区分标记。

[0244] 可以使用任何合适的方法(例如,通过使用合适的技术测量寿命或通过确定发射的时间相关特性)来确定发光标记的发光寿命。在一些实施方案中,确定一个标记的发光寿命包括确定相对于另一标记的寿命。在一些实施方案中,确定标记的发光寿命包括确定相对于参照的寿命。在一些实施方案中,确定标记的发光寿命包括测量寿命(例如荧光寿命)。在一些实施方案中,确定标记的发光寿命包括确定一种或多种指示寿命的时间特性。在一些实施方案中,可以基于多个发射事件(例如,1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、30、40、50、60、70、80、90、100或更多个发射事件)发生在相对于激发脉冲的一个或多个时间门控窗口的分布来确定标记的发光寿命。例如,可以基于关于激发脉冲测量的光子到达时间的分布将标记的发光寿命与具有不同发光寿命的多个标记区分开来。

[0245] 应当理解,发光标记的发光寿命指示在标记达到激发态之后发射的光子的时序,并且可以通过指示光子的时序的信息来区分标记。一些实施方案可以包括基于标记的发光寿命通过测量与所述标记发射的光子相关联的时间来区分标记与多个标记。时间分布可以提供发光寿命的指示,该指示可以从分布中确定。在一些实施方案中,可以基于时间分布区分标记与多个标记,例如通过将时间分布与对应于已知标记的参照分布进行比较。在一些实施方案中,发光寿命的值由时间分布确定。

[0246] 如本文所用,在一些实施方案中,发光强度是指每单位时间由发光标记发射的发射光子的数量,该发光标记通过递送脉冲激发能量而被激发。在一些实施方案中,发光强度是指每单位时间检测到的发射光子的数量,这些光子由通过脉冲激发能量的递送而被激发的标记发射并且由特定传感器或传感器组检测。

[0247] 如本文所用,在一些实施方案中,亮度是指报告每个发光标记的平均发射强度的参数。因此,在一些实施方案中,“发射强度”可用于一般指包含一种或多种标记的组合物的亮度。在一些实施方案中,标记的亮度等于其量子产率和消光系数的乘积。

[0248] 如本文所用,在一些实施方案中,发光量子产率是指在给定波长或在给定光谱范围内导致发射事件的激发事件的分数,并且通常小于1。在一些实施方案中,本文所述的发光标记的发光量子产率在0和约0.001之间、在约0.001和约0.01之间、在约0.01和约0.1之间、在约0.1和约0.5之间、在约0.5和0.9之间或在约0.9和1之间。在一些实施方案中,通过确定或估计发光量子产率来鉴定标记。

[0249] 如本文所用,在一些实施方案中,激发能量是来自光源的光脉冲。在一些实施方案中,激发能量在可见光谱中。在一些实施方案中,激发能量在紫外光谱中。在一些实施方案中,激发能量在红外光谱中。在一些实施方案中,激发能量处于或接近发光标记的吸收最大值,从该发光标记中检测多个发射光子。在某些实施方案中,激发能量在约500nm和约700nm之间(例如,在约500nm和约600nm之间、在约600nm和约700nm之间、在约500nm和约550nm之间、在约550nm和约600nm之间、在约600nm和约650nm之间或在约650nm和约700nm之间)。在某些实施方案中,激发能量可以是单色的或限制在光谱范围内。在一些实施方案中,光谱范围具有约0.1nm至约1nm、约1nm至约2nm或约2nm至约5nm的范围。在一些实施方案中,光谱范围具有约5nm至约10nm、约10nm至约50nm或约50nm至约100nm的范围。

[0250] IV. 用于样品制备的试剂盒

[0251] 在一些方面,本公开涉及用于制备用于测序的多肽样品(例如,富集样品)的试剂盒。试剂盒可能足以制备用于测序的一个或多个多肽样品(例如,富集样品)。在一些实施方

案中,试剂盒足以制备单个多肽样品。在其他实施方案中,试剂盒足以制备至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少14个、至少15个、至少20个、至少25个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90或至少100个多肽样品。

[0252] 在一些实施方案中,试剂盒包含如本文所述的包含多个富集分子的富集组分。参见“多肽富集方法”。在一些实施方案中,试剂盒包含如本文所述的修饰剂。参见“多肽富集方法”。在一些实施方案中,试剂盒包含如本文所述的亲和试剂。参见“多肽测序方法学”。在一些实施方案中,试剂盒包含如本文所述的标记的肽酶。参见“多肽测序方法学”。

[0253] 试剂盒可以对一种或多种生物体(例如,一种或多种单细胞和/或多细胞生物体)是特异性的。在一些实施方案中,试剂盒包含修饰一种或多种生物体的多肽、与其结合、被其结合等的组分(例如,富集分子、修饰剂或其组合)。例如,在一些实施方案中,试剂盒包含修饰人类蛋白质组中的一种或多种已知多肽、与其结合、被其结合等的组分。

[0254] 在一些实施方案中,试剂盒对一种或多种疾病或病况具有特异性。例如,试剂盒可以是肿瘤学试剂盒、心脏病学试剂盒、遗传性疾病试剂盒、细菌毒力因子试剂盒、抗生素耐药性试剂盒或其组合。

[0255] 肿瘤学试剂盒可以包含与ABL1、ABL2、ACSL3、ACVR2A、ADAMTS20、ADGRA2、ADGRB3、ADGRL3、AFF1、AFF3、AKAP9、AKT1、AKT2、AKT3、ALK、AMER1、APC、AR、ARID1A、ARID2、ARNT、ASXL1、ATF1、ATM、ATR、ATRX、AURKA、AURKB、AURKC、AXL、BAP1、BCL10、BCL11A、BCL11B、BCL2、BCL2L1、BCL2L2、BCL3、BCL6、BCL7A、BCL9、BCR、BIRC2、BIRC3、BIRC5、BLM、BLNK、BMPR1A、BRAF、BRCA1、BRCA2、BRD3、BRIP1、BTK、BUB1B、CACNA1D、CARD11、CASC5、CASP8、CBFA2T3、CBFB、CBL、CCND1、CCND2、CCNE1、CD79A、CD79B、CDC73、CDH1、CDH11、CDH2、CDH20、CDH5、CDK12、CDK4、CDK6、CDK8、CDKN2A、CDKN2B、CDKN2C、CEBPA、CHEK1、CHEK2、CIC、CKS1B、CMPK1、COL1A1、CRBN、CREB1、CREBBP、CRKL、CRLF2、CRTC1、CSF1R、CSMD3、CTNNA1、CTNNB1、CYLD、CYP2C19、CYP2D6、DAXX、DCC、DDB2、DDIT3、DDR2、DEK、DICER1、DNMT3A、DPYD、DST、EGFR、EML4、EP300、EP400、EPA3、EPA7、EPHB1、EPHB4、EPHB6、ERBB2、ERBB3、ERBB4、ERCC1、ERCC2、ERCC3、ERCC4、ERCC5、ERG、ESR1、ETS1、ETV1、ETV4、EXT1、EXT2、EZH2、FANCA、FANCC、FANCD2、FANCF、FANCG、FAS、FBXW7、FCGR2B、FGFR1、FGFR2、FGFR3、FGFR4、FH、FLCN、FLI1、FLT1、FLT3、FLT4、FN1、FOXA1、FOXL2、FOXO1、FOXO3、FOXP1、FOXP4、FZR1、G6PD、GATA1、GATA2、GATA3、GDNF、GNA11、GNAQ、GNAS、GPC3、GRM8、GUCY1A2、HCAR1、HEY1、HIF1A、HIST1H3B、HLF、HMGA1、HNF1A、HOOK3、HOXA13、HOXD11、HRAS、HSP90AA1、HSP90AB1、ICK、IDH1、IDH2、IGF1R、IGF2、IGF2R、IKBKB、IKBKE、IKZF1、IL2、IL21R、IL6ST、IL7R、ING4、IRF4、IRS2、ITGA10、ITGA9、ITGB2、ITGB3、JAK1、JAK2、JAK3、JUN、KAT6A、KAT6B、KDM5C、KDM6A、KDR、KEAP1、KIAA1549、KIT、KLF6、KMT2A、KMT2C、KMT2D、KRAS、LAMP1、LCK、LIFR、LPP、LRP1B、LTF、LTK、MAF、MAFB、MAGEA1、MAGI1、MALT1、MAML2、MAP2K1、MAP2K2、MAP2K4、MAP3K7、MAPK1、MAPK8、MARK1、MARK4、MBD1、MCL1、MDM2、MDM4、MEN1、MET、MITF、MLH1、MLLT10、MLLT4、MLLT6、MMP2、MN1、MPL、MRE11A、MSH2、MSH6、MTCP1、MTOR、MTR、MTRR、MUC1、MUTYH、MYB、MYC、MYCL、MYCN、MYD88、MYH11、MYH9、NBN、NCOA1、NCOA2、NCOA4、NF1、NF2、NFE2L2、NFKB1、NFKB2、NIN、NKX2-1、NLRP1、NOTCH1、NOTCH2、NOTCH4、NPM1、NR4A3、NRAS、NSD1、NTRK1、NTRK3、NUMA1、NUP214、NUP98、NUTM2A、NUTM2B、OMD、P2RY8、PAK3、PALB2、PARP1、PAX3、PAX5、PAX7、PAX8、PBRM1、PBX1、

KCNQ4、KIAA0196、KLHL7、KRAS、KRT14、KRT5、L1CAM、LAMB3、LAMP2、LDB3、LMNA、LRAT、LRRK2、MAPT、MC1R、MECP2、MED12、MEN1、MERTK、MFN2、MLH1、MMAA、MMAB、MMACHC、MPZ、MSH2、MTM1、MUT、MYBPC3、MYH11、MYH6、MYH7、MYL2、MYL3、MYLK、MYO7A、MYOZ2、NF1、NF2、NIPBL、NKX2-5、NME8、NPC1、NPC2、NR2E3、NRAS、NSD1、OCA2、OCRL、OTC、PABPN1、PAFAH1B1、PAH、PAX3、PAX6、PCDH15、PEX1、PEX10、PEX13、PEX14、PEX19、PEX26、PEX3、PEX5、PINK1、PKD1、PKD2、PKHD1、PKP2、PLEC、PLN、PLOD1、PMM2、PMP22、POLG、PPT1、PRCD、PRKAG2、PROM1、PRPF31、PRPF8、PRPH2、PSEN1、PSEN2、PTCH1、PTPN11、RAF1、RAG1、RAG2、RAI1、RAPSN、RB1、RDH12、RET、RHO、ROR2、RP9、RPE65、RPGR、RPGRIP1、RPL11、RPL35A、RPS10、RPS19、RPS24、RPS26、RPS6KA3、RPS7、RS1、RSPH4A、RSPH9、RYY1、RYY2、SALL4、SCN1B、SCN3B、SCN4B、SCN5A、SCN9A、SEMA4A、SERPINA1、SERPING1、SGCD、SH3BP2、SIX1、SIX5、SLC25A13、SLC25A4、SLC26A4、SMAD3、SMAD4、SNCA、SNRNP200、SNTA1、SOD1、SOS1、SOX9、SPATA7、SPG7、STARD3、TAF1、TAZ、TBX5、TCOF1、TGFB1、TGFB2、TMEM43、TNNC1、TNNT1、TNNT2、TNXB、TOPORS、TP53、TPM1、TSC1、TSC2、TTPA、TTR、TULP1、TWIST1、TYR、USH1C、USH2A、VCL、VHL、WAS、WRN、WT1或其任意组合结合(或被其结合)的富集分子。

[0258] 细菌毒力因子试剂盒可以包含与 α -C蛋白、 α -溶血素、 β -C蛋白、 β -溶血素/溶细胞素、 β -溶血素、 δ -溶血素、 γ -溶血素、T2SS、AAFs、ACF、AI-2、ALO、AS、Ace、酸性磷酸酶、不动杆菌素(Acinetobactin)、Acm、AcrAB、ActA、AdeFGH外排泵、粘附菌毛、Adr1、Adr2、AdsA、气杆菌素(Aerobactin)、气溶素(Aerolysin)、Afa/Dr家族、Agf、AhpC、Ail、AipA、藻酸盐、碱性蛋白酶、尿囊素利用、Ami、AnsP、炭疽毒素、抗原85、ArgP、AslA、Asp14、AtxA、Aureolysin、Auto、自溶素(Autolysin)、BFP、BSH、BabA、BadA/Vomp、Bap、BapC、BfmRS、BimA、生物素合成、BoNT、BoaA、BoaB、BopD、Brk、Bsa T3SS、BslA、BtpA/Btp1/TcpB、BtpB、BvgAS、BvrR-BvrS、C2毒素、C3毒素、C5a肽酶、C β G、CAI-1、CAMP因子、CARDS毒素、CBPs、CDT、CHIPS、CNA、CNF-1、CNF_y、CPAF、CPE、CT、CadF、CagA、Capsule、Capsule I、CbpA/PspC、CcmC、CdpA、CdtB、Chu、CiaB、CiaC、Cif、ClpC、ClpE、ClpP、凝集因子(Clumping factor)、Colibactin、CsrA、Csu菌毛、Cya、CytK、Cytadherence细胞器、溶细胞素、DNA酶、DT、DevRS、DipA、Dispersin、Dnt、Dot/Icm、Dot/Icm T4SS、Dr粘附素、EAST1、ECP、EF-Tu、ESAT-6/CFP-10、ESX-1、ESX-3、ESX-5、Eap/Map、Ebp pili、EbpS、EcbA、Efa-1/LifA、EfaA、Ent、肠杆菌素、Erp、Esp、EspA、EspB、EspC、EspD、EspF、EspG、EspH、EspP、EtpA、Exe T2SS、表皮剥脱毒素、ExoA、ExoS、ExoT、ExoU、ExoY、F1抗原、F1C菌毛、FBPs、FHA、FadD33、FarAB、FbpA、FbpABC、FbsA、FbsB、FdeC、FeoAB、菌毛、鞭毛、Flp IV型菌毛、FmvB、FnBPs、FrgA、FsaP、Fsr、FupA、Fur、GGT、GRAB、GadC、明胶酶、GrvA、GspA、GtcA、HBL、HMW1/HMW2、HP-NAP、HSI-I、血凝菌毛、Hap、HbhA、不耐热毒素(LT)、耐热毒素(ST)、溶血素、Hgp、HhuA、Hia/Hsf、HitABC、HmbR、HopZ、Hpt、HpuAB、Hsp60、HspR、HspX、HxuABC、透明质酸裂解酶、透明质酸胶囊、透明质酸酶、Ibes、IcsA(VirG)、IcsP(SopA)、IdeR、IdeS、IgA1蛋白酶、IleP、IlpA、InhA、InlA、InlB、InlC、InlF、InlJ、InlK、InlP、细胞间粘附蛋白、紧密粘附素(Intimin)、侵袭素、侵袭素B/Ifp、侵袭素C/Ilp、侵袭素D、IraAB、Iron、Isd、异柠檬酸裂合酶、JlpA、K1胶囊、KatA、KatAB、KatG、LAM、LLO、LLS、LOS、LPS、Lap、LapB、LasA、LasB、侧鞭毛、Lbp、Legiobactin、Ler、LetA/S、Lewis抗原、LigA、LipF、脂肪酶、Lmb、LntA、LpeA、Lpf、LplA1、Lsp、Lymphostatin/LifA、M蛋白、MAM7、MARTX、MOMP、MSHA菌毛、MSHA IV型菌毛、Map、MgtBC、MgtC、Mig-14、Mig-

5、Mip、MisL、MmaA4、MntABC、Mpl、MprAB、MsrAB、MtrCDE、分枝杆菌素、Myf/pH6抗原、神经氨酸酶、Nhe、硝酸还原酶、NleA/EspI、NleC、NleD、NspA、O抗原、OapA、OatA、OipA、OmpA、OmpU、Opa、Opc、P菌毛、P2蛋白、P44/Msp2家族、P5蛋白、P97/P102横向同源物家族、PDIM、PE/PE-PGRS、PEB1、PI-1、PI-2、PI-2a、PLC、PNAG、PVL、Paa、PanC/PanD、PavA、PavB、PbpG、PcaA、Pef、Per、百日咳杆菌粘附素、Pet、PfbA、Pgda、PhoP、PhoPQ、磷脂酶A2、磷脂酶C、磷脂酶D、Pht、Pic、菌毛、Pla、PlcA、PlcB、Pld、肺炎球菌溶血素、极性鞭毛、孔蛋白(Porin)、PrfA、PrsA2、PsaA、PspA、Ptx、嘌呤生物合成、Pyochelin、绿脓杆菌素、Pyoverdine、嘧啶生物合成、Quorum感应、Quorum感应、Quorum-感应、RatB、Rck、RcsAB、RecN、RelA、鼠李糖脂、Rhizoferrin、RicA、RickA、RipA、RmpA、RpoS、RtxA、Rvh T4SS、S菌毛、SCIN、SDr、SE、SIC、SLO、SLS、SMase、SabA、Sal、Sat、Sbi、Sca1、Sca2、Sca4、Scm、SgrA、ShET1、ShET2、ShdA、志贺毒素、Shu、SigA、SigE、SigF、SigH、SinH、SodA、SodB、SodC、SodCI、SpA、SpaP、SpeB、Spes、SprE、Spv、Staphopain、葡萄球菌凝固酶、葡萄球菌激酶、StcE、链激酶、Stx、表面脂蛋白、SvpA、T2SS、T3SS、T3SS1、T3SS2、T6SS、T6SS-1、T7SS、TCP、TCT、TDH、TRH、TSST-1、TTSS、TTSS (SPI-1encode)、TTSS (SPI-2encode)、Tap IV型菌毛、Tbp、TcdA、TcdB、TcfA、TcpC、TeNT、Tir、TlyC、ToxB、TraJ、Trw IV型分泌系统、Tsh、1型菌毛、3型菌毛、I型菌毛、I型菌毛、IV型菌毛、IV型分泌系统、VII型分泌系统、脲酶、V8蛋白酶、VCC、VacA、Vi抗原、Vip、VirB IV型分泌系统、VirB/VirD4 IV型分泌系统、VpadF、WhiB3、YadA、YapC、YapE、YapJ、YapK、YapV、YaxAB、Ybt、耶尔森菌素、Ymt、Yst、Zot、 α -梭菌蛋白酶、 α -毒素(CpPLC)、 α -毒素(novy)、 α -毒素(septum)、 β -毒素、 β 2-毒素、enh基因座、 ϵ -毒素、fHbp、iota-毒素、 κ -毒素、 μ -毒素、p60、菌毛、pmiA、rOmpA/Sca0、rOmpB/Sca5、唾液酸酶、 θ -毒素/PF0、vWbp、xcp分泌系统或其任意组合结合(或被其结合)的富集分子。

[0259] 抗生素耐药性试剂盒可以包含与AAC(1)-I、AAC(2')-IIa、AAC(2')-IIb、AAC(2')-Ia、AAC(2')-Ib、AAC(2')-Ic、AAC(2')-Id、AAC(2')-Ie、AAC(3)-IIIa、AAC(3)-IIIb、AAC(3)-IIIc、AAC(3)-IIa、AAC(3)-IIb、AAC(3)-IIc、AAC(3)-IIId、AAC(3)-IIe、AAC(3)-IV、AAC(3)-IXa、AAC(3)-Ia、AAC(3)-Ib、AAC(3)-Ib/AAC(6')-Ib'、AAC(3)-Ic、AAC(3)-Id、AAC(3)-VIIa、AAC(3)-VIIa、AAC(3)-VIa、AAC(3)-Xa、AAC(6')-29a、AAC(6')-29b、AAC(6')-30/AAC(6')-Ib'融合蛋白、AAC(6')-31、AAC(6')-32、AAC(6')-33、AAC(6')-34、AAC(6')-I30、AAC(6')-IIa、AAC(6')-IIb、AAC(6')-IIc、AAC(6')-Ia、AAC(6')-Iaa、AAC(6')-Iad、AAC(6')-Iae、AAC(6')-Iaf、AAC(6')-Iag、AAC(6')-Iai、AAC(6')-Iaj、AAC(6')-Iak、AAC(6')-Ian、AAC(6')-Ib、AAC(6')-Ib'、AAC(6')-Ib-杭州、AAC(6')-Ib-SK、AAC(6')-Ib-苏州、AAC(6')-Ib-cr、AAC(6')-Ib10、AAC(6')-Ib11、AAC(6')-Ib3、AAC(6')-Ib4、AAC(6')-Ib7、AAC(6')-Ib8、AAC(6')-Ib9、AAC(6')-Ic、AAC(6')-Ie-APH(2'')-Ia、AAC(6')-If、AAC(6')-Ig、AAC(6')-Ih、AAC(6')-Ii、AAC(6')-Iid、AAC(6')-Iih、AAC(6')-Ij、AAC(6')-Ik、AAC(6')-Il、AAC(6')-Im、AAC(6')-Ip、AAC(6')-Iq、AAC(6')-Ir、AAC(6')-Is、AAC(6')-Isa、AAC(6')-It、AAC(6')-Iu、AAC(6')-Iv、AAC(6')-Iw、AAC(6')-Ix、AAC(6')-Iy、AAC(6')-Iz、ACC-1、ACC-2、ACC-3、ACC-4、ACC-5、ACI-1、ACT-1、ACT-10、ACT-12、ACT-13、ACT-14、ACT-15、ACT-16、ACT-17、ACT-18、ACT-19、ACT-2、ACT-20、ACT-21、ACT-22、ACT-23、ACT-24、ACT-25、ACT-27、ACT-28、ACT-29、ACT-3、ACT-30、ACT-31、ACT-32、ACT-33、ACT-35、ACT-36、ACT-37、ACT-38、ACT-4、ACT-5、ACT-6、ACT-7、ACT-8、ACT-9、ADC-1、ADC-10、ADC-11、ADC-12、ADC-13、ADC-

14、ADC-15、ADC-16、ADC-17、ADC-18、ADC-19、ADC-20、ADC-21、ADC-22、ADC-23、ADC-25、ADC-30、ADC-31、ADC-39、ADC-41、ADC-42、ADC-43、ADC-44、ADC-5、ADC-56、ADC-58、ADC-59、ADC-60、ADC-61、ADC-62、ADC-67、ADC-68、ADC-7、ADC-73、ADC-74、ADC-75、ADC-76、ADC-77、ADC-78、ADC-79、ADC-8、ADC-81、ADC-82、AER-1、AIM-1、ANT(2'')-Ia、ANT(3'')-IIa、ANT(3'')-IIb、ANT(3'')-IIc、ANT(3'')-Ii-AAC(6')-IIId融合蛋白、ANT(4')-IIa、ANT(4')-IIb、ANT(4')-Ia、ANT(4')-Ib、ANT(6)-Ia、ANT(6)-Ib、ANT(9)-Ia、APH(2'')-IIIa、APH(2'')-IIa、APH(2'')-IVa、APH(2'')-Ie、APH(2'')-If、APH(2'')-Ig、APH(3'')-Ia、APH(3'')-Ib、APH(3'')-Ic、APH(3')-IIIa、APH(3')-IIa、APH(3')-IIb、APH(3')-IIc、APH(3')-IVa、APH(3')-IX、APH(3')-Ia、APH(3')-Ib、APH(3')-VI、APH(3')-VIIa、APH(3')-VIIb、APH(3')-VIIa、APH(3')-VIa、APH(3')-Va、APH(3')-Vb、APH(3')-Vc、APH(4)-Ia、APH(4)-Ib、APH(6)-Ia、APH(6)-Ib、APH(6)-Ic、APH(6)-Id、APH(7'')-Ia、APH(9)-Ia、APH(9)-Ib、AQU-1、AQU-2、AQU-3、ARL-1、ARL-2、ARL-3、ARL-4、ARL-5、ARL-6、AST-1、AZECL-25、鲍曼不动杆菌(*Acinetobacter baumannii*)AbaF、鲍曼不动杆菌AbaQ、鲍曼不动杆菌AbuO、鲍曼不动杆菌AmvA、赋予对亚胺培南耐药性的鲍曼不动杆菌OprD、鲍曼不动杆菌ampC β -内酰胺酶、赋予对氟喹诺酮类耐药性的鲍曼不动杆菌gyrA、赋予对氟喹诺酮类耐药性的鲍曼不动杆菌parC、AcrE、AcrF、AcrS、农杆菌(*Agrobacterium fabrum*)氯霉素乙酰转移酶、ArmR、AxyX、AxyY、AxyZ、BAT-1、BCL-1、BEL-1、BEL-2、BEL-3、BES-1、BIC-1、BIL-1、BJP-1、BKC-1、BPU-1、BRO-1、BRO-2、BRP(MBL)、BUT-1、克劳氏芽孢杆菌(*Bacillus clausii*)氯霉素乙酰转移酶、短小芽孢杆菌(*Bacillus pumilus*)cat86、枯草芽孢杆菌mprF、具有赋予对达托霉素耐药性的突变的枯草芽孢杆菌pgsA、BahA、赋予对氟喹诺酮类耐药性的杆菌状巴尔通氏体(*Bartonella bacilliformis*)gyrA、赋予对氨基香豆素耐药性的杆菌状巴尔通氏体gyrB、BcI、BcII、赋予对利福平耐药性的青春双歧杆菌(*Bifidobacterium adolescentis*)rpoB突变体、赋予对莫匹罗星耐药性的双歧杆菌(*Bifidobacterium*)ileS、Bla1、Bla2、赋予对庆大霉素耐药性的伯氏疏螺旋体(*Borrelia burgdorferi*)16S rRNA突变、赋予对卡那霉素耐药性的伯氏疏螺旋体16S rRNA突变、赋予对壮观霉素耐药性的伯氏疏螺旋体16S rRNA突变、赋予对磷霉素耐药性的伯氏疏螺旋体murA突变、具有赋予对泰乐菌素耐药性的突变的猪痢疾短螺旋体(*Brachyspira hyodysenteriae*)23S rRNA突变、猪布鲁氏菌(*Brucella suis*)mprF、类鼻疽伯克霍尔德菌(*Burkholderia pseudomallei*)Omp38、CAM-1、CARB-1、CARB-10、CARB-12、CARB-14、CARB-16、CARB-17、CARB-18、CARB-19、CARB-2、CARB-20、CARB-21、CARB-22、CARB-23、CARB-3、CARB-4、CARB-5、CARB-6、CARB-7、CARB-8、CARB-9、CAU-1、CBP-1、CFE-1、CFE-2、CGA-1、CGB-1、CIA-1、CIA-2、CIA-3、CIA-4、CKO-1、CME-1、CMH-1、CMY-1、CMY-10、CMY-100、CMY-101、CMY-102、CMY-103、CMY-104、CMY-105、CMY-106、CMY-108、CMY-11、CMY-110、CMY-111、CMY-112、CMY-113、CMY-114、CMY-115、CMY-116、CMY-117、CMY-118、CMY-119、CMY-12、CMY-13、CMY-131、CMY-132、CMY-133、CMY-135、CMY-14、CMY-15、CMY-16、CMY-17、CMY-18、CMY-19、CMY-2、CMY-20、CMY-21、CMY-22、CMY-23、CMY-24、CMY-25、CMY-26、CMY-27、CMY-28、CMY-29、CMY-30、CMY-31、CMY-32、CMY-33、CMY-34、CMY-35、CMY-36、CMY-37、CMY-38、CMY-39、CMY-4、CMY-40、CMY-41、CMY-42、CMY-43、CMY-44、CMY-45、CMY-46、CMY-47、CMY-48、CMY-49、CMY-5、CMY-50、CMY-51、CMY-53、CMY-54、CMY-55、CMY-56、CMY-57、CMY-58、CMY-59、CMY-6、CMY-60、CMY-61、CMY-62、CMY-63、CMY-64、CMY-65、CMY-66、CMY-67、CMY-68、

CMY-69、CMY-7、CMY-70、CMY-71、CMY-72、CMY-73、CMY-74、CMY-75、CMY-76、CMY-77、CMY-78、CMY-79、CMY-8、CMY-80、CMY-81、CMY-82、CMY-83、CMY-84、CMY-85、CMY-86、CMY-87、CMY-9、CMY-90、CMY-93、CMY-94、CMY-95、CMY-98、CMY-99、CPS-1、CRP、CTX-M-1、CTX-M-10、CTX-M-100、CTX-M-101、CTX-M-102、CTX-M-103、CTX-M-104、CTX-M-105、CTX-M-106、CTX-M-107、CTX-M-108、CTX-M-109、CTX-M-11、CTX-M-110、CTX-M-111、CTX-M-112、CTX-M-113、CTX-M-114、CTX-M-115、CTX-M-116、CTX-M-117、CTX-M-12、CTX-M-121、CTX-M-122、CTX-M-123、CTX-M-124、CTX-M-125、CTX-M-126、CTX-M-129、CTX-M-13、CTX-M-130、CTX-M-131、CTX-M-132、CTX-M-134、CTX-M-136、CTX-M-137、CTX-M-139、CTX-M-14、CTX-M-141、CTX-M-142、CTX-M-144、CTX-M-147、CTX-M-148、CTX-M-15、CTX-M-151、CTX-M-152、CTX-M-155、CTX-M-156、CTX-M-157、CTX-M-158、CTX-M-159、CTX-M-16、CTX-M-160、CTX-M-17、CTX-M-19、CTX-M-2、CTX-M-20、CTX-M-21、CTX-M-22、CTX-M-23、CTX-M-24、CTX-M-25、CTX-M-26、CTX-M-27、CTX-M-28、CTX-M-29、CTX-M-3、CTX-M-30、CTX-M-31、CTX-M-32、CTX-M-33、CTX-M-34、CTX-M-35、CTX-M-36、CTX-M-37、CTX-M-38、CTX-M-39、CTX-M-4、CTX-M-40、CTX-M-41、CTX-M-42、CTX-M-43、CTX-M-44、CTX-M-45、CTX-M-46、CTX-M-47、CTX-M-48、CTX-M-49、CTX-M-5、CTX-M-50、CTX-M-51、CTX-M-52、CTX-M-53、CTX-M-54、CTX-M-55、CTX-M-56、CTX-M-58、CTX-M-59、CTX-M-6、CTX-M-60、CTX-M-61、CTX-M-62、CTX-M-63、CTX-M-64、CTX-M-65、CTX-M-66、CTX-M-67、CTX-M-68、CTX-M-69、CTX-M-7、CTX-M-71、CTX-M-72、CTX-M-74、CTX-M-75、CTX-M-76、CTX-M-77、CTX-M-78、CTX-M-79、CTX-M-8、CTX-M-80、CTX-M-81、CTX-M-82、CTX-M-83、CTX-M-84、CTX-M-85、CTX-M-86、CTX-M-87、CTX-M-88、CTX-M-89、CTX-M-9、CTX-M-90、CTX-M-91、CTX-M-92、CTX-M-93、CTX-M-94、CTX-M-95、CTX-M-96、CTX-M-98、CTX-M-99、大肠弯曲杆菌 (*Campylobacter coli*) 氯霉素乙酰转移酶、具有赋予对红霉素耐药性的突变的空肠弯曲菌 (*Campylobacter jejuni*) 23S rRNA、赋予对氟喹诺酮类耐药性的空肠弯曲菌 *gyrA*、赋予氟喹诺酮类耐药性的牙龈二氧化碳嗜纤维菌 (*Capnocytophaga gingivalis*) *gyrA*、*CatU*、*CblA-1*、*CcrA*、*CepS*、*CfxA*、*CfxA2*、*CfxA3*、*CfxA4*、*CfxA5*、*CfxA6*、具有赋予对大环内酯类抗生素耐药性的突变的沙眼衣原体 (*Chlamydia trachomatis*) 23S rRNA 突变、赋予对磷霉素耐药性的沙眼衣原体固有 *murA*、赋予对链霉素耐药性的莱茵衣藻 (*Chlamydomonas reinhardtii*) 16S rRNA (*rrnS*) 突变、具有赋予对红霉素耐药性的突变的莱茵衣藻 23S rRNA、赋予对壮观霉素耐药性的鸚鵡热衣原体 (*Chlamydophila psittaci*) 16S rRNA 突变、脑膜脓毒性金黄杆菌 (*Chryseobacterium meningosepticum*) *BlaB*、具有赋予对红霉素和克林霉素耐药性的突变的艰难梭菌 (*Clostridioides difficile*) 23S rRNA、赋予对 *elfamycin* 耐药性的艰难梭菌 EF-Tu 突变体、赋予对氟喹诺酮类耐药性的艰难梭菌 *gyrA*、赋予对氟喹诺酮类耐药性的艰难梭菌 *gyrB*、具有赋予对万古霉素耐药性的突变的艰难梭菌 *murG*、具有赋予对利福平耐药性的突变的艰难梭菌 *rpoB*、具有赋予对万古霉素耐药性的突变的艰难梭菌 *rpoC*、丁酸梭菌 (*Clostridium butyricum*) *catB*、产气荚膜梭菌 (*Clostridium perfringens*) *mprF*、纹状体棒状杆菌 (*Corynebacterium striatum*) *tetA*、*CrpP*、赋予对四环素耐药性的 *Cutibacterium acnes* 16S rRNA 突变、赋予对氟喹诺酮类耐药性的 *Cutibacterium acnes* *gyrA*、D-Ala-D-Ala 连接酶、DES-1、DHA-1、DHA-10、DHA-12、DHA-13、DHA-14、DHA-15、DHA-16、DHA-17、DHA-18、DHA-19、DHA-2、DHA-20、DHA-21、DHA-22、DHA-3、DHA-5、DHA-6、DHA-7、DHA-9、DIM-1、DnaA、EBR-1 β -内酰胺酶、EBR-2、ERP-1、ESP-1、EXO-1、

EdeQ、阴沟肠杆菌 (*Enterobacter cloacae*) *acrA*、阴沟肠杆菌 *rob*、具有赋予达托霉素耐药性的突变的粪肠球菌 (*Enterococcus faecalis*) *YvlB*、具有赋予达托霉素耐药性的突变的粪肠球菌 *YybT*、粪肠球菌氯霉素乙酰转移酶、具有赋予对达托霉素耐药性的突变的粪肠球菌 *cls*、具有赋予达托霉素耐药性的突变的粪肠球菌 *drmA*、具有赋予达托霉素耐药性的突变的粪肠球菌 *gdpD*、具有赋予达托霉素耐药性的突变的粪肠球菌 *gshF*、赋予达托霉素耐药性的粪肠球菌 *liaF* 突变体、赋予达托霉素耐药性的粪肠球菌 *liaR* 突变体、赋予达托霉素耐药性的粪肠球菌 *liaS* 突变体、赋予对 GE2270A 耐药性的粪肠球菌 EF-Tu 突变体、粪肠球菌氯霉素乙酰转移酶、赋予对达托霉素耐药性的粪肠球菌 *cls*、赋予达托霉素耐药性的粪肠球菌 *liaF* 突变体、赋予达托霉素耐药性的粪肠球菌 *liaR* 突变体、赋予达托霉素耐药性的粪肠球菌 *liaS* 突变体、*EreA*、*EreA2*、*EreB*、*EreD*、*Erm* (30)、*Erm* (31)、*Erm* (33)、*Erm* (34)、*Erm* (35)、*Erm* (36)、*Erm* (37)、*Erm* (38)、*Erm* (39)、*Erm* (41)、*Erm* (42)、*Erm* (43)、*Erm* (44) v、*Erm* (47)、*Erm* (48)、*Erm* (49)、*Erm* (K)、*Erm* (O) - *lrm*、*ErmA*、*ErmB*、*ErmC*、*ErmD*、*ErmE*、*ErmF*、*ErmG*、*ErmH*、*ErmN*、*ErmO*-*srmA*、*ErmQ*、*ErmR*、*ErmS*、*ErmT*、*ErmU*、*ErmV*、*ErmW*、*ErmX*、*ErmY*、赋予对壮观霉素耐药性的大肠杆菌 16S rRNA (*rrnB*) 突变、赋予对链霉素耐药性的大肠杆菌 16S rRNA (*rrnB*) 突变、赋予对四环素耐药性的大肠杆菌 16S rRNA (*rrnB*) 突变、赋予对 G418 耐药性的大肠杆菌 16SrRNA (*rrsB*) 突变、赋予对庆大霉素 C 耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对卡那霉素 A 耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对新霉素耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对巴龙霉素耐药性的大肠杆菌 16SrRNA (*rrsB*) 突变、赋予对壮观霉素耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对链霉素耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对四环素耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对妥布霉素耐药性的大肠杆菌 16S rRNA (*rrsB*) 突变、赋予对 kasugamicin 耐药性的大肠杆菌 16S rRNA (*rrsC*) 突变、赋予对壮观霉素耐药性的大肠杆菌 16S rRNA (*rrsH*) 突变、赋予对伊短菌素 (edeine) 耐药性的大肠杆菌 16S rRNA 突变、具有赋予对氯霉素耐药性的突变的大肠杆菌 23SrRNA、具有赋予对 UI 克拉霉素耐药性的突变的大肠杆菌 23S rRNA、具有赋予对克林霉素耐药性的突变的大肠杆菌 23S rRNA、具有对赋予红霉素和泰利霉素耐药性的突变的大肠杆菌 23S rRNA、具有赋予对恶唑烷酮类抗生素耐药性的突变的大肠杆菌 23S rRNA、大肠杆菌 *CpxR*、具有赋予对磷霉素耐药性的突变的大肠杆菌 *CyaA*、赋予对 Enacyloxin IIa 耐药性的大肠杆菌 EF-Tu 突变体、赋予对 Pulvomycin 耐药性的大肠杆菌 EF-Tu 突变体、赋予对黄色霉素耐药性的大肠杆菌 EF-Tu 突变体、具有赋予对磷霉素耐药性的突变的大肠杆菌 *GlpT*、大肠杆菌 *LamB*、具有赋予对磷霉素耐药性的突变的大肠杆菌 *PtsI*、具有赋予对磷霉素耐药性的突变的大肠杆菌 *UhpA*、具有赋予对磷霉素耐药性的突变的大肠杆菌 *UhpT*、大肠杆菌 *acrA*、具有赋予多药抗生素耐药性的突变的大肠杆菌 *acrR*、大肠杆菌 *ampC β -内酰胺酶*、大肠杆菌 *ampC1 β -内酰胺酶*、大肠杆菌 *ampH β -内酰胺酶*、大肠杆菌 *emrE*、赋予对三氯生耐药性的大肠杆菌 *fabG* 突变、赋予对异烟肼和三氯生耐药性的大肠杆菌 *fabI* 突变、具有赋予对磺胺类耐药性的突变的大肠杆菌 *folP*、赋予对氟喹诺酮类耐药性的大肠杆菌 *gyrA*、具有赋予对三氯生耐药性的突变的大肠杆菌 *gyrA*、赋予对氨基香豆素耐药性的大肠杆菌 *gyrB*、赋予抗生素耐药性的大肠杆菌 *marR* 突变体、大肠杆菌 *mdfA*、大肠杆菌 *mipA*、具有赋予对磷霉素耐药性的突变的大肠杆菌 *murA*、赋予对呋喃妥因耐药性的大肠杆菌 *nfsA* 突变、具有赋予对呋喃妥因耐药性的突变的大肠杆菌 *nfsB*、具有赋予 β -内酰胺类抗生素耐药性的突变的大肠杆菌 *ompF*、赋予对氟喹诺酮类耐

药性的大肠杆菌parC、赋予对氟喹诺酮类耐药性的大肠杆菌parE、大肠杆菌rob、赋予对利福平耐药性的大肠杆菌rpoB突变体、具有赋予抗生素耐药性的突变的大肠杆菌soxR、具有赋予抗生素耐药性的突变的大肠杆菌soxS、FAR-1、FEZ-1、FIM-1、FONA-1、FONA-2、FONA-3、FONA-4、FONA-5、FONA-6、FOX-1、FOX-10、FOX-2、FOX-3、FOX-4、FOX-5、FOX-7、FOX-8、FOX-9、FPH-1、FRI-1、FRI-2、FRI-3、FTU-1、FomA、FomB、FosA、FosA2、FosA3、FosA4、FosA5、FosA6、FosA7、FosB、FosB1、FosB3、FosB4、FosB5、FosB6、FosC、FosC2、FosD、FosK、FosX、FusF、GES-1、GES-10、GES-11、GES-12、GES-13、GES-14、GES-15、GES-16、GES-17、GES-18、GES-19、GES-2、GES-20、GES-21、GES-22、GES-23、GES-24、GES-26、GES-3、GES-4、GES-5、GES-6、GES-7、GES-8、GES-9、GIM-1、GIM-2、GOB-1、GOB-10、GOB-11、GOB-12、GOB-13、GOB-14、GOB-15、GOB-16、GOB-18、GOB-2、GOB-3、GOB-4、GOB-5、GOB-6、GOB-7、GOB-8、GOB-9、H-NS、HERA-1、HERA-2、HERA-3、HMB-1、赋予对 β -内酰胺类抗生素耐药性的流感嗜血杆菌 (*Haemophilus influenzae*) PBP3、赋予对氟喹诺酮类耐药性的副流感嗜血杆菌 (*Haemophilus parainfluenzae*) gyrA、赋予对氟喹诺酮类耐药性的副流感嗜血杆菌parC、赋予对氯霉素耐药性的嗜盐杆菌 (*Halobacterium halobium*) 23S rRNA突变、赋予对帕他霉素 (pactamycin) 耐药性的盐生盐杆菌 (*Halobacterium salinarum*) 16S rRNA突变、赋予对四环素耐药性的幽门螺杆菌16S rRNA突变、具有赋予对克拉霉素耐药性的突变的幽门螺杆菌23S rRNA、ICR-Mc、ICR-Mo、IMI-1、IMI-2、IMI-3、IMI-4、IMI-7、IMP-1、IMP-10、IMP-11、IMP-12、IMP-13、IMP-14、IMP-15、IMP-16、IMP-18、IMP-19、IMP-2、IMP-20、IMP-21、IMP-22、IMP-24、IMP-25、IMP-26、IMP-27、IMP-28、IMP-29、IMP-3、IMP-30、IMP-31、IMP-32、IMP-33、IMP-34、IMP-35、IMP-37、IMP-38、IMP-4、IMP-40、IMP-41、IMP-42、IMP-43、IMP-44、IMP-45、IMP-48、IMP-5、IMP-51、IMP-55、IMP-56、IMP-6、IMP-7、IMP-8、IMP-9、IND-1、IND-10、IND-11、IND-12、IND-14、IND-15、IND-2、IND-2a、IND-3、IND-4、IND-5、IND-6、IND-7、IND-8、IND-9、JOHN-1、KHM-1、KPC-10、KPC-11、KPC-12、KPC-13、KPC-14、KPC-15、KPC-16、KPC-17、KPC-19、KPC-2、KPC-22、KPC-24、KPC-3、KPC-4、KPC-5、KPC-6、KPC-7、KPC-8、KPC-9、产气克雷伯菌 (*Klebsiella aerogenes*) Omp36、具有赋予多药抗生素耐药性的突变的产气克雷伯菌acrR、赋予对粘菌素抗生素耐药性的克雷伯菌突变体PhoP、肺炎克雷伯菌 (*Klebsiella pneumoniae*) KpnE、肺炎克雷伯菌KpnF、肺炎克雷伯菌KpnG、肺炎克雷伯菌KpnH、肺炎克雷伯菌OmpK35、肺炎克雷伯菌OmpK36、肺炎克雷伯菌OmpK37、肺炎克雷伯菌acrA、具有赋予多药抗生素耐药性的突变的肺炎克雷伯菌acrR、肺炎克雷伯菌ramR突变体、L1 β -内酰胺酶、LAT-1、LCR-1、LEN-1、LEN-10、LEN-11、LEN-12、LEN-13、LEN-14、LEN-15、LEN-16、LEN-18、LEN-19、LEN-2、LEN-20、LEN-21、LEN-22、LEN-23、LEN-24、LEN-26、LEN-3、LEN-4、LEN-5、LEN-6、LEN-7、LEN-8、LEN-9、LRA-1、LRA-10、LRA-12、LRA-13、LRA-17、LRA-18、LRA-19、LRA-2、LRA-3、LRA-5、LRA-7、LRA-8、LRA-9、罗伊氏乳杆菌 (*Lactobacillus reuteri*) cat-TC、香港海鸥菌 (*Laribacter hongkongensis*) ampC β -内酰胺酶、单核细胞增生性李斯特氏菌mprF、LlmA 23S核糖体RNA甲基转移酶、LnuP、LpeA、LpeB、LpxA、LpxC、LpxD、MCR-1.1、MCR-1.10、MCR-1.11、MCR-1.12、MCR-1.13、MCR-1.2、MCR-1.3、MCR-1.4、MCR-1.5、MCR-1.6、MCR-1.7、MCR-1.8、MCR-1.9、MCR-2.1、MCR-2.2、MCR-3.1、MCR-3.10、MCR-3.11、MCR-3.12、MCR-3.2、MCR-3.3、MCR-3.4、MCR-3.5、MCR-3.6、MCR-3.7、MCR-3.8、MCR-3.9、MCR-4.1、MCR-4.2、MCR-4.3、MCR-4.4、MCR-4.5、MCR-5.1、MCR-5.2、MCR-6.1、MCR-7.1、MCR-8.1、MCR-9.1、MIR-1、MIR-10、

MIR-11、MIR-12、MIR-13、MIR-14、MIR-15、MIR-16、MIR-17、MIR-2、MIR-3、MIR-4、MIR-5、MIR-6、MIR-8、MIR-9、MOX-1、MOX-2、MOX-3、MOX-4、MOX-5、MOX-6、MOX-7、MOX-8、MOX-9、MSI-1、MSI-OXA、MUS-1、MUS-2、MdtK、Mef (En2)、MexA、MexB、MexC、MexD、MexE、MexF、MexG、MexH、MexI、MexJ、MexK、MexL、MexR、MexS、MexT、MexV、MexW、MexZ、具有赋予对大环内酯类抗生素耐药性的突变的卡他莫拉菌 (*Moraxella catarrhalis*) 23S rRNA、卡他莫拉菌M35、赋予对氟喹诺酮类耐药性的摩氏摩根氏菌 (*Morganella morganii*) gyrB、MuxA、MuxB、MuxC、MvaT、具有赋予对克拉霉素耐药性的突变的鸟分枝杆菌 (*Mycobacterium avium*) 23S rRNA、具有赋予对阿奇霉素耐药性的突变的细胞内分枝杆菌 (*Mycobacterium intracellulare*) 23S rRNA、具有赋予对克拉霉素耐药性的突变的细胞内分枝杆菌23S rRNA、具有赋予对克拉霉素耐药性的突变的堪萨斯分枝杆菌 (*Mycobacterium kansasii*) 23S rRNA、具有赋予对氨苯砒耐药性的突变的麻风分枝杆菌 (*Mycobacterium leprae*) folP、赋予对氟喹诺酮类耐药性的麻风分枝杆菌gyrB、赋予对利福平耐药性的麻风分枝杆菌rpoB突变、赋予对阿米卡星耐药性的结核分枝杆菌 (*Mycobacterium tuberculosis*) 16S rRNA突变、赋予对卡那霉素耐药性的结核分枝杆菌16S rRNA突变、赋予对链霉素耐药性的结核分枝杆菌16S rRNA突变、赋予对紫霉素耐药性的结核分枝杆菌16S rRNA突变、赋予对乙胺丁醇耐药性的结核分枝杆菌embA突变体、具有赋予对乙胺丁醇耐药性的突变的结核分枝杆菌embB、具有赋予对利福平耐药性的结核分枝杆菌embB、赋予对乙胺丁醇耐药性的结核分枝杆菌embR突变体、具有赋予对乙硫异烟胺耐药性的突变的结核分枝杆菌ethA、具有赋予对对氨基水杨酸耐药性的突变的结核分枝杆菌folC、赋予对链霉素耐药性的结核分枝杆菌gidB突变、赋予对氟喹诺酮类耐药性的结核分枝杆菌gyrA、赋予对氟喹诺酮类耐药性的结核分枝杆菌gyrB突变体、赋予对异烟肼耐药性的结核分枝杆菌inhA突变体、赋予对乙胺丁醇耐药性的结核分枝杆菌iniA突变体、具有赋予对乙胺丁醇耐药性的突变的结核分枝杆菌iniB、赋予对乙胺丁醇耐药性的结核分枝杆菌iniC突变体、赋予对磷霉素耐药性的结核分枝杆菌固有murA、赋予对异烟肼耐药性的结核分枝杆菌kasA突变体、赋予对异烟肼耐药性的结核分枝杆菌katG突变、赋予对乙胺丁醇耐药性的结核分枝杆菌突变体embC、具有赋予对异烟肼耐药性的突变的结核分枝杆菌ndh、赋予对吡嗪酰胺耐药性的结核分枝杆菌pncA突变、具有赋予对对氨基水杨酸耐药性的突变的结核分枝杆菌ribD、赋予对利福平耐药性的结核分枝杆菌rpoB突变体、赋予对吡嗪酰胺耐药性的结核分枝杆菌rpsA突变、赋予对链霉素耐药性的结核分枝杆菌rpsL突变、具有赋予对对氨基水杨酸耐药性的突变的结核分枝杆菌thyA、赋予对氨基糖苷类耐药性的结核分枝杆菌tlyA突变、具有赋予对乙胺丁醇耐药性的突变的结核分枝杆菌变体bovis embB、具有赋予对异烟肼耐药性的突变的结核分枝杆菌变体bovis ndh突变、赋予对阿米卡星耐药性的脓肿分枝杆菌 (*Mycobacteroides abscessus*) 16S rRNA突变、赋予对庆大霉素耐药性的脓肿分枝杆菌16S rRNA突变、赋予对卡那霉素耐药性的脓肿分枝杆菌16S rRNA突变、赋予对新霉素耐药性的脓肿分枝杆菌16SrRNA突变、赋予对妥布霉素耐药性的脓肿分枝杆菌16S rRNA突变、具有赋予对克拉霉素耐药性的突变的脓肿分枝杆菌23S rRNA、赋予对阿米卡星耐药性的龟分枝杆菌 (*Mycobacteroides chelonae*) 16S rRNA突变、赋予对庆大霉素C耐药性的龟分枝杆菌16S rRNA突变、赋予对卡那霉素A耐药性的龟分枝杆菌16SrRNA突变、赋予对新霉素耐药性的龟分枝杆菌16S rRNA突变、赋予对妥布霉素耐药性的龟分枝杆菌16S rRNA突变、具有赋予对克拉霉素耐药性的突变的龟分枝杆菌23S rRNA、

赋予对氟喹诺酮类耐药性的麻风分枝杆菌 (*Mycobacterium leprae*) gyrA、赋予对潮霉素B耐药性的耻垢分枝杆菌 (*Mycolicibacterium smegmatis*) 16S rRNA (rrsA) 突变、赋予对卡那霉素A耐药性的耻垢分枝杆菌16SrRNA (rrsA) 突变、赋予对新霉素耐药性的耻垢分枝杆菌16S rRNA (rrsA) 突变、赋予对紫霉素耐药性的耻垢分枝杆菌16S rRNA (rrsA) 突变、赋予对潮霉素B耐药性的耻垢分枝杆菌16S rRNA (rrsB) 突变、赋予对卡那霉素A耐药性的耻垢分枝杆菌16S rRNA (rrsB) 突变、赋予对新霉素耐药性的耻垢分枝杆菌16S rRNA (rrsB) 突变、赋予对链霉素耐药性的耻垢分枝杆菌16S rRNA (rrsB) 突变、赋予对紫霉素耐药性的耻垢分枝杆菌16S rRNA (rrsB) 突变、具有赋予对克拉霉素耐药性的突变的耻垢分枝杆菌23S rRNA、具有赋予对异烟肼耐药性的突变的耻垢分枝杆菌ndh、具有赋予对大环内酯类抗生素耐药性的突变的发酵支原体 (*Mycoplasma fermentans*) 23S rRNA、赋予对截短侧耳素 (pleuromutilin) 抗生素耐药性的鸡毒支原体 (*Mycoplasma gallisepticum*) 23S rRNA突变、赋予对氟喹诺酮类和大环内酯类抗生素耐药性的生殖支原体 (*Mycoplasma genitalium*) 23S rRNA突变、赋予对氟喹诺酮类耐药性的生殖支原体gyrA突变、赋予对莫西沙星耐药性的生殖支原体parC突变、具有赋予对大环内酯类抗生素耐药性的突变的人型支原体 (*Mycoplasma hominis*) 23S rRNA、赋予对氟喹诺酮类耐药性的人型支原体parC、赋予对红霉素耐药性的肺炎支原体23S rRNA突变、NDM-1、NDM-10、NDM-11、NDM-12、NDM-13、NDM-14、NDM-17、NDM-2、NDM-3、NDM-4、NDM-5、NDM-6、NDM-7、NDM-8、NDM-9、NPS-1、赋予对壮观霉素耐药性的淋病奈瑟菌 (*Neisseria gonorrhoeae*) 16S rRNA突变、赋予对氟喹诺酮类耐药性的淋病奈瑟菌gyrA、赋予对氟喹诺酮类耐药性的淋病奈瑟菌parC、淋病奈瑟菌孔蛋白PIB (por)、赋予对壮观霉素耐药性的脑膜炎奈瑟菌 (*Neisseria meningitidis*) 16S rRNA突变、赋予对β-内酰胺耐药性的脑膜炎奈瑟菌PBP2、Nmca、NmCR、OCH-1、OCH-2、OCH-3、OCH-4、OCH-5、OCH-6、OCH-7、OCH-8、OKP-A-1、OKP-A-10、OKP-A-11、OKP-A-12、OKP-A-13、OKP-A-14、OKP-A-15、OKP-A-16、OKP-A-2、OKP-A-3、OKP-A-4、OKP-A-5、OKP-A-6、OKP-A-7、OKP-A-8、OKP-A-9、OKP-B-1、OKP-B-10、OKP-B-11、OKP-B-12、OKP-B-13、OKP-B-17、OKP-B-18、OKP-B-19、OKP-B-2、OKP-B-20、OKP-B-3、OKP-B-4、OKP-B-5、OKP-B-6、OKP-B-7、OKP-B-8、OKP-B-9、OXA-1、OXA-10、OXA-100、OXA-101、OXA-103、OXA-104、OXA-106、OXA-107、OXA-108、OXA-109、OXA-111、OXA-110、OXA-111、OXA-112、OXA-113、OXA-114a、OXA-115、OXA-116、OXA-117、OXA-118、OXA-119、OXA-12、OXA-120、OXA-121、OXA-122、OXA-123、OXA-124、OXA-125、OXA-126、OXA-127、OXA-128、OXA-129、OXA-13、OXA-130、OXA-131、OXA-132、OXA-133、OXA-134、OXA-135、OXA-136、OXA-137、OXA-138、OXA-139、OXA-14、OXA-140、OXA-141、OXA-142、OXA-143、OXA-144、OXA-145、OXA-146、OXA-147、OXA-148、OXA-149、OXA-15、OXA-150、OXA-151、OXA-152、OXA-153、OXA-154、OXA-155、OXA-156、OXA-157、OXA-158、OXA-16、OXA-160、OXA-161、OXA-162、OXA-163、OXA-164、OXA-165、OXA-166、OXA-167、OXA-168、OXA-169、OXA-17、OXA-170、OXA-171、OXA-172、OXA-173、OXA-174、OXA-175、OXA-176、OXA-177、OXA-178、OXA-179、OXA-18、OXA-180、OXA-181、OXA-182、OXA-183、OXA-184、OXA-185、OXA-19、OXA-192、OXA-194、OXA-195、OXA-196、OXA-197、OXA-198、OXA-199、OXA-2、OXA-20、OXA-200、OXA-201、OXA-202、OXA-203、OXA-204、OXA-205、OXA-206、OXA-207、OXA-208、OXA-209、OXA-21、OXA-210、OXA-211、OXA-212、OXA-213、OXA-214、OXA-215、OXA-216、OXA-217、OXA-219、OXA-22、OXA-223、OXA-224、OXA-225、OXA-226、OXA-228、OXA-229、OXA-23、OXA-230、OXA-231、OXA-232、OXA-

233、OXA-234、OXA-235、OXA-236、OXA-237、OXA-239、OXA-24、OXA-240、OXA-241、OXA-242、OXA-243、OXA-244、OXA-245、OXA-246、OXA-247、OXA-248、OXA-249、OXA-25、OXA-250、OXA-251、OXA-252、OXA-253、OXA-254、OXA-255、OXA-256、OXA-257、OXA-258、OXA-259、OXA-26、OXA-260、OXA-261、OXA-262、OXA-263、OXA-264、OXA-265、OXA-266、OXA-267、OXA-268、OXA-269、OXA-27、OXA-270、OXA-271、OXA-272、OXA-273、OXA-274、OXA-275、OXA-276、OXA-277、OXA-278、OXA-279、OXA-28、OXA-280、OXA-281、OXA-282、OXA-283、OXA-284、OXA-285、OXA-286、OXA-287、OXA-288、OXA-29、OXA-291、OXA-292、OXA-293、OXA-294、OXA-295、OXA-296、OXA-297、OXA-298、OXA-299、OXA-3、OXA-300、OXA-301、OXA-302、OXA-303、OXA-304、OXA-305、OXA-306、OXA-308、OXA-309、OXA-31、OXA-312、OXA-313、OXA-314、OXA-315、OXA-316、OXA-317、OXA-32、OXA-320、OXA-322、OXA-323、OXA-324、OXA-325、OXA-326、OXA-327、OXA-328、OXA-329、OXA-33、OXA-330、OXA-331、OXA-332、OXA-333、OXA-334、OXA-335、OXA-336、OXA-337、OXA-338、OXA-339、OXA-34、OXA-340、OXA-341、OXA-342、OXA-343、OXA-344、OXA-345、OXA-346、OXA-347、OXA-348、OXA-349、OXA-35、OXA-350、OXA-351、OXA-352、OXA-353、OXA-354、OXA-355、OXA-356、OXA-357、OXA-358、OXA-359、OXA-36、OXA-360、OXA-361、OXA-362、OXA-363、OXA-364、OXA-365、OXA-366、OXA-368、OXA-37、OXA-370、OXA-371、OXA-372、OXA-373、OXA-374、OXA-375、OXA-376、OXA-377、OXA-378、OXA-379、OXA-380、OXA-381、OXA-382、OXA-383、OXA-384、OXA-385、OXA-386、OXA-387、OXA-388、OXA-389、OXA-390、OXA-391、OXA-392、OXA-397、OXA-398、OXA-4、OXA-400、OXA-401、OXA-402、OXA-403、OXA-404、OXA-405、OXA-406、OXA-407、OXA-408、OXA-409、OXA-411、OXA-412、OXA-413、OXA-414、OXA-415、OXA-416、OXA-417、OXA-418、OXA-42、OXA-420、OXA-421、OXA-422、OXA-423、OXA-424、OXA-425、OXA-426、OXA-427、OXA-429、OXA-43、OXA-430、OXA-431、OXA-432、OXA-433、OXA-435、OXA-436、OXA-437、OXA-438、OXA-439、OXA-440、OXA-441、OXA-442、OXA-443、OXA-444、OXA-446、OXA-447、OXA-448、OXA-449、OXA-45、OXA-450、OXA-451、OXA-452、OXA-453、OXA-454、OXA-455、OXA-457、OXA-458、OXA-459、OXA-46、OXA-460、OXA-461、OXA-464、OXA-465、OXA-466、OXA-47、OXA-470、OXA-471、OXA-472、OXA-473、OXA-474、OXA-475、OXA-476、OXA-477、OXA-478、OXA-479、OXA-48、OXA-480、OXA-482、OXA-483、OXA-484、OXA-485、OXA-486、OXA-488、OXA-49、OXA-5、OXA-50、OXA-51、OXA-53、OXA-535、OXA-54、OXA-55、OXA-56、OXA-57、OXA-58、OXA-59、OXA-60、OXA-61、OXA-62、OXA-63、OXA-64、OXA-65、OXA-66、OXA-663、OXA-664、OXA-665、OXA-67、OXA-68、OXA-69、OXA-7、OXA-70、OXA-71、OXA-72、OXA-73、OXA-74、OXA-75、OXA-76、OXA-77、OXA-78、OXA-79、OXA-80、OXA-82、OXA-83、OXA-84、OXA-85、OXA-86、OXA-87、OXA-88、OXA-89、OXA-9、OXA-90、OXA-91、OXA-92、OXA-93、OXA-94、OXA-95、OXA-96、OXA-97、OXA-98、OXA-99、OXY-1-1、OXY-1-2、OXY-1-3、OXY-1-4、OXY-1-6、OXY-2-1、OXY-2-10、OXY-2-2、OXY-2-3、OXY-2-4、OXY-2-5、OXY-2-6、OXY-2-7、OXY-2-8、OXY-2-9、OXY-3-1、OXY-4-1、OXY-5-1、OXY-5-2、OXY-6-1、OXY-6-2、OXY-6-3、OXY-6-4、OpmB、OpmD、OpmH、OprA、OprJ、OprM、OprN、OprZ、PC1 β -内酰胺酶 (blaZ)、PDC-1、PDC-10、PDC-2、PDC-3、PDC-4、PDC-5、PDC-6、PDC-7、PDC-73、PDC-74、PDC-75、PDC-76、PDC-77、PDC-78、PDC-79、PDC-8、PDC-80、PDC-81、PDC-82、PDC-83、PDC-84、PDC-85、PDC-86、PDC-87、PDC-88、PDC-89、PDC-9、PDC-90、PDC-91、PDC-92、PDC-93、PEDO-1、PEDO-2、PEDO-3、PER-1、PER-2、PER-3、PER-4、PER-5、PER-6、PER-7、PNGM-1、赋予对壮观霉素耐药性的多杀性巴斯德氏菌 (*Pasteurella multocida*)

16S rRNA突变、赋予对抑制剂GE2270A耐药性的玫瑰游动双孢菌 (*Planobispora rosea*) EF-Tu突变体、PmpM、PmrF、具有赋予对大环内酯类抗生素耐药性的突变的丙酸杆菌23SrRNA、铜绿假单胞菌CpxR、铜绿假单胞菌catB6、铜绿假单胞菌catB7、铜绿假单胞菌emrE、赋予对氟喹诺酮类耐药性的铜绿假单胞菌gyrA和parC、赋予对氟喹诺酮类耐药性的铜绿假单胞菌gyrA、具有赋予对亚胺培南耐药性的突变的铜绿假单胞菌oprD、赋予对氟喹诺酮类耐药性的铜绿假单胞菌parE、铜绿假单胞菌soxR、赋予对粘菌素耐药性的假单胞菌突变体PhoP、赋予对粘菌素耐药性的假单胞菌突变体PhoQ、PvrR、QepA1、QepA2、QepA3、QepA4、QnrA1、QnrA2、QnrA3、QnrA4、QnrA5、QnrA6、QnrA7、QnrB1、QnrB10、QnrB11、QnrB12、QnrB13、QnrB14、QnrB15、QnrB16、QnrB17、QnrB18、QnrB19、QnrB2、QnrB20、QnrB21、QnrB22、QnrB23、QnrB24、QnrB25、QnrB26、QnrB27、QnrB28、QnrB29、QnrB3、QnrB30、QnrB31、QnrB32、QnrB33、QnrB34、QnrB35、QnrB36、QnrB37、QnrB38、QnrB4、QnrB40、QnrB41、QnrB42、QnrB43、QnrB44、QnrB45、QnrB46、QnrB47、QnrB48、QnrB49、QnrB5、QnrB50、QnrB54、QnrB55、QnrB56、QnrB57、QnrB58、QnrB59、QnrB6、QnrB60、QnrB61、QnrB62、QnrB64、QnrB65、QnrB66、QnrB67、QnrB68、QnrB69、QnrB7、QnrB70、QnrB71、QnrB72、QnrB73、QnrB74、QnrB8、QnrB9、QnrC、QnrD1、QnrD2、QnrS1、QnrS10、QnrS11、QnrS12、QnrS15、QnrS2、QnrS3、QnrS4、QnrS5、QnrS6、QnrS7、QnrS8、QnrS9、QnrVC1、QnrVC3、QnrVC4、QnrVC5、QnrVC6、QnrVC7、R39、RCP-1、ROB-1、RSA-1、RSA-2、RbpA、球形红细菌 (*Rhodobacter sphaeroides*) ampCB-内酰胺酶、*Rhodococcus fascians* cmr、RlmA (II)、Rm3、SAT-2、SAT-3、SAT-4、SFB-1、SFH-1、SHV-1、SHV-100、SHV-101、SHV-102、SHV-103、SHV-104、SHV-105、SHV-106、SHV-107、SHV-108、SHV-109、SHV-11、SHV-110、SHV-111、SHV-112、SHV-119、SHV-12、SHV-120、SHV-121、SHV-122、SHV-123、SHV-124、SHV-125、SHV-126、SHV-127、SHV-128、SHV-129、SHV-13、SHV-133、SHV-134、SHV-135、SHV-137、SHV-14、SHV-140、SHV-141、SHV-142、SHV-143、SHV-144、SHV-145、SHV-147、SHV-148、SHV-149、SHV-15、SHV-150、SHV-151、SHV-152、SHV-153、SHV-154、SHV-155、SHV-156、SHV-157、SHV-158、SHV-159、SHV-16、SHV-160、SHV-161、SHV-162、SHV-163、SHV-164、SHV-165、SHV-167、SHV-168、SHV-172、SHV-173、SHV-178、SHV-179、SHV-18、SHV-180、SHV-182、SHV-183、SHV-185、SHV-186、SHV-187、SHV-188、SHV-189、SHV-19、SHV-2、SHV-20、SHV-21、SHV-22、SHV-23、SHV-24、SHV-25、SHV-26、SHV-27、SHV-28、SHV-29、SHV-2A、SHV-3、SHV-30、SHV-31、SHV-32、SHV-33、SHV-34、SHV-35、SHV-36、SHV-37、SHV-38、SHV-39、SHV-40、SHV-41、SHV-42、SHV-43、SHV-44、SHV-45、SHV-46、SHV-48、SHV-49、SHV-5、SHV-50、SHV-51、SHV-52、SHV-53、SHV-55、SHV-56、SHV-57、SHV-59、SHV-6、SHV-60、SHV-61、SHV-62、SHV-63、SHV-64、SHV-65、SHV-66、SHV-67、SHV-69、SHV-7、SHV-70、SHV-71、SHV-72、SHV-73、SHV-74、SHV-75、SHV-76、SHV-77、SHV-78、SHV-79、SHV-8、SHV-80、SHV-81、SHV-82、SHV-83、SHV-84、SHV-85、SHV-86、SHV-89、SHV-9、SHV-92、SHV-93、SHV-94、SHV-95、SHV-96、SHV-97、SHV-98、SHV-99、SIM-1、SLB-1、SMB-1、SME-1、SME-2、SME-3、SME-4、SME-5、SPG-1、SPM-1、SRT-1、SRT-2、赋予对壮观霉素耐药性的肠道沙门氏菌16S rRNA (rrsD) 突变、肠道沙门氏菌cmlA、赋予对氟喹诺酮类耐药性的肠道沙门氏菌gyrA、具有赋予对三氯生耐药性的突变的肠道沙门氏菌gyrA、赋予对氟喹诺酮类耐药性的肠道沙门氏菌parC、肠道沙门氏菌ramR突变体、具有赋予抗生素耐药性的突变的肠道沙门氏菌soxR、赋予对氟喹诺酮类耐药性的沙门氏菌血清型gyrB、赋予对氟喹诺酮类耐药性的沙门氏菌血清型parE、具有赋予抗生素耐药性的突变的沙门氏菌血清型soxS、

Sed-1、粘质沙雷氏菌 (*Serratia marcescens*) Omp1、福氏志贺氏菌 (*Shigella flexneri*) 氯霉素乙酰转移酶、赋予对氟喹诺酮类耐药性的福氏志贺氏菌 *gyrA*、赋予对氟喹诺酮类耐药性的福氏志贺氏菌 *parC*、具有赋予对利奈唑胺耐药性的突变的金黄色葡萄球菌 23S rRNA、金黄色葡萄球菌 *FosB*、具有赋予对磷霉素耐药性的突变的金黄色葡萄球菌 *GlpT*、具有赋予对磷霉素耐药性的突变的金黄色葡萄球菌 *UhpT*、具有赋予对达托霉素耐药性的金黄色葡萄球菌 *agrA*、赋予对达托霉素耐药性的金黄色葡萄球菌 *cls*、具有赋予对夫西地酸耐药性的突变的金黄色葡萄球菌 *fusA*、具有赋予对夫西地酸耐药性的突变的金黄色葡萄球菌 *fusE*、赋予对氟喹诺酮类耐药性的金黄色葡萄球菌 *gyrA*、赋予对氨基香豆素耐药性的金黄色葡萄球菌 *gyrB*、具有赋予对莫匹罗星耐药性的突变的金黄色葡萄球菌 *ileS*、具有赋予对 *lysocin* 耐药性的突变的金黄色葡萄球菌 *menA*、金黄色葡萄球菌 *mprF*、具有赋予对达托霉素耐药性的突变的金黄色葡萄球菌 *mprF*、具有赋予对磷霉素耐药性的突变的金黄色葡萄球菌 *murA*、金黄色葡萄球菌 *norA*、赋予对氟喹诺酮类耐药性的金黄色葡萄球菌 *parC*、赋予对氨基香豆素耐药性的金黄色葡萄球菌 *parE*、赋予对氟喹诺酮类耐药性的金黄色葡萄球菌 *parE*、赋予对达托霉素耐药性的金黄色葡萄球菌 *pgsA* 突变、赋予对达托霉素耐药性的金黄色葡萄球菌 *rpoB* 突变体、赋予对利福平耐药性的金黄色葡萄球菌 *rpoB* 突变体、赋予对达托霉素耐药性的金黄色葡萄球菌 *rpoC*、具有赋予对达托霉素耐药性的突变的金黄色葡萄球菌 *walK*、中间葡萄球菌 (*Staphylococcus intermedius*) 氯霉素乙酰转移酶、赋予对莫匹罗星耐药性的葡萄球菌 *mupA*、赋予对莫匹罗星耐药性的葡萄球菌 *mupB*、金黄色葡萄球菌 *LmrS*、无乳链球菌 (*Streptococcus agalactiae*) *mprF*、具有赋予达托霉素耐药性的突变的轻链球菌 (*Streptococcus mitis*) *CdsA*、赋予对大环内酯类和链阳菌素类抗生素耐药性的肺炎链球菌 23S rRNA 突变、具有赋予对大环内酯类抗生素耐药性的突变的肺炎链球菌 23S rRNA、赋予对阿莫西林耐药性的肺炎链球菌 *PBP1a*、赋予对阿莫西林耐药性的肺炎链球菌 *PBP2b*、赋予对阿莫西林耐药性的肺炎链球菌 *PBP2x*、赋予对氟喹诺酮类耐药性的肺炎链球菌 *parC*、具有赋予对磺胺类耐药性的突变的化脓性链球菌 (*Streptococcus pyogenes*) *folP*、猪链球菌 (*Streptococcus suis*) 氯霉素乙酰转移酶、具有赋予对大环内酯类抗生素耐药性的突变的生二素链霉菌 (*Streptomyces ambofaciens*) 23S rRNA、赋予对 *elfamycin* 耐药性的肉桂链霉菌 (*Streptomyces cinnamoneus*) *EF-Tu* 突变体、变青链霉菌 (*Streptomyces lividans*) *cmlR*、赋予对氨基香豆素耐药性的利尻链霉菌 (*Streptomyces rishiriensis*) *parY* 突变体、TEM-1、TEM-10、TEM-101、TEM-102、TEM-104、TEM-105、TEM-106、TEM-107、TEM-108、TEM-109、TEM-11、TEM-110、TEM-111、TEM-112、TEM-113、TEM-114、TEM-115、TEM-116、TEM-117、TEM-118、TEM-12、TEM-120、TEM-121、TEM-122、TEM-123、TEM-124、TEM-125、TEM-126、TEM-127、TEM-128、TEM-129、TEM-130、TEM-131、TEM-132、TEM-133、TEM-134、TEM-135、TEM-136、TEM-137、TEM-138、TEM-139、TEM-141、TEM-142、TEM-143、TEM-144、TEM-145、TEM-146、TEM-147、TEM-148、TEM-149、TEM-15、TEM-150、TEM-151、TEM-152、TEM-153、TEM-154、TEM-155、TEM-156、TEM-157、TEM-158、TEM-159、TEM-16、TEM-160、TEM-162、TEM-163、TEM-164、TEM-166、TEM-167、TEM-168、TEM-169、TEM-17、TEM-171、TEM-176、TEM-177、TEM-178、TEM-182、TEM-183、TEM-184、TEM-185、TEM-186、TEM-187、TEM-188、TEM-189、TEM-19、TEM-190、TEM-191、TEM-192、TEM-193、TEM-194、TEM-195、TEM-196、TEM-197、TEM-198、TEM-199、TEM-2、TEM-20、TEM-201、TEM-205、TEM-206、TEM-207、TEM-208、TEM-209、TEM-21、TEM-211、TEM-213、TEM-

214、TEM-215、TEM-216、TEM-217、TEM-219、TEM-22、TEM-220、TEM-24、TEM-26、TEM-28、TEM-29、TEM-3、TEM-30、TEM-33、TEM-34、TEM-4、TEM-40、TEM-42、TEM-43、TEM-45、TEM-47、TEM-48、TEM-49、TEM-52、TEM-53、TEM-54、TEM-55、TEM-57、TEM-59、TEM-6、TEM-60、TEM-63、TEM-67、TEM-68、TEM-7、TEM-70、TEM-71、TEM-72、TEM-73、TEM-75、TEM-76、TEM-78、TEM-79、TEM-8、TEM-80、TEM-81、TEM-82、TEM-83、TEM-84、TEM-85、TEM-86、TEM-87、TEM-88、TEM-89、TEM-90、TEM-91、TEM-92、TEM-93、TEM-94、TEM-95、TEM-96、THIN-B、TLA-1、TLA-2、TLA-3、TMB-1、TMB-2、TRU-1、TUS-1、TaeA、Tet (47) 、Tet (X3) 、Tet (X4) 、TolC、TriA、TriB、TriC、A型NfxB、B型NfxB、赋予对氟喹诺酮类耐药性的解脲脲原体 (*Ureaplasma urealyticum*) gyrB、赋予对氟喹诺酮类耐药性的解脲脲原体parC、VCC-1、VEB-1、VEB-1b、VEB-2、VEB-3、VEB-4、VEB-5、VEB-6、VEB-7、VEB-8、VEB-9、VIM-1、VIM-10、VIM-11、VIM-12、VIM-13、VIM-14、VIM-15、VIM-16、VIM-17、VIM-18、VIM-19、VIM-2、VIM-20、VIM-23、VIM-24、VIM-25、VIM-26、VIM-27、VIM-28、VIM-29、VIM-3、VIM-30、VIM-31、VIM-32、VIM-33、VIM-34、VIM-35、VIM-36、VIM-37、VIM-38、VIM-39、VIM-4、VIM-42、VIM-43、VIM-5、VIM-6、VIM-7、VIM-8、VIM-9、VatI、鳃弧菌 (*Vibrio anguillarum*) 氯霉素乙酰转移酶、霍乱弧菌 (*Vibrio cholerae*) OmpT、霍乱弧菌 OmpU、霍乱弧菌 varG、YojI、aacA43、aad (6) 、aadA、aadA10、aadA11、aadA12、aadA13、aadA14、aadA15、aadA16、aadA17、aadA2、aadA21、aadA22、aadA23、aadA24、aadA25、aadA27、aadA3、aadA4、aadA5、aadA6、aadA6/aadA10、aadA7、aadA8、aadA8b、aadA9、aadK、aadS、abcA、abeM、abeS、acrB、acrD、adeA、adeB、adeC、adeF、adeG、adeH、adeI、adeJ、adeK、adeL、adeN、adeR、adeS、almG、ampS、amrA、amrB、aphA15、apmA、arlR、arlS、armA、arnA、arr-1、arr-2、arr-3、arr-4、arr-5、arr-7、arr-8、bacA、baeR、baeS、basR、basS、bcr-1、bcrA、bcrB、bcrC、blaF、blaI、blaR1、blt、bmr、carA、carO、catA4、catA8、catB10、catB11、catB2、catB3、catB8、catB9、catI、catII、来自大肠杆菌K-12的catII、catIII、catP、catQ、catS、catV、cdeA、ceoA、ceoB、cepA、cfr (B) 、cfrA、cfrC、chrB、cipA、clbA、clbB、clbC、clcD、cmeA、cmeB、cmeC、cmeR、cmlA1、cmlA4、cmlA5、cmlA6、cmlA8、cmlB、cmlB1、cmlv、cmrA、cmx、cpaA、cphA2、cphA3、cphA4、cphA5、cphA6、cphA7、cphA8、cpxA、dfrA1、dfrA10、dfrA12、dfrA13、dfrA14、dfrA15、dfrA15b、dfrA16、dfrA17、dfrA18、dfrA19、dfrA20、dfrA21、dfrA22、dfrA23、dfrA24、dfrA25、dfrA26、dfrA27、dfrA28、dfrA29、dfrA3、dfrA30、dfrA32、dfrA3b、dfrA5、dfrA6、来自奇异变形杆菌 (*Proteus mirabilis*) 的dfrA6、dfrA7、dfrA8、dfrA9、dfrB1、dfrB2、dfrB3、dfrB4、dfrB5、dfrB6、dfrB7、dfrC、dfrD、dfrE、dfrF、dfrG、dfrI、dfrK、eatA、efmA、efpA、efrA、efrB、emeA、emrA、emrB、emrD、emrK、emrR、emrY、emtA、eptA、erm (32) 、erm (40) 、erm (45) 、erm (46) 、ermZ、evgA、evgS、facT、farA、farB、fexA、floR、fusB、fusC、fusD、fusH、gadW、gadX、gimA、golS、hmrM、hp1181、hp1184、imiH、imiS、iri、kamB、kdpE、lfrA、lin、linG、lmrA、lmrB、lmrC、lmrD、lmrP、lnuA、lnuB、lnuC、lnuD、lnuE、lnuF、lnuG、lsaA、lsaB、lsaC、lsaE、macA、macB、marA、mdsA、mdsB、mdsC、mdtA、mdtB、mdtC、mdtE、mdtF、mdtG、mdtH、mdtM、mdtN、mdtO、mdtP、mecA、mecB、mecC、mecD、mecI、mecR1、mef (B) 、mefC、mefE、mel、mepA、mepR、mexM、mexN、mexP、mexQ、mexX、mexY、mfd、mfpA、mgrA、mgrB、mgtA、mphA、mphB、mphC、mphE、mphF、mphG、mphH、mphI、mphJ、mphK、mphL、mphM、mphN、mphO、msbA、msrA、msrB、msrC、msrE、mtrA、mtrC、mtrD、mtrE、mtrR、myrA、nalC、nalD、norA、norB、novA、npmA、oleB、oleC、oleD、oleI、opcM、opmE、optrA、oqx A、oqx B、otr (A) 、otr (B) 、otrC、patA、patB、pexA、pgpB、

plasmid-encoded cat(pp-cat)、pmrA、porin OmpC、poxtA、pp-flo、qacA、qacB、qacH、qnrE1、qnrE2、ramA、rgt1438、rmtA、rmtB、rmtC、rmtD、rmtD2、rmtE、rmtE2、rmtF、rmtG、rmtH、rosA、rosB、rphA、rphB、rpoB2、rpsJ、salA、sav1866、sdiA、sgm、smeA、smeB、smeC、smeD、smeE、smeF、smeR、smeS、spd、srmB、sta、sul1、sul2、sul3、sul4、tap、tcmA、tcr3、tet(30)、tet(31)、tet(33)、tet(35)、tet(38)、tet(39)、tet(40)、tet(41)、tet(42)、tet(43)、tet(44)、tet(45)、tet(48)、tet(49)、tet(50)、tet(51)、tet(52)、tet(53)、tet(54)、tet(55)、tet(56)、tet(59)、tet(A)、tet(B)、tet(C)、tet(D)、tet(E)、tet(G)、tet(H)、tet(J)、tet(K)、tet(L)、tet(V)、tet(W/N/W)、tet(Y)、tet(Z)、tet32、tet34、tet36、tet37、tetA(46)、tetA(58)、tetA(60)、tetA(P)、tetB(46)、tetB(58)、tetB(60)、tetB(P)、tetM、tetO、tetQ、tetR、tetR(G)、tetS、tetT、tetU、tetW、tetX、赋予泰乐菌素耐药性的tlrB、tlrC、tmrB、tsnR、tva(A)、ugd、vanA、vanB、vanC、vanD、vanE、vanF、vanG、vanHA、vanHB、vanHD、vanHF、vanHM、vanHO、vanI、vanJ、vanKI、vanL、vanM、vanN、vanO、vanRA、vanRB、vanRC、vanRD、vanRE、vanRF、vanRG、vanRI、vanRL、vanRM、vanRN、vanRO、vanSA、vanSB、vanSC、vanSD、vanSE、vanSF、vanSG、vanSI、vanSL、vanSM、vanSN、vanSO、vanTC、vanTE、vanTG、vanTN、vanTmL、vanTrL、vanUG、vanVB、vanWB、vanWG、vanWI、vanXA、vanXB、vanXD、vanXF、vanXI、vanXM、vanXO、vanXYC、vanXYE、vanXYG、vanXYL、vanXYN、vanYA、vanYB、vanYD、vanYF、vanYG1、vanYM、vanZA、vanZF、vatA、vatB、vatC、vatD、vatE、vatF、vatH、vga(E) 科氏葡萄球菌(*Staphylococcus cohnii*)、vgaA、vgaALC、vgaB、vgaC、vgaD、vgaE、vgbA、vgbB、vgbC、vmlR、vph、y56β-内酰胺酶、ykkC、ykkD或其任意组合结合(或被其结合)的富集分子。

[0260] 在一些实施方案中,试剂盒中的至少一种组分以干燥或冻干形式提供。在其他实施方案中,试剂盒的至少一种组分以溶解的形式提供。

[0261] 本文提供的试剂盒采用合适的包装。合适的包装包括但不限于小瓶、瓶子、罐子、软包装等。还考虑了与特定装置结合使用的包装。参见“用于样品制备和样品测序的装置”。试剂盒可以具有无菌入口(例如,容器可以是具有可被皮下注射针刺穿的塞子的静脉注射溶液袋或小瓶)。容器还可以具有无菌入口。

[0262] 试剂盒任选地可以提供另外的部件,例如缓冲液和解释信息。在一些实施方案中,试剂盒进一步包含至少一种缓冲液。适用于本文所述的方法的缓冲液先前已描述。在一些实施方案中,试剂盒可以另外包含在本文所述的任何方法中使用的说明。

[0263] 在一些实施方案中,本公开提供了包含上述试剂盒内容物的制品。

[0264] V. 用于样品制备和样品测序的装置

[0265] 在一些方面,本公开涉及用于样品制备和/或样品测序的装置。在一些实施方案中,所述装置包括样品制备模块。在一些实施方案中,所述装置包括样品测序模块。在一些实施方案中,所述装置包括样品制备模块和样品测序模块。

[0266] A. 用于样品制备的装置

[0267] 通常提供在制备用于分析的样品的过程中使用的包括装置、盒(例如,包含通道(例如,微流体通道))和/或泵(例如,蠕动泵)的装置。根据本公开,可以使用装置来使来自生物样品的靶分子的富集、浓缩、操作和/或检测成为可能。在一些实施方案中,提供了用于自动处理样品以产生用于下一代测序和/或其他下游分析技术的材料的装置和相关方法。装置和相关方法可用于执行化学和/或生物反应,包括根据本文别处描述的样品制备或样

品分析过程的核酸和/或多肽加工反应。

[0268] 在一些实施方案中,设置样品制备装置以将靶分子或包含多个分子(例如,靶核酸或靶多肽)的样品递送或转移至测序模块或装置。在一些实施方案中,样品制备装置直接连接到(例如,物理连接到)或间接连接到测序装置。

[0269] 在一些实施方案中,装置包括序列制备模块,其被设置为接收一个或多个盒。在一些实施方案中,盒包括一个或多个储存器或反应容器,其被设置为接收流体和/或包含在样品制备过程中使用的一种或多种试剂。在一些实施方案中,盒包含一个或多个通道(例如,微流体通道),其被设置为容纳和/或传输在样品制备过程中使用的流体(例如,包含一种或多种试剂的流体)。试剂包括缓冲液、酶促试剂、聚合物基质、富集分子、捕获试剂、尺寸特异性选择试剂、序列特异性选择试剂和/或纯化试剂。在样品制备过程中使用的其他试剂在本文别处描述。

[0270] 在一些实施方案中,盒包括一种或多种(例如,适合重构为液体形式的液体或冻干形式的)储存试剂。盒的储存试剂包括适合于执行所需过程的试剂和/或适合于处理所需样品类型的试剂。在一些实施方案中,盒是单次使用的盒(例如,一次性盒)或多次使用的盒(例如,可重复使用的盒)。在一些实施方案中,盒被设置为接收用户提供的样品。用户提供的样品可以在装置接收所述盒之前或之后被添加到所述盒中,例如,由用户手动或以自动化过程。

[0271] 在一些实施方案中,所述装置可有助于在根据本公开的方法中靶分子的富集。参见“多肽富集方法”。以这种方式,所述装置能够利用分子以高度多重的方式富集目的多肽。

[0272] 在一些实施方案中,使用电泳方法富集样品中的靶分子。在一些实施方案中,使用亲和力和SCODA富集样品中的靶分子。在一些实施方案中,使用反转场凝胶电泳(FIGE)富集样品中的靶分子。在一些实施方案中,使用脉冲场凝胶电泳(PFGE)富集样品中的靶分子。

[0273] 在一些实施方案中,装置包含样品制备模块,所述样品制备模块包含在富集过程中使用的基质(例如,多孔介质、电泳聚合物凝胶),所述基质包含固定化的捕获探针,所述捕获探针与样品中存在的靶分子(直接或间接)结合。在一些实施方案中,在富集过程中使用的基质包含1、2、3、4、5或更多个独特的固定化的捕获探针,每个探针结合独特靶分子和/或以不同的结合亲和力结合相同的靶分子。

[0274] 在一些实施方案中,固定化的捕获探针是与靶多肽或多肽片段结合的多肽捕获探针。例如,在一些实施方案中,固定化的捕获探针是如本文所述的富集分子。

[0275] 在一些实施方案中,多肽捕获探针以 10^{-9} 至 10^{-8} M、 10^{-8} 至 10^{-7} M、 10^{-7} 至 10^{-6} M、 10^{-6} 至 10^{-5} M、 10^{-5} 至 10^{-4} M、 10^{-4} 至 10^{-3} M或 10^{-3} 至 10^{-2} M的结合亲和力与靶多肽(或多肽片段)结合。在一些实施方案中,结合亲和力在皮摩尔至纳摩尔范围内(例如,在约 10^{-12} 和约 10^{-9} M之间)。在一些实施方案中,结合亲和力在纳摩尔至微摩尔范围内(例如,在约 10^{-9} 和约 10^{-6} M之间)。在一些实施方案中,结合亲和力在微摩尔至毫摩尔范围内(例如,在约 10^{-6} 和约 10^{-3} M之间)。在一些实施方案中,结合亲和力在皮摩尔至微摩尔范围内(例如,在约 10^{-12} 和约 10^{-6} M之间)。在一些实施方案中,结合亲和力在纳摩尔至毫摩尔范围内(例如,在约 10^{-9} 和约 10^{-3} M之间)。

[0276] 在一些实施方案中,固定化的捕获探针是与靶核酸杂交的寡核苷酸捕获探针。在一些实施方案中,寡核苷酸捕获探针与靶核酸至少50%、60%、70%、80%、90%、95%或100%互补。在一些实施方案中,单个寡核苷酸捕获探针可用于富集具有至少50%、60%、

70%、80%、90%、95%或99%的序列同一性的多个相关靶核酸(例如,2、3、4、5、6、7、8、9、10、20、30、40、50或更多个相关靶核酸)。多种相关靶核酸的富集可以允许产生宏基因组文库。在一些实施方案中,寡核苷酸捕获探针可以实现相关靶核酸的差异富集。在一些实施方案中,寡核苷酸捕获探针可以实现靶核酸相对于与其修饰状态(例如甲基化状态、乙酰化状态)不同的相同序列核酸的富集。

[0277] 在一些实施方案中,为了富集长度为0.5-2k碱基的核酸靶分子,寡核苷酸捕获探针可以使用5' Acrydite部分被共价固定在丙烯酰胺基质中。在一些实施方案中,为了富集更大的核酸靶分子(例如,长度>2k碱基),寡核苷酸捕获探针可以被固定在琼脂糖基质中。在一些实施方案中,寡核苷酸捕获探针可以使用硫醇-环氧化物化学(例如,通过将硫醇修饰的寡核苷酸共价连接到交联的琼脂糖珠粒)固定在琼脂糖基质中。连接到琼脂糖珠粒的寡核苷酸捕获探针可以在标准琼脂糖基质(例如,以相同的琼脂糖百分比)中结合和固化。

[0278] 在一些实施方案中,多种捕获探针(例如,多种捕获探针类型的群体,例如,与诸如腺病毒、葡萄球菌、肺炎或结核病等传染原的确定性靶分子结合的群体)可以被固定在富集基质中。将样品应用于具有多个确定性捕获探针的富集基质可能会导致疾病或病况的诊断(例如,存在传染原)。

[0279] 在一些实施方案中,在根据本公开的方法中,在去除非靶分子之后,装置可以有助于靶分子从富集基质中的释放。在一些实施方案中,可以通过增加富集基质的温度从富集基质中释放靶分子。调整基质的温度会进一步影响迁移速率,因为升高的温度会提供更高的捕获探针严格性,从而需要靶分子和捕获探针之间更大的结合亲和力。在一些实施方案中,在富集相关靶分子时,可以逐步提高基质温度,从而以逐步增加同源性来释放和分离靶分子。这可以允许对与初始参考靶分子的关系越来越远的靶多肽或靶核酸进行测序,从而能够发现新的蛋白质(例如,酶)或功能(例如,酶促功能或基因功能)。在一些实施方案中,当使用多个捕获探针(例如,多个确定性捕获探针)时,基质温度可以逐步或以梯度方式增加,允许不同靶分子的温度依赖性释放并导致产生一系列条形码释放带,其代表存在或不存在控制分子和目标分子。

[0280] 根据本公开的装置通常包含机械和电子和/或光学部件,其可用于操作如本文所述的盒。在一些实施方案中,装置部件运行以在盒上或在盒的特定区域上实现和维持特定温度。在一些实施方案中,装置部件运行以向盒的电极施加特定长时的特定电压。在一些实施方案中,装置部件运行以将液体移入、移出盒的储存器和/或反应容器或移至盒的储存器和/或反应容器之间。在一些实施方案中,装置部件运行以将液体移动通过盒的通道,例如,移入、移出盒的储存器和/或反应容器或移至盒的储存器和/或反应容器之间。在一些实施方案中,装置部件通过与盒的弹性体、试剂特异性储存器或反应容器相互作用的蠕动泵送机构(例如,装置)移动液体。在一些实施方案中,装置部件通过蠕动泵送机构(例如,装置)移动液体,所述蠕动泵送机构被设置为与盒的通道相关联的弹性体部件(例如,包括弹性体的表面层)相互作用以泵送流体通过通道。装置部件可以包括计算机资源,例如,用于驱动可以输入样品信息、可以选择特定过程以及可以报告运行结果的用户界面。

[0281] 以下非限制性实例旨在说明本文所述的装置、方法和组合物的方面。根据本公开的样品制备装置的使用可以进行以下描述的步骤中的一个或多个。用户可以打开装置的盖子并插入支持所需过程的盒。然后,用户可以将可以与特定裂解溶液结合的样品添加到盒

上的样品端口。然后,用户可以关闭装置盖子,通过装置上的触摸屏界面输入任何样品特定信息,选择任何过程特定参数(例如,所需大小选择的范围、靶分子捕获所需的同源度等),并启动样品制备过程运行。

[0282] 运行后,用户可能会收到相关的运行数据(例如,运行成功完成的确认、运行特定指标等)以及过程特定信息(例如,生成的样本量、特定靶序列的存在或不存在等)。通过运行生成的数据可以进行后续的生物信息学分析,该分析可以是本地的或基于云的。根据进程,可以从盒中提取完成的样本以供后续使用(例如,基因组测序、qPCR定量、克隆等)。然后可以打开该装置,然后可以取出盒。

[0283] 图8提供了描绘用于执行富集的示例性装置的图示。参见例如美国专利号8608929,其全部内容通过引用的方式并入本文。

[0284] B. 测序装置

[0285] 通常还提供包括在对包含多肽的样品(例如,富集样品)进行测序的过程中使用的装置、盒(例如,包含通道(例如,微流体通道))和/或泵(例如,蠕动泵)的装置。在一些方面,可以使用允许单分子分析和/或单分子测序并行的系统来进行根据本公开的核酸或多肽的测序。所述系统可以包括测序装置和被设置为与测序装置接口的仪器。

[0286] 测序装置可以包括包含像素阵列的测序模块,其中各个像素包括样品孔和至少一个光检测器。测序装置的样品孔可以形成在测序装置的表面上或穿过测序装置的表面,并且被设置为接收放置在测序装置表面上的样品。在一些实施方案中,样品孔是可以被插入装置中的盒(例如,一次性或单次使用的盒)的部件。总的来说,样品孔可以被认为是样品孔的阵列。多个样品孔可以具有合适的尺寸和形状,使得样品孔的至少一部分接收单个靶分子或包含多个分子(例如,靶核酸或靶多肽)的样品。在一些实施方案中,样品孔内的分子数量可以分布在测序装置的样品孔中,使得一些样品孔包含一个分子(例如,靶核酸或靶多肽),而其他样品孔包含零个、两个或更多个分子。

[0287] 在一些实施方案中,测序装置设置在从样品制备装置接收包含多个分子(例如,一种或多种目的多肽)的样品的位置。在一些实施方案中,测序装置直接连接(例如,物理连接到)或间接连接到样品制备装置。

[0288] 测序装置可以包括像素阵列,其中各个像素包括样品孔和至少一个光检测器。测序装置的样品孔可以形成在测序装置的表面上或穿过测序装置的表面,并且被设置为接收放置在测序装置表面上的样品。总的来说,样品孔可以被认为是样品孔的阵列。多个样品孔可以具有合适的尺寸和形状,使得样品孔的至少一部分接收单个样品(例如,单个分子,例如多肽)。在一些实施方案中,样品孔内的样品数量可以分布在测序装置的样品孔中,使得一些样品孔包含一个样品,而其他样品孔包含零个、两个或更多个样品。

[0289] 从一个或多个光源向测序装置提供激发光,所述光源可以在测序装置的外部或内部。测序装置的光学部件可以接收来自光源的激发光并将光引导到测序装置的样品孔阵列并照亮样品孔内的照明区域。在一些实施方案中,样品孔可以具有允许样品保持在样品孔表面附近的构造,这可以容易地将激发光递送到样品和检测来自样品的发射光。位于照明区域内的样品可以响应于被激发光照亮而发射光。例如,可以用荧光标志物标记样品,所述荧光标志物响应于通过激发光的照射实现激发态而发射光。由样品发射的发射光然后可以由对应于样品孔的像素内的一个或多个光检测器检测,其中样品被分析。根据一些实施方

案,当在数量范围可以在大约10,000像素到1,000,000像素之间的样品孔阵列上执行时,可以并行分析多个样品。

[0290] 测序装置可以包括用于接收激发光并将激发光引导到样品孔阵列之间的光学系统。光学系统可以包括一个或多个被设置为将激发光耦合到测序装置并将激发光引导到其他光学部件的光栅耦合器。光学系统可以包括将来自光栅耦合器的激发光引导到样品孔阵列的光学部件。这样的光学部件可以包括分光器、光学组合器和波导。在一些实施方案中,一个或多个分光器可以耦合来自光栅耦合器的激发光并将激发光递送到至少一个波导。根据一些实施方案,分光器可以具有允许激发光在所有波导上基本均匀地传递的构造,使得每个波导接收基本相似量的激发光。这样的实施方案可以通过提高测序装置的样品孔接收的激发光的均匀性来提高测序装置的性能。例如,用于将激发光耦合到样品孔和/或将发射光引导到光检测器以包括在测序装置中的合适部件的实例在2015年8月7日提交的标题为“INTEGRATED DEVICE FOR PROBING, DETECTING AND ANALYZING MOLECULES”的美国专利申请号14/821,688以及2014年11月17日提交的标题为“INTEGRATED DEVICE WITH EXTERNAL LIGHT SOURCE FOR PROBING, DETECTING, AND ANALYZING MOLECULES”的美国专利申请号14/543,865中进行了描述,两者的全部内容均通过引用的方式并入本文。可以在测序装置中实施的合适的光栅耦合器和波导的实例在2017年12月15日提交的标题为“OPTICAL COUPLER AND WAVEGUIDE SYSTEM”的美国专利申请号15/844,403中进行了描述,其全部内容通过引用的方式并入本文。

[0291] 另外的光激性结构可以定位在样品孔和光检测器之间,并且被设置为减少或防止激发光到达光检测器,否则这可能会导致检测发射光时的信号噪声。在一些实施方案中,可以充当测序装置的电路的金属层也可以充当空间滤光器。合适的光激性结构的实例可以包括光谱滤光器、偏振滤光器和空间滤光器,并且在2018年7月23日提交的标题为“OPTICAL REJECTION PHOTONIC STRUCTURES”的美国专利申请号16/042,968中进行了描述,其全部内容通过引用的方式并入本文。

[0292] 位于测序装置之外的部件可用于将激发源定位和对准到测序装置。这样的部件可以包括光学部件,包括透镜、镜子、棱镜、窗口、孔径、衰减器和/或光纤。仪器中可以包括另外的机械部件,以允许控制一个或多个对准部件。这样的机械部件可以包括致动器、步进电机和/或旋钮。合适的激发源和对准机构的实例在2016年5月20日提交的标题为“PULSED LASER AND SYSTEM”的美国专利申请号15/161,088中进行了描述,其全部内容通过引用的方式并入本文。光束控制模块的另一个实例在2017年12月14日提交的标题为“COMPACT BEAM SHAPING AND STEERING ASSEMBLY”的美国专利申请号15/842,720中进行了描述,其通过引用的方式并入本文。合适的激发源的另外的实例在2015年8月7日提交的标题为“INTEGRATED DEVICE FOR PROBING, DETECTING AND ANALYZING MOLECULES”的美国专利申请号14/821,688中进行了描述,其全部内容通过引用的方式并入本文。

[0293] 与测序装置的单个像素一起定位的光检测器可以被设置和定位以检测来自像素的相应样品孔的发射光。合适的光检测器的实例在2015年8月7日提交的标题为“INTEGRATED DEVICE FOR TEMPORAL BINNING OF RECEIVED PHOTONS”的美国专利申请号14/821,656中进行了描述,其全部内容通过引用的方式并入本文。在一些实施方案中,样品孔及其相应的光检测器可以沿着公共轴线对齐。以这种方式,光检测器可以与像素内的样

本孔重叠。

[0294] 检测到的发射光的特性可以提供用于鉴定与发射光相关的标志物的指示。这样的特性可以包括任何合适类型的特性,包括由光检测器检测到的光子的到达时间、由光检测器随时间累积的光子量和/或跨两个或更多个光检测器的光子分布。在一些实施方案中,光检测器可以具有允许检测与样品的发射光(例如,发光寿命)相关的一个或多个时序特性的构造。在激发光脉冲传播通过测序装置之后,光检测器可以检测光子到达时间的分布,并且到达时间的分布可以提供样品发射光的时序特性的指示(例如,发光寿命的代表)。在一些实施方案中,一个或多个光检测器提供由标志物发射的发射光的概率(例如,发光强度)的指示。在一些实施方案中,多个光检测器的尺寸和布置可被设置为捕获发射光的空间分布。来自一个或多个光检测器的输出信号然后可用于将标志物与多个标志物区分开来,其中多个标志物可用于鉴定样品内的样品。在一些实施方案中,样品可以被多种激发能量激发,并且样品响应于多种激发能量而发射的发射光和/或发射光的时序特性可以将标志物与多个标志物区分开来。

[0295] 在操作中,样品孔内的样品的并行分析是通过使用激发光激发孔内的一些或所有样品并用光检测器检测来自样品发射的信号来进行的。来自样品的发射光可以由相应的光检测器检测并转换为至少一个电信号。电信号可以沿着测序装置的电路中的导线传输,所述导线可以连接到与测序装置接口的仪器。随后可以处理和/或分析电信号。电信号的处理或分析可以在位于仪器上或仪器外的合适的计算设备上进行。

[0296] 所述仪器可以包括用于控制仪器和/或测序装置的操作的用户界面。用户界面可以被设置为允许用户将信息输入到仪器中,例如用于控制仪器功能的命令和/或设置。在一些实施方案中,用户界面可以包括用于语音命令的按钮、开关、拨号盘和麦克风。用户界面可以允许用户接收关于仪器和/或测序装置的性能的反馈,例如同轴度(proper alignment)和/或通过来自测序装置上的光检测器读出信号获得的信息。在一些实施方案中,用户界面可以使用扬声器提供听觉反馈来提供反馈。在一些实施方案中,用户界面可以包括用于向用户提供视觉反馈的指示灯和/或显示屏。

[0297] 在一些实施方案中,所述仪器可以包括被设置为与计算设备连接的计算机接口。计算机接口可以是USB接口、火线接口或任何其他合适的计算机接口。计算设备可以是任何通用计算机,例如膝上型计算机或台式计算机。在一些实施方案中,计算设备可以是经由合适的计算机接口在无线网络上可访问的服务器(例如,基于云的服务器)。计算机接口可以促进仪器和计算设备之间的信息通信。用于控制和/或配置仪器的输入信息可以被提供给计算设备并通过计算机接口传输给仪器。由仪器生成的输出信息可以通过计算机接口由计算设备接收。输出信息可以包括关于仪器性能、测序装置性能和/或从光检测器的读出信号产生的数据的反馈。

[0298] 在一些实施方案中,所述仪器可以包括被设置为分析从测序装置的一个或多个光检测器接收的数据和/或将控制信号传输到激发源的处理装置。在一些实施方案中,处理装置可以包括通用处理器、专门适配的处理器(例如,中央处理单元(CPU),例如一个或多个微处理器或微控制器内核、现场可编程门阵列(FPGA)、专用集成电路(ASIC)、定制集成电路、数字信号处理器(DSP)或其组合)。在一些实施方案中,来自一个或多个光检测器的数据的处理可以由仪器的处理装置和外部计算设备两者来执行。在其他实施方案中,可以省略外

部计算设备,并且可以仅由测序装置的处理装置执行来自一个或多个光检测器的数据处理。

[0299] 根据一些实施方案,被设置为基于发光发射特性来分析样品的仪器可以检测不同发光分子之间的发光寿命和/或强度的差异,和/或相同发光分子在不同环境中的寿命和/或强度之间的差异。发明人已经认识到并理解,发光发射寿命的差异可用于辨别不同发光分子的存在与否和/或辨别发光分子所经受的不同环境或条件。在一些情况下,根据寿命(例如,而不是发射波长)辨别发光分子可以简化系统的方面。作为实例,当基于寿命辨别发光分子时,波长区分光学器件(例如波长过滤器、每个波长的专用检测器、不同波长的专用脉冲光源和/或衍射光学器件)可以在数量上减少或被消除。在一些情况下,以单一特征波长操作的单一脉冲光源可用于激发在光谱的相同波长区域内发射但具有可测量的不同寿命的不同发光分子。使用单个脉冲光源而不是在不同波长下工作的多个光源来激发和辨别在相同波长范围内发射的不同发光分子的分析系统操作和维护的复杂度更低,更紧凑,并且可以以更低成本制造。

[0300] 尽管基于发光寿命分析的分析系统可能具有某些好处,但通过允许另外的检测技术可以增加由分析系统获得的信息量和/或检测精度。例如,系统的一些实施方案可以另外被设置成基于发光波长和/或发光强度来辨别样品的一种或多种特性。在一些实施方式中,发光强度可以另外地或替代地用于区分不同的发光标记。例如,一些发光标记可以以显著不同的强度发射或在它们的激发概率上有显著差异(例如,至少约35%的差异),即使它们的衰减率可能相似。通过将分箱信号参考测量的激发光,可以根据强度水平区分不同的发光标记。

[0301] 根据一些实施方案,不同的发光寿命可以用被设置为在发光标记激发之后对发光发射事件进行时间分箱(time-bin)的光检测器来区分。时间分箱可以发生在光检测器的单个电荷累积周期期间。电荷累积周期是读出事件之间的间隔,在该期间光生载流子累积在时间分箱的光检测器的仓中。时间分箱的光检测器的实例在2015年8月7日提交的标题为“INTEGRATED DEVICE FOR TEMPORAL BINNING OF RECEIVED PHOTONS”的美国专利申请号14/821,656中进行了描述,其通过引用的方式并入本文。在一些实施方案中,时间分箱的光检测器可以在光子吸收/载流子产生区域中产生电荷载流子并且将电荷载流子直接转移到电荷载流子存储仓中的电荷载流子存储仓。在这样的实施方案中,时间分箱的光检测器可以不包括载流子行进/捕获区域。这样的时间分箱的光检测器可以被称为“直接分箱像素”。包括直接分箱像素的时间分箱的光检测器的实例在2017年12月22日提交的标题为“INTEGRATED PHOTODETECTOR WITH DIRECT BINNING PIXEL”的美国专利申请号15/852,571中进行了描述,其通过引用的方式并入本文。

[0302] 在一些实施方案中,相同类型的不同数量的荧光团可以与样品中的不同试剂连接,从而可以基于发光强度来鉴定每种试剂。例如,两个荧光团可以连接到第一标记的亲试剂,四个或更多个荧光团可以连接到第二标记的亲试剂。由于不同数量的荧光团,可能存在与不同亲试剂相关的不同激发和荧光团发射概率。例如,在信号累积间隔期间,第二标记的亲试剂可能有更多的发射事件,因此仓的表观强度明显高于第一标记的亲试剂。

[0303] 发明人已经认识到并理解,基于荧光团衰减率和/或荧光团强度区分核苷酸或任

何其他生物或化学样品可以简化光学激发和检测系统。例如,可以用单波长源(例如,产生一个特征波长而不是多个源的源或以多个不同特征波长操作的源)来执行光激发。此外,检测系统中可能不需要波长识别光学器件和滤光器。此外,每个样品孔可以使用单个光检测器来检测来自不同荧光团的发射。短语“特征波长”或“波长”用于指代有限辐射带宽内的中心或主要波长(例如,脉冲光源输出的20nm带宽内的中心或峰值波长)。在一些情况下,“特征波长”或“波长”可用于指代源辐射输出的总带宽内的峰值波长。

[0304] 等同和范围

[0305] 在权利要求书中,除非有相反的指示或从上下文中明显看出,否则诸如“一个/一种(a)”、“一个/一种(an)”和“该/所述(the)”之类的冠词可以表示一个或多个。除非有相反的指示或从上下文中明显看出,否则如果一个、多于一个或所有组成员存在于给定产品或方法中、在给定产品或方法中使用、或以其他方式与给定产品或方法相关,则认为在组的一个或多个成员之间包含“或”的权利要求或描述是令人满意的。本发明包括其中该组的一个成员恰好存在于给定产品或方法中、在给定产品或方法中使用、或以其他方式与给定产品或方法相关的实施方案。本发明包括其中多于一个或所有组成员存在于给定产品或方法中、在给定产品或方法中使用、或以其他方式与给定产品或方法相关的实施方案。

[0306] 此外,本发明涵盖所有变化、组合和排列,其中将来自一个或多个所列权利要求的一个或多个限制、要素、条款和描述性术语引入另一权利要求。例如,可以修改从属于另一权利要求的任何权利要求,以包括在从属于相同基本权利要求的任何其他权利要求中发现的一个或多个限制。在要素作为列表,例如以马库什组格式呈现的情况下,要素的每个子组也被公开,并且可以从所述组中删除任何要素。应当理解,一般而言,在本发明或本发明的方面被称为包括特定要素和/或特征的情况下,本发明的某些实施方案或本发明的方面由或基本上由这样的要素和/或特征组成。为简单起见,这些实施方案并未在本文中具体阐述。

[0307] 如本文在说明书和权利要求书中使用的短语“和/或”应理解为表示如此结合的要素中的“一个或两个”,即,在一些情况下要素结合存在而在其他情况下要素分离存在。用“和/或”列出的多个要素应该以相同的方式解释,即“一个或多个”这样结合的元素。除了由“和/或”子句具体标识的要素之外,可以任选地存在其他要素,无论与那些具体标识的要素相关或不相关。因此,作为非限制性实例,当与诸如“包含”之类的开放式语言结合使用时,对“A和/或B”的引用在一个实施方案中可以仅指A(任选地包括除B之外的要素);在另一个实施方案中,仅指B(任选地包括除A之外的要素);在又一个实施方案中,指A和B两者(任选地包括其他元素);等等。

[0308] 如本文在说明书和权利要求书中使用的,“或”应理解为与如上定义的“和/或”具有相同的含义。例如,当分开列表中的项目时,“或”或“和/或”应被解释为是包含的,即包含至少一个,但也包含要素的数量或列表以及任选地其他未列出的项目中的多于一个。只有明确指出相反的术语,例如“仅一个”或“恰好一个”,或在权利要求中使用“由……组成”将指包含要素的数量或列表中的恰好一个要素。一般而言,当前面带有排他性术语,例如“任一”、“其中一个”、“只有一个”或“恰好一个”时,本文使用的术语“或”仅应解释为表示排他性的替代方案(即“一个或另一个但不是两者”)。当在权利要求中使用“基本上由……组成”应具有专利法领域所使用的一般含义。

[0309] 如本文在说明书和权利要求书中使用的,短语“至少一个”在提及一个或多个要素的列表时,应理解为表示选自要素列表中的任何一个或多个要素的至少一个元素,但不一定包括要素列表中具体列出的每个要素中的至少一个,并且不排除要素列表中的任何要素组合。该定义还允许除了在短语“至少一个”所指的要素列表中具体标识的要素之外的元素可以任选地存在,无论是否与那些具体标识的要素相关或不相关。因此,作为非限制性实例,“A和B中的至少一个”(或等效地,“A或B中的至少一个”,或等效地“A和/或B中的至少一个”)可以,在一个实施方案中,指至少一个,任选地包括多于一个的A,不存在B(并且任选地包括除B之外的要素);在另一个实施方案中,指至少一个,任选地包括多于一个的B,不存在A(并且任选地包括除A之外的要素);在又一个实施方案中,指至少一个,任选地包括多于一个的A,和至少一个,任选地包括多于一个的B(并且任选地包括其它要素);等等。

[0310] 还应该理解的是,除非有明确的相反指示,否则在本文要求保护的任何包括一个以上步骤或动作的方法中,该方法的步骤或动作的顺序不一定限于所记载的方法的步骤或动作的顺序。

[0311] 在权利要求以及上述说明书中,所有过渡短语,例如“包含”、“包括”、“携带”、“具有”、“含有”、“涉及”、“持有”、“由……组成”,等应被理解为是开放式的,即意味着包括但不限于。如美国专利局专利审查程序手册第2111.03节所述,只有过渡短语“由……组成”和“基本上由……组成”应分别为封闭或半封闭式过渡短语。应当理解,在本文档中使用开放式过渡短语(例如,“包含”)描述的实施方案在替代实施方案中还被考虑为“由”开放式过渡短语描述的特征“组成”和“基本上由”开放式过渡短语描述的特征“组成”。例如,如果本申请描述了“包含A和B的组合物”,则该申请还考虑了替代实施方案“由A和B组成的组合物”和“基本上由A和B组成的组合物”。

[0312] 在给出范围的地方,端点包括在内。此外,除非另有说明或从上下文和本领域普通技术人员的理解中以其他方式显而易见,则表示为范围的值可以假定在本发明的不同实施方案中的所述范围内的任何特定值或子范围,以范围下限单位的十分之一,除非上下文另有明确规定。

[0313] 本申请涉及各种已发布的专利、公开的专利申请、期刊文章和其他出版物,所有这些都通过引用的方式并入本文。如果任何并入的参考文献与本说明书之间存在冲突,则以本说明书为准。此外,落入现有技术的本发明的任何特定实施方案可以明确地排除在任何一项或多项权利要求之外。因为这样的实施方案被认为是本领域普通技术人员已知的,所以即使本文没有明确阐述排除,它们也可以被排除。出于任何原因,无论是否与现有技术存在相关,本发明的任何特定实施方案都可以从任何权利要求中排除。

[0314] 本领域技术人员将认识到或能够仅使用常规实验来确定本文描述的特定实施方案的许多等同物。本文描述的本实施方案的范围不旨在限于以上描述,而是如所附权利要求书中所述。本领域的普通技术人员将理解,在不背离本发明的精神或范围的情况下,可以对本说明书进行各种改变和修改,如所附权利要求书所定义的。

[0315] 本文对变量的任何定义中记载的化学基团列表包括将该变量定义为任何单个基团或所列基团的组合。本文对变量的实施方案的记载包括作为任何单个实施方案或与其他实施方案或其部分组合的实施方案。本文对实施方案的记载包括作为任何单个实施方案或与其他实施方案或其部分组合的实施方案。

序列表

- <110> 宽腾矽公司
 <120> 制备用于多肽测序的富集样品的方法
 <130> R0708.70076W000
 <140> 尚未指定
 <141> 与此同时
 <150> US 62/926,897
 <151> 2019-10-28
 <160> 25
 <170> PatentIn version 3.5
 <210> 1
 <211> 921
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成的
 <400> 1

```

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
1          5          10          15
Arg Gly Ser His Met Met Val Lys Gln Gly Val Phe Met Lys Thr Asp
          20          25          30
Gln Ser Lys Val Lys Lys Leu Ser Asp Tyr Lys Ser Leu Asp Tyr Phe
          35          40          45
Val Ile His Val Asp Leu Gln Ile Asp Leu Ser Lys Lys Pro Val Glu
          50          55          60
Ser Lys Ala Arg Leu Thr Val Val Pro Asn Leu Asn Val Asp Ser His
65          70          75          80
Ser Asn Asp Leu Val Leu Asp Gly Glu Asn Met Thr Leu Val Ser Leu
          85          90          95
Gln Met Asn Asp Asn Leu Leu Lys Glu Asn Glu Tyr Glu Leu Thr Lys
          100         105         110
Asp Ser Leu Ile Ile Lys Asn Ile Pro Gln Asn Thr Pro Phe Thr Ile
          115         120         125
Glu Met Thr Ser Leu Leu Gly Glu Asn Thr Asp Leu Phe Gly Leu Tyr
          130         135         140
Glu Thr Glu Gly Val Ala Leu Val Lys Ala Glu Ser Glu Gly Leu Arg
145         150         155         160
Arg Val Phe Tyr Leu Pro Asp Arg Pro Asp Asn Leu Ala Thr Tyr Lys

```

				165					170					175			
Thr	Thr	Ile	Ile	Ala	Asn	Gln	Glu	Asp	Tyr	Pro	Val	Leu	Leu	Ser	Asn		
				180					185					190			
Gly	Val	Leu	Ile	Glu	Lys	Lys	Glu	Leu	Pro	Leu	Gly	Leu	His	Ser	Val		
				195					200					205			
Thr	Trp	Leu	Asp	Asp	Val	Pro	Lys	Pro	Ser	Tyr	Leu	Phe	Ala	Leu	Val		
				210					215					220			
Ala	Gly	Asn	Leu	Gln	Arg	Ser	Val	Thr	Tyr	Tyr	Gln	Thr	Lys	Ser	Gly		
225							230				235				240		
Arg	Glu	Leu	Pro	Ile	Glu	Phe	Tyr	Val	Pro	Pro	Ser	Ala	Thr	Ser	Lys		
				245							250				255		
Cys	Asp	Phe	Ala	Lys	Glu	Val	Leu	Lys	Glu	Ala	Met	Ala	Trp	Asp	Glu		
				260							265				270		
Arg	Thr	Phe	Asn	Leu	Glu	Cys	Ala	Leu	Arg	Gln	His	Met	Val	Ala	Gly		
				275							280				285		
Val	Asp	Lys	Tyr	Ala	Ser	Gly	Ala	Ser	Glu	Pro	Thr	Gly	Leu	Asn	Leu		
				290							295				300		
Phe	Asn	Thr	Glu	Asn	Leu	Phe	Ala	Ser	Pro	Glu	Thr	Lys	Thr	Asp	Leu		
305							310				315				320		
Gly	Ile	Leu	Arg	Val	Leu	Glu	Val	Val	Ala	His	Glu	Phe	Phe	His	Tyr		
				325							330				335		
Trp	Ser	Gly	Asp	Arg	Val	Thr	Ile	Arg	Asp	Trp	Phe	Asn	Leu	Pro	Leu		
				340							345				350		
Lys	Glu	Gly	Leu	Thr	Thr	Phe	Arg	Ala	Ala	Met	Phe	Arg	Glu	Glu	Leu		
				355							360				365		
Phe	Gly	Thr	Asp	Leu	Ile	Arg	Leu	Leu	Asp	Gly	Lys	Asn	Leu	Asp	Glu		
				370							375				380		
Arg	Ala	Pro	Arg	Gln	Ser	Ala	Tyr	Thr	Ala	Val	Arg	Ser	Leu	Tyr	Thr		
385							390				395				400		
Ala	Ala	Ala	Tyr	Glu	Lys	Ser	Ala	Asp	Ile	Phe	Arg	Met	Met	Met	Leu		
				405							410				415		
Phe	Ile	Gly	Lys	Glu	Pro	Phe	Ile	Glu	Ala	Val	Ala	Lys	Phe	Phe	Lys		
				420							425				430		
Asp	Asn	Asp	Gly	Gly	Ala	Val	Thr	Leu	Glu	Asp	Phe	Ile	Glu	Ser	Ile		
				435							440				445		
Ser	Asn	Ser	Ser	Gly	Lys	Asp	Leu	Arg	Ser	Phe	Leu	Ser	Trp	Phe	Thr		
				450							455				460		
Glu	Ser	Gly	Ile	Pro	Glu	Leu	Ile	Val	Thr	Asp	Glu	Leu	Asn	Pro	Asp		
465							470				475				480		

Thr	Lys	Gln	Tyr	Phe	Leu	Lys	Ile	Lys	Thr	Val	Asn	Gly	Arg	Asn	Arg			
					485					490					495			
Pro	Ile	Pro	Ile	Leu	Met	Gly	Leu	Leu	Asp	Ser	Ser	Gly	Ala	Glu	Ile			
					500					505					510			
Val	Ala	Asp	Lys	Leu	Leu	Ile	Val	Asp	Gln	Glu	Glu	Ile	Glu	Phe	Gln			
					515					520					525			
Phe	Glu	Asn	Ile	Gln	Thr	Arg	Pro	Ile	Pro	Ser	Leu	Leu	Arg	Ser	Phe			
					530					535					540			
Ser	Ala	Pro	Val	His	Met	Lys	Tyr	Glu	Tyr	Ser	Tyr	Gln	Asp	Leu	Leu			
					545					550					555			560
Leu	Leu	Met	Gln	Phe	Asp	Thr	Asn	Leu	Tyr	Asn	Arg	Cys	Glu	Ala	Ala			
					565					570					575			
Lys	Gln	Leu	Ile	Ser	Ala	Leu	Ile	Asn	Asp	Phe	Cys	Ile	Gly	Lys	Lys			
					580					585					590			
Ile	Glu	Leu	Ser	Pro	Gln	Phe	Phe	Ala	Val	Tyr	Lys	Ala	Leu	Leu	Ser			
					595					600					605			
Asp	Asn	Ser	Leu	Asn	Glu	Trp	Met	Leu	Ala	Glu	Leu	Ile	Thr	Leu	Pro			
					610					615					620			
Ser	Leu	Glu	Glu	Leu	Ile	Glu	Asn	Gln	Asp	Lys	Pro	Asp	Phe	Glu	Lys			
					625					630					635			640
Leu	Asn	Glu	Gly	Arg	Gln	Leu	Ile	Gln	Asn	Ala	Leu	Ala	Asn	Glu	Leu			
					645					650					655			
Lys	Thr	Asp	Phe	Tyr	Asn	Leu	Leu	Phe	Arg	Ile	Gln	Ile	Ser	Gly	Asp			
					660					665					670			
Asp	Asp	Lys	Gln	Lys	Leu	Lys	Gly	Phe	Asp	Leu	Lys	Gln	Ala	Gly	Leu			
					675					680					685			
Arg	Arg	Leu	Lys	Ser	Val	Cys	Phe	Ser	Tyr	Leu	Leu	Asn	Val	Asp	Phe			
					690					695					700			
Glu	Lys	Thr	Lys	Glu	Lys	Leu	Ile	Leu	Gln	Phe	Glu	Asp	Ala	Leu	Gly			
					705					710					715			720
Lys	Asn	Met	Thr	Glu	Thr	Ala	Leu	Ala	Leu	Ser	Met	Leu	Cys	Glu	Ile			
					725					730					735			
Asn	Cys	Glu	Glu	Ala	Asp	Val	Ala	Leu	Glu	Asp	Tyr	Tyr	His	Tyr	Trp			
					740					745					750			
Lys	Asn	Asp	Pro	Gly	Ala	Val	Asn	Asn	Trp	Phe	Ser	Ile	Gln	Ala	Leu			
					755					760					765			
Ala	His	Ser	Pro	Asp	Val	Ile	Glu	Arg	Val	Lys	Lys	Leu	Met	Arg	His			
					770					775					780			
Gly	Asp	Phe	Asp	Leu	Ser	Asn	Pro	Asn	Lys	Val	Tyr	Ala	Leu	Leu	Gly			

785	790	795	800
Ser Phe Ile Lys Asn Pro Phe Gly Phe His Ser Val Thr Gly Glu Gly			
	805	810	815
Tyr Gln Leu Val Ala Asp Ala Ile Phe Asp Leu Asp Lys Ile Asn Pro			
	820	825	830
Thr Leu Ala Ala Asn Leu Thr Glu Lys Phe Thr Tyr Trp Asp Lys Tyr			
	835	840	845
Asp Val Asn Arg Gln Ala Met Met Ile Ser Thr Leu Lys Ile Ile Tyr			
	850	855	860
Ser Asn Ala Thr Ser Ser Asp Val Arg Thr Met Ala Lys Lys Gly Leu			
865	870	875	880
Asp Lys Val Lys Glu Asp Leu Pro Leu Pro Ile His Leu Thr Phe His			
	885	890	895
Gly Gly Ser Thr Met Gln Asp Arg Thr Ala Gln Leu Ile Ala Asp Gly			
	900	905	910
Asn Lys Glu Asn Ala Tyr Gln Leu His			
	915	920	
<210> 2			
<211> 273			
<212> PRT			
<213> 人工序列			
<220>			
<223> 合成的			
<400> 2			
Met Ala His His His His His His Met Gly Thr Ala Ile Ser Ile Lys			
1	5	10	15
Thr Pro Glu Asp Ile Glu Lys Met Arg Val Ala Gly Arg Leu Ala Ala			
	20	25	30
Glu Val Leu Glu Met Ile Glu Pro Tyr Val Lys Pro Gly Val Ser Thr			
	35	40	45
Gly Glu Leu Asp Arg Ile Cys Asn Asp Tyr Ile Val Asn Glu Gln His			
	50	55	60
Ala Val Ser Ala Cys Leu Gly Tyr His Gly Tyr Pro Lys Ser Val Cys			
65	70	75	80
Ile Ser Ile Asn Glu Val Val Cys His Gly Ile Pro Asp Asp Ala Lys			
	85	90	95
Leu Leu Lys Asp Gly Asp Ile Val Asn Ile Asp Val Thr Val Ile Lys			
	100	105	110
Asp Gly Phe His Gly Asp Thr Ser Lys Met Phe Ile Val Gly Lys Pro			

115	120	125
Thr Ile Met Gly Glu Arg Leu Cys Arg Ile Thr Gln Glu Ser Leu Tyr		
130	135	140
Leu Ala Leu Arg Met Val Lys Pro Gly Ile Asn Leu Arg Glu Ile Gly		
145	150	155
Ala Ala Ile Gln Lys Phe Val Glu Ala Glu Gly Phe Ser Val Val Arg		
165	170	175
Glu Tyr Cys Gly His Gly Ile Gly Arg Gly Phe His Glu Glu Pro Gln		
180	185	190
Val Leu His Tyr Asp Ser Arg Glu Thr Asn Val Val Leu Lys Pro Gly		
195	200	205
Met Thr Phe Thr Ile Glu Pro Met Val Asn Ala Gly Lys Lys Glu Ile		
210	215	220
Arg Thr Met Lys Asp Gly Trp Thr Val Lys Thr Lys Asp Arg Ser Leu		
225	230	235
Ser Ala Gln Tyr Glu His Thr Ile Val Val Thr Asp Asn Gly Cys Glu		
245	250	255
Ile Leu Thr Leu Arg Lys Asp Asp Thr Ile Pro Ala Ile Ile Ser His		
260	265	270

Asp

<210> 3

<211> 330

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 3

Met Ala His His His His His Met Gly Thr Leu Glu Ala Asn Thr		
1	5	10
Asn Gly Pro Gly Ser Met Leu Ser Arg Met Pro Val Ser Ser Arg Thr		
20	25	30
Val Pro Phe Gly Asp His Glu Thr Trp Val Gln Val Thr Thr Pro Glu		
35	40	45
Asn Ala Gln Pro His Ala Leu Pro Leu Ile Val Leu His Gly Gly Pro		
50	55	60
Gly Met Ala His Asn Tyr Val Ala Asn Ile Ala Ala Leu Ala Asp Glu		
65	70	75
Thr Gly Arg Thr Val Ile His Tyr Asp Gln Val Gly Cys Gly Asn Ser		
85	90	95

Thr His Leu Pro Asp Ala Pro Ala Asp Phe Trp Thr Pro Gln Leu Phe		
100	105	110
Val Asp Glu Phe His Ala Val Cys Thr Ala Leu Gly Ile Glu Arg Tyr		
115	120	125
His Val Leu Gly Gln Ser Trp Gly Gly Met Leu Gly Ala Glu Ile Ala		
130	135	140
Val Arg Gln Pro Ser Gly Leu Val Ser Leu Ala Ile Cys Asn Ser Pro		
145	150	155
Ala Ser Met Arg Leu Trp Ser Glu Ala Ala Gly Asp Leu Arg Ala Gln		
165	170	175
Leu Pro Ala Glu Thr Arg Ala Ala Leu Asp Arg His Glu Ala Ala Gly		
180	185	190
Thr Ile Thr His Pro Asp Tyr Leu Gln Ala Ala Ala Glu Phe Tyr Arg		
195	200	205
Arg His Val Cys Arg Val Val Pro Thr Pro Gln Asp Phe Ala Asp Ser		
210	215	220
Val Ala Gln Met Glu Ala Glu Pro Thr Val Tyr His Thr Met Asn Gly		
225	230	235
Pro Asn Glu Phe His Val Val Gly Thr Leu Gly Asp Trp Ser Val Ile		
245	250	255
Asp Arg Leu Pro Asp Val Thr Ala Pro Val Leu Val Ile Ala Gly Glu		
260	265	270
His Asp Glu Ala Thr Pro Lys Thr Trp Gln Pro Phe Val Asp His Ile		
275	280	285
Pro Asp Val Arg Ser His Val Phe Pro Gly Thr Ser His Cys Thr His		
290	295	300
Leu Glu Lys Pro Glu Glu Phe Arg Ala Val Val Ala Gln Phe Leu His		
305	310	315
Gln His Asp Leu Ala Ala Asp Ala Arg Val		
325	330	
<210>	4	
<211>	452	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	4	
Met Thr Gln Gln Glu Tyr Gln Asn Arg Arg Gln Ala Leu Leu Ala Lys		
1	5	10
		15

Met	Ala	Pro	Gly	Ser	Ala	Ala	Ile	Ile	Phe	Ala	Ala	Pro	Glu	Ala	Thr
20				25				30							
Arg	Ser	Ala	Asp	Ser	Glu	Tyr	Pro	Tyr	Arg	Gln	Asn	Ser	Asp	Phe	Ser
35				40				45							
Tyr	Leu	Thr	Gly	Phe	Asn	Glu	Pro	Glu	Ala	Val	Leu	Ile	Leu	Val	Lys
50				55				60							
Ser	Asp	Glu	Thr	His	Asn	His	Ser	Val	Leu	Phe	Asn	Arg	Ile	Arg	Asp
65				70				75				80			
Leu	Thr	Ala	Glu	Ile	Trp	Phe	Gly	Arg	Arg	Leu	Gly	Gln	Glu	Ala	Ala
85				90				95							
Pro	Thr	Lys	Leu	Ala	Val	Asp	Arg	Ala	Leu	Pro	Phe	Asp	Glu	Ile	Asn
100				105				110							
Glu	Gln	Leu	Tyr	Leu	Leu	Leu	Asn	Arg	Leu	Asp	Val	Ile	Tyr	His	Ala
115				120				125							
Gln	Gly	Gln	Tyr	Ala	Tyr	Ala	Asp	Asn	Ile	Val	Phe	Ala	Ala	Leu	Glu
130				135				140							
Lys	Leu	Arg	His	Gly	Phe	Arg	Lys	Asn	Leu	Arg	Ala	Pro	Ala	Thr	Leu
145				150				155				160			
Thr	Asp	Trp	Arg	Pro	Trp	Leu	His	Glu	Met	Arg	Leu	Phe	Lys	Ser	Ala
165				170				175							
Glu	Glu	Ile	Ala	Val	Leu	Arg	Arg	Ala	Gly	Glu	Ile	Ser	Ala	Leu	Ala
180				185				190							
His	Thr	Arg	Ala	Met	Glu	Lys	Cys	Arg	Pro	Gly	Met	Phe	Glu	Tyr	Gln
195				200				205							
Leu	Glu	Gly	Glu	Ile	Leu	His	Glu	Phe	Thr	Arg	His	Gly	Ala	Arg	Tyr
210				215				220							
Pro	Ala	Tyr	Asn	Thr	Ile	Val	Gly	Gly	Gly	Glu	Asn	Gly	Cys	Ile	Leu
225				230				235				240			
His	Tyr	Thr	Glu	Asn	Glu	Cys	Glu	Leu	Arg	Asp	Gly	Asp	Leu	Val	Leu
245				250				255							
Ile	Asp	Ala	Gly	Cys	Glu	Tyr	Arg	Gly	Tyr	Ala	Gly	Asp	Ile	Thr	Arg
260				265				270							
Thr	Phe	Pro	Val	Asn	Gly	Lys	Phe	Thr	Pro	Ala	Gln	Arg	Ala	Val	Tyr
275				280				285							
Asp	Ile	Val	Leu	Ala	Ala	Ile	Asn	Lys	Ser	Leu	Thr	Leu	Phe	Arg	Pro
290				295				300							
Gly	Thr	Ser	Ile	Arg	Glu	Val	Thr	Glu	Glu	Val	Val	Arg	Ile	Met	Val
305				310				315				320			
Val	Gly	Leu	Val	Glu	Leu	Gly	Ile	Leu	Lys	Gly	Asp	Ile	Glu	Gln	Leu

				325					330					335		
Ile	Ala	Glu	Gln	Ala	His	Arg	Pro	Phe	Phe	Met	His	Gly	Leu	Ser	His	
			340					345					350			
Trp	Leu	Gly	Met	Asp	Val	His	Asp	Val	Gly	Asp	Tyr	Gly	Ser	Ser	Asp	
		355					360					365				
Arg	Gly	Arg	Ile	Leu	Glu	Pro	Gly	Met	Val	Leu	Thr	Val	Glu	Pro	Gly	
	370					375					380					
Leu	Tyr	Ile	Ala	Pro	Asp	Ala	Asp	Val	Pro	Pro	Gln	Tyr	Arg	Gly	Ile	
385				390					395						400	
Gly	Ile	Arg	Ile	Glu	Asp	Asp	Ile	Val	Ile	Thr	Ala	Thr	Gly	Asn	Glu	
			405					410					415			
Asn	Leu	Thr	Ala	Ser	Val	Val	Lys	Asp	Pro	Asp	Asp	Ile	Glu	Ala	Leu	
	420					425					430					
Met	Ala	Leu	Asn	His	Ala	Gly	Glu	Asn	Leu	Tyr	Phe	Gln	Glu	His	His	
	435					440					445					
His	His	His	His													
	450															
<210>	5															
<211>	303															
<212>	PRT															
<213>	人工序列															
<220>																
<223>	合成的															
<400>	5															
Met	Asp	Thr	Glu	Lys	Leu	Met	Lys	Ala	Gly	Glu	Ile	Ala	Lys	Lys	Val	
1				5					10				15			
Arg	Glu	Lys	Ala	Ile	Lys	Leu	Ala	Arg	Pro	Gly	Met	Leu	Leu	Leu	Glu	
			20						25				30			
Leu	Ala	Glu	Ser	Ile	Glu	Lys	Met	Ile	Met	Glu	Leu	Gly	Gly	Lys	Pro	
		35					40					45				
Ala	Phe	Pro	Val	Asn	Leu	Ser	Ile	Asn	Glu	Ile	Ala	Ala	His	Tyr	Thr	
	50					55					60					
Pro	Tyr	Lys	Gly	Asp	Thr	Thr	Val	Leu	Lys	Glu	Gly	Asp	Tyr	Leu	Lys	
65				70					75					80		
Ile	Asp	Val	Gly	Val	His	Ile	Asp	Gly	Phe	Ile	Ala	Asp	Thr	Ala	Val	
			85					90					95			
Thr	Val	Arg	Val	Gly	Met	Glu	Glu	Asp	Glu	Leu	Met	Glu	Ala	Ala	Lys	
	100					105					110					
Glu	Ala	Leu	Asn	Ala	Ala	Ile	Ser	Val	Ala	Arg	Ala	Gly	Val	Glu	Ile	

115	120	125
Lys Glu Leu Gly Lys Ala Ile	Glu Asn Glu Ile Arg Lys Arg Gly Phe	
130	135	140
Lys Pro Ile Val Asn Leu Ser Gly His Lys Ile Glu Arg Tyr Lys Leu		
145	150	155
His Ala Gly Ile Ser Ile Pro Asn Ile Tyr Arg Pro His Asp Asn Tyr		
165	170	175
Val Leu Lys Glu Gly Asp Val Phe Ala Ile Glu Pro Phe Ala Thr Ile		
180	185	190
Gly Ala Gly Gln Val Ile Glu Val Pro Pro Thr Leu Ile Tyr Met Tyr		
195	200	205
Val Arg Asp Val Pro Val Arg Val Ala Gln Ala Arg Phe Leu Leu Ala		
210	215	220
Lys Ile Lys Arg Glu Tyr Gly Thr Leu Pro Phe Ala Tyr Arg Trp Leu		
225	230	235
Gln Asn Asp Met Pro Glu Gly Gln Leu Lys Leu Ala Leu Lys Thr Leu		
245	250	255
Glu Lys Ala Gly Ala Ile Tyr Gly Tyr Pro Val Leu Lys Glu Ile Arg		
260	265	270
Asn Gly Ile Val Ala Gln Phe Glu His Thr Ile Ile Val Glu Lys Asp		
275	280	285
Ser Val Ile Val Thr Gln Asp Met Ile Asn Lys Ser Thr Leu Glu		
290	295	300
<210> 6		
<211> 428		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成的		
<400> 6		
His Met Ser Ser Pro Leu His Tyr Val Leu Asp Gly Ile His Cys Glu		
1	5	10
Pro His Phe Phe Thr Val Pro Leu Asp His Gln Gln Pro Asp Asp Glu		
20	25	30
Glu Thr Ile Thr Leu Phe Gly Arg Thr Leu Cys Arg Lys Asp Arg Leu		
35	40	45
Asp Asp Glu Leu Pro Trp Leu Leu Tyr Leu Gln Gly Gly Pro Gly Phe		
50	55	60
Gly Ala Pro Arg Pro Ser Ala Asn Gly Gly Trp Ile Lys Arg Ala Leu		

65	70	75	80
Gln Glu Phe Arg Val Leu Leu Leu Asp Gln Arg Gly Thr Gly His Ser			
	85	90	95
Thr Pro Ile His Ala Glu Leu Leu Ala His Leu Asn Pro Arg Gln Gln			
	100	105	110
Ala Asp Tyr Leu Ser His Phe Arg Ala Asp Ser Ile Val Arg Asp Ala			
	115	120	125
Glu Leu Ile Arg Glu Gln Leu Ser Pro Asp His Pro Trp Ser Leu Leu			
	130	135	140
Gly Gln Ser Phe Gly Gly Phe Cys Ser Leu Thr Tyr Leu Ser Leu Phe			
145	150	155	160
Pro Asp Ser Leu His Glu Val Tyr Leu Thr Gly Gly Val Ala Pro Ile			
	165	170	175
Gly Arg Ser Ala Asp Glu Val Tyr Arg Ala Thr Tyr Gln Arg Val Ala			
	180	185	190
Asp Lys Asn Arg Ala Phe Phe Ala Arg Phe Pro His Ala Gln Ala Ile			
	195	200	205
Ala Asn Arg Leu Ala Thr His Leu Gln Arg His Asp Val Arg Leu Pro			
	210	215	220
Asn Gly Gln Arg Leu Thr Val Glu Gln Leu Gln Gln Gln Gly Leu Asp			
225	230	235	240
Leu Gly Ala Ser Gly Ala Phe Glu Glu Leu Tyr Tyr Leu Leu Glu Asp			
	245	250	255
Ala Phe Ile Gly Glu Lys Leu Asn Pro Ala Phe Leu Tyr Gln Val Gln			
	260	265	270
Ala Met Gln Pro Phe Asn Thr Asn Pro Val Phe Ala Ile Leu His Glu			
	275	280	285
Leu Ile Tyr Cys Glu Gly Ala Ala Ser His Trp Ala Ala Glu Arg Val			
	290	295	300
Arg Gly Glu Phe Pro Ala Leu Ala Trp Ala Gln Gly Lys Asp Phe Ala			
305	310	315	320
Phe Thr Gly Glu Met Ile Phe Pro Trp Met Phe Glu Gln Phe Arg Glu			
	325	330	335
Leu Ile Pro Leu Lys Glu Ala Ala His Leu Leu Ala Glu Lys Ala Asp			
	340	345	350
Trp Gly Pro Leu Tyr Asp Pro Val Gln Leu Ala Arg Asn Lys Val Pro			
	355	360	365
Val Ala Cys Ala Val Tyr Ala Glu Asp Met Tyr Val Glu Phe Asp Tyr			
	370	375	380

Ser Arg Glu Thr Leu Lys Gly Leu Ser Asn Ser Arg Ala Trp Ile Thr
 385 390 395 400
 Asn Glu Tyr Glu His Asn Gly Leu Arg Val Asp Gly Glu Gln Ile Leu
 405 410 415
 Asp Arg Leu Ile Arg Leu Asn Arg Asp Cys Leu Glu
 420 425
 <210> 7
 <211> 348
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成的
 <400> 7
 Met Lys Glu Arg Leu Glu Lys Leu Val Lys Phe Met Asp Glu Asn Ser
 1 5 10 15
 Ile Asp Arg Val Phe Ile Ala Lys Pro Val Asn Val Tyr Tyr Phe Ser
 20 25 30
 Gly Thr Ser Pro Leu Gly Gly Gly Tyr Ile Ile Val Asp Gly Asp Glu
 35 40 45
 Ala Thr Leu Tyr Val Pro Glu Leu Glu Tyr Glu Met Ala Lys Glu Glu
 50 55 60
 Ser Lys Leu Pro Val Val Lys Phe Lys Lys Phe Asp Glu Ile Tyr Glu
 65 70 75 80
 Ile Leu Lys Asn Thr Glu Thr Leu Gly Ile Glu Gly Thr Leu Ser Tyr
 85 90 95
 Ser Met Val Glu Asn Phe Lys Glu Lys Ser Asn Val Lys Glu Phe Lys
 100 105 110
 Lys Ile Asp Asp Val Ile Lys Asp Leu Arg Ile Ile Lys Thr Lys Glu
 115 120 125
 Glu Ile Glu Ile Ile Glu Lys Ala Cys Glu Ile Ala Asp Lys Ala Val
 130 135 140
 Met Ala Ala Ile Glu Glu Ile Thr Glu Gly Lys Arg Glu Arg Glu Val
 145 150 155 160
 Ala Ala Lys Val Glu Tyr Leu Met Lys Met Asn Gly Ala Glu Lys Pro
 165 170 175
 Ala Phe Asp Thr Ile Ile Ala Ser Gly His Arg Ser Ala Leu Pro His
 180 185 190
 Gly Val Ala Ser Asp Lys Arg Ile Glu Arg Gly Asp Leu Val Val Ile
 195 200 205

Asp	Leu	Gly	Ala	Leu	Tyr	Asn	His	Tyr	Asn	Ser	Asp	Ile	Thr	Arg	Thr
210						215					220				
Ile	Val	Val	Gly	Ser	Pro	Asn	Glu	Lys	Gln	Arg	Glu	Ile	Tyr	Glu	Ile
225					230					235					240
Val	Leu	Glu	Ala	Gln	Lys	Arg	Ala	Val	Glu	Ala	Ala	Lys	Pro	Gly	Met
				245					250					255	
Thr	Ala	Lys	Glu	Leu	Asp	Ser	Ile	Ala	Arg	Glu	Ile	Ile	Lys	Glu	Tyr
			260					265					270		
Gly	Tyr	Gly	Asp	Tyr	Phe	Ile	His	Ser	Leu	Gly	His	Gly	Val	Gly	Leu
		275					280					285			
Glu	Ile	His	Glu	Trp	Pro	Arg	Ile	Ser	Gln	Tyr	Asp	Glu	Thr	Val	Leu
	290					295					300				
Lys	Glu	Gly	Met	Val	Ile	Thr	Ile	Glu	Pro	Gly	Ile	Tyr	Ile	Pro	Lys
305					310					315					320
Leu	Gly	Gly	Val	Arg	Ile	Glu	Asp	Thr	Val	Leu	Ile	Thr	Glu	Asn	Gly
				325					330					335	
Ala	Lys	Arg	Leu	Thr	Lys	Thr	Glu	Arg	Glu	Leu	Leu				
			340					345							

<210> 8

<211> 298

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 8

Met	Ile	Pro	Ile	Thr	Thr	Pro	Val	Gly	Asn	Phe	Lys	Val	Trp	Thr	Lys
1				5				10					15		
Arg	Phe	Gly	Thr	Asn	Pro	Lys	Ile	Lys	Val	Leu	Leu	Leu	His	Gly	Gly
		20					25					30			
Pro	Ala	Met	Thr	His	Glu	Tyr	Met	Glu	Cys	Phe	Glu	Thr	Phe	Phe	Gln
		35				40					45				
Arg	Glu	Gly	Phe	Glu	Phe	Tyr	Glu	Tyr	Asp	Gln	Leu	Gly	Ser	Tyr	Tyr
	50					55				60					
Ser	Asp	Gln	Pro	Thr	Asp	Glu	Lys	Leu	Trp	Asn	Ile	Asp	Arg	Phe	Val
65				70					75					80	
Asp	Glu	Val	Glu	Gln	Val	Arg	Lys	Ala	Ile	His	Ala	Asp	Lys	Glu	Asn
				85				90						95	
Phe	Tyr	Val	Leu	Gly	Asn	Ser	Trp	Gly	Gly	Ile	Leu	Ala	Met	Glu	Tyr
			100					105					110		

Ala Leu Lys Tyr Gln Gln Asn Leu Lys Gly Leu Ile Val Ala Asn Met		
115	120	125
Met Ala Ser Ala Pro Glu Tyr Val Lys Tyr Ala Glu Val Leu Ser Lys		
130	135	140
Gln Met Lys Pro Glu Val Leu Ala Glu Val Arg Ala Ile Glu Ala Lys		
145	150	155
Lys Asp Tyr Ala Asn Pro Arg Tyr Thr Glu Leu Leu Phe Pro Asn Tyr		
	165	170
Tyr Ala Gln His Ile Cys Arg Leu Lys Glu Trp Pro Asp Ala Leu Asn		
	180	185
Arg Ser Leu Lys His Val Asn Ser Thr Val Tyr Thr Leu Met Gln Gly		
	195	200
Pro Ser Glu Leu Gly Met Ser Ser Asp Ala Arg Leu Ala Lys Trp Asp		
	210	215
Ile Lys Asn Arg Leu His Glu Ile Ala Thr Pro Thr Leu Met Ile Gly		
225	230	235
Ala Arg Tyr Asp Thr Met Asp Pro Lys Ala Met Glu Glu Gln Ser Lys		
	245	250
Leu Val Gln Lys Gly Arg Tyr Leu Tyr Cys Pro Asn Gly Ser His Leu		
	260	265
Ala Met Trp Asp Asp Gln Lys Val Phe Met Asp Gly Val Ile Lys Phe		
	275	280
Ile Lys Asp Val Asp Thr Lys Ser Phe Asn		
290	295	
<210> 9		
<211> 428		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成的		
<400> 9		
His Met Ser Ser Pro Leu His Tyr Val Leu Asp Gly Ile His Cys Glu		
1	5	10
Pro His Phe Phe Thr Val Pro Leu Asp His Gln Gln Pro Asp Asp Glu		
	20	25
Glu Thr Ile Thr Leu Phe Gly Arg Thr Leu Cys Arg Lys Asp Arg Leu		
	35	40
Asp Asp Glu Leu Pro Trp Leu Leu Tyr Leu Gln Gly Gly Pro Gly Phe		
50	55	60

Gly	Ala	Pro	Arg	Pro	Ser	Ala	Asn	Gly	Gly	Trp	Ile	Lys	Arg	Ala	Leu
65					70					75					80
Gln	Glu	Phe	Arg	Val	Leu	Leu	Leu	Asp	Gln	Arg	Gly	Thr	Gly	His	Ser
				85						90					95
Thr	Pro	Ile	His	Ala	Glu	Leu	Leu	Ala	His	Leu	Asn	Pro	Arg	Gln	Gln
				100						105					110
Ala	Asp	Tyr	Leu	Ser	His	Phe	Arg	Ala	Asp	Ser	Ile	Val	Arg	Asp	Ala
				115						120					125
Glu	Leu	Ile	Arg	Glu	Gln	Leu	Ser	Pro	Asp	His	Pro	Trp	Ser	Leu	Leu
				130						135					140
Gly	Gln	Ser	Phe	Gly	Gly	Phe	Cys	Ser	Leu	Thr	Tyr	Leu	Ser	Leu	Phe
145										150					155
Pro	Asp	Ser	Leu	His	Glu	Val	Tyr	Leu	Thr	Gly	Gly	Val	Ala	Pro	Ile
										165					170
Gly	Arg	Ser	Ala	Asp	Glu	Val	Tyr	Arg	Ala	Thr	Tyr	Gln	Arg	Val	Ala
										180					185
Asp	Lys	Asn	Arg	Ala	Phe	Phe	Ala	Arg	Phe	Pro	His	Ala	Gln	Ala	Ile
										195					200
Ala	Asn	Arg	Leu	Ala	Thr	His	Leu	Gln	Arg	His	Asp	Val	Arg	Leu	Pro
										210					215
Asn	Gly	Gln	Arg	Leu	Thr	Val	Glu	Gln	Leu	Gln	Gln	Gln	Gly	Leu	Asp
225										230					235
Leu	Gly	Ala	Ser	Gly	Ala	Phe	Glu	Glu	Leu	Tyr	Tyr	Leu	Leu	Glu	Asp
										245					250
Ala	Phe	Ile	Gly	Glu	Lys	Leu	Asn	Pro	Ala	Phe	Leu	Tyr	Gln	Val	Gln
										260					265
Ala	Met	Gln	Pro	Phe	Asn	Thr	Asn	Pro	Val	Phe	Ala	Ile	Leu	His	Glu
										275					280
Leu	Ile	Tyr	Cys	Glu	Gly	Ala	Ala	Ser	His	Trp	Ala	Ala	Glu	Arg	Val
										285					290
Arg	Gly	Glu	Phe	Pro	Ala	Leu	Ala	Trp	Ala	Gln	Gly	Lys	Asp	Phe	Ala
305										310					315
Phe	Thr	Gly	Glu	Met	Ile	Phe	Pro	Trp	Met	Phe	Glu	Gln	Phe	Arg	Glu
										325					330
Leu	Ile	Pro	Leu	Lys	Glu	Ala	Ala	His	Leu	Leu	Ala	Glu	Lys	Ala	Asp
										340					345
Trp	Gly	Pro	Leu	Tyr	Asp	Pro	Val	Gln	Leu	Ala	Arg	Asn	Lys	Val	Pro
										355					360
Val	Ala	Cys	Ala	Val	Tyr	Ala	Glu	Asp	Met	Tyr	Val	Glu	Phe	Asp	Tyr
										365					

370	375	380
Ser Arg Glu Thr Leu Lys Gly Leu Ser Asn Ser Arg Ala Trp Ile Thr		
385	390	395
Asn Glu Tyr Glu His Asn Gly Leu Arg Val Asp Gly Glu Gln Ile Leu		400
	405	410
Asp Arg Leu Ile Arg Leu Asn Arg Asp Cys Leu Glu		415
420	425	
<210> 10		
<211> 310		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成的		
<400> 10		
Met Tyr Glu Ile Lys Gln Pro Phe His Ser Gly Tyr Leu Gln Val Ser		
1	5	10
Glu Ile His Gln Ile Tyr Trp Glu Glu Ser Gly Asn Pro Asp Gly Val		15
	20	25
Pro Val Ile Phe Leu His Gly Gly Pro Gly Ala Gly Ala Ser Pro Glu		30
	35	40
Cys Arg Gly Phe Phe Asn Pro Asp Val Phe Arg Ile Val Ile Ile Asp		45
50	55	60
Gln Arg Gly Cys Gly Arg Ser His Pro Tyr Ala Cys Ala Glu Asp Asn		
65	70	75
Thr Thr Trp Asp Leu Val Ala Asp Ile Glu Lys Val Arg Glu Met Leu		80
	85	90
Gly Ile Gly Lys Trp Leu Val Phe Gly Gly Ser Trp Gly Ser Thr Leu		95
	100	105
Ser Leu Ala Tyr Ala Gln Thr His Pro Glu Arg Val Lys Gly Leu Val		110
	115	120
Leu Arg Gly Ile Phe Leu Cys Arg Pro Ser Glu Thr Ala Trp Leu Asn		125
	130	135
Glu Ala Gly Gly Val Ser Arg Ile Tyr Pro Glu Gln Trp Gln Lys Phe		140
145	150	155
Val Ala Pro Ile Ala Glu Asn Arg Arg Asn Arg Leu Ile Glu Ala Tyr		160
	165	170
His Gly Leu Leu Phe His Gln Asp Glu Glu Val Cys Leu Ser Ala Ala		175
	180	185
Lys Ala Trp Ala Asp Trp Glu Ser Tyr Leu Ile Arg Phe Glu Pro Glu		190

195	200	205
Gly Val Asp Glu Asp Ala Tyr Ala Ser Leu Ala Ile Ala Arg Leu Glu		
210	215	220
Asn His Tyr Phe Val Asn Gly Gly Trp Leu Gln Gly Asp Lys Ala Ile		
225	230	235
240		
Leu Asn Asn Ile Gly Lys Ile Arg His Ile Pro Thr Val Ile Val Gln		
245	250	255
Gly Arg Tyr Asp Leu Cys Thr Pro Met Gln Ser Ala Trp Glu Leu Ser		
260	265	270
Lys Ala Phe Pro Glu Ala Glu Leu Arg Val Val Gln Ala Gly His Cys		
275	280	285
Ala Phe Asp Pro Pro Leu Ala Asp Ala Leu Val Gln Ala Val Glu Asp		
290	295	300
Ile Leu Pro Arg Leu Leu		
305	310	
<210> 11		
<211> 891		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成的		
<400> 11		
Met Gly Ser Ser His His His His His His Ser Ser Gly Glu Asn Leu		
1	5	10
15		
Tyr Phe Gln Gly His Met Thr Gln Gln Pro Gln Ala Lys Tyr Arg His		
20	25	30
Asp Tyr Arg Ala Pro Asp Tyr Gln Ile Thr Asp Ile Asp Leu Thr Phe		
35	40	45
Asp Leu Asp Ala Gln Lys Thr Val Val Thr Ala Val Ser Gln Ala Val		
50	55	60
Arg His Gly Ala Ser Asp Ala Pro Leu Arg Leu Asn Gly Glu Asp Leu		
65	70	75
80		
Lys Leu Val Ser Val His Ile Asn Asp Glu Pro Trp Thr Ala Trp Lys		
85	90	95
Glu Glu Glu Gly Ala Leu Val Ile Ser Asn Leu Pro Glu Arg Phe Thr		
100	105	110
Leu Lys Ile Ile Asn Glu Ile Ser Pro Ala Ala Asn Thr Ala Leu Glu		
115	120	125
Gly Leu Tyr Gln Ser Gly Asp Ala Leu Cys Thr Gln Cys Glu Ala Glu		

130	135	140
Gly Phe Arg His Ile Thr Tyr Tyr Leu Asp Arg Pro Asp Val Leu Ala		
145	150	155
Arg Phe Thr Thr Lys Ile Ile Ala Asp Lys Ile Lys Tyr Pro Phe Leu		
165	170	175
Leu Ser Asn Gly Asn Arg Val Ala Gln Gly Glu Leu Glu Asn Gly Arg		
180	185	190
His Trp Val Gln Trp Gln Asp Pro Phe Pro Lys Pro Cys Tyr Leu Phe		
195	200	205
Ala Leu Val Ala Gly Asp Phe Asp Val Leu Arg Asp Thr Phe Thr Thr		
210	215	220
Arg Ser Gly Arg Glu Val Ala Leu Glu Leu Tyr Val Asp Arg Gly Asn		
225	230	235
Leu Asp Arg Ala Pro Trp Ala Met Thr Ser Leu Lys Asn Ser Met Lys		
245	250	255
Trp Asp Glu Glu Arg Phe Gly Leu Glu Tyr Asp Leu Asp Ile Tyr Met		
260	265	270
Ile Val Ala Val Asp Phe Phe Asn Met Gly Ala Met Glu Asn Lys Gly		
275	280	285
Leu Asn Ile Phe Asn Ser Lys Tyr Val Leu Ala Arg Thr Asp Thr Ala		
290	295	300
Thr Asp Lys Asp Tyr Leu Asp Ile Glu Arg Val Ile Gly His Glu Tyr		
305	310	315
Phe His Asn Trp Thr Gly Asn Arg Val Thr Cys Arg Asp Trp Phe Gln		
325	330	335
Leu Ser Leu Lys Glu Gly Leu Thr Val Phe Arg Asp Gln Glu Phe Ser		
340	345	350
Ser Asp Leu Gly Ser Arg Ala Val Asn Arg Ile Asn Asn Val Arg Thr		
355	360	365
Met Arg Gly Leu Gln Phe Ala Glu Asp Ala Ser Pro Met Ala His Pro		
370	375	380
Ile Arg Pro Asp Met Val Ile Glu Met Asn Asn Phe Tyr Thr Leu Thr		
385	390	395
Val Tyr Glu Lys Gly Ala Glu Val Ile Arg Met Ile His Thr Leu Leu		
405	410	415
Gly Glu Glu Asn Phe Gln Lys Gly Met Gln Leu Tyr Phe Glu Arg His		
420	425	430
Asp Gly Ser Ala Ala Thr Cys Asp Asp Phe Val Gln Ala Met Glu Asp		
435	440	445

Ala Ser Asn Val Asp Leu Ser His Phe Arg Arg Trp Tyr Ser Gln Ser		
450	455	460
Gly Thr Pro Ile Val Thr Val Lys Asp Asp Tyr Asn Pro Glu Thr Glu		
465	470	475 480
Gln Tyr Thr Leu Thr Ile Ser Gln Arg Thr Pro Ala Thr Pro Asp Gln		
	485	490 495
Ala Glu Lys Gln Pro Leu His Ile Pro Phe Ala Ile Glu Leu Tyr Asp		
	500	505 510
Asn Glu Gly Lys Val Ile Pro Leu Gln Lys Gly Gly His Pro Val Asn		
	515	520 525
Ser Val Leu Asn Val Thr Gln Ala Glu Gln Thr Phe Val Phe Asp Asn		
	530	535 540
Val Tyr Phe Gln Pro Val Pro Ala Leu Leu Cys Glu Phe Ser Ala Pro		
545	550	555 560
Val Lys Leu Glu Tyr Lys Trp Ser Asp Gln Gln Leu Thr Phe Leu Met		
	565	570 575
Arg His Ala Arg Asn Asp Phe Ser Arg Trp Asp Ala Ala Gln Ser Leu		
	580	585 590
Leu Ala Thr Tyr Ile Lys Leu Asn Val Ala Arg His Gln Gln Gly Gln		
	595	600 605
Pro Leu Ser Leu Pro Val His Val Ala Asp Ala Phe Arg Ala Val Leu		
	610	615 620
Leu Asp Glu Lys Ile Asp Pro Ala Leu Ala Ala Glu Ile Leu Thr Leu		
625	630	635 640
Pro Ser Val Asn Glu Met Ala Glu Leu Phe Asp Ile Ile Asp Pro Ile		
	645	650 655
Ala Ile Ala Glu Val Arg Glu Ala Leu Thr Arg Thr Leu Ala Thr Glu		
	660	665 670
Leu Ala Asp Glu Leu Leu Ala Ile Tyr Asn Ala Asn Tyr Gln Ser Glu		
	675	680 685
Tyr Arg Val Glu His Glu Asp Ile Ala Lys Arg Thr Leu Arg Asn Ala		
	690	695 700
Cys Leu Arg Phe Leu Ala Phe Gly Glu Thr His Leu Ala Asp Val Leu		
705	710	715 720
Val Ser Lys Gln Phe His Glu Ala Asn Asn Met Thr Asp Ala Leu Ala		
	725	730 735
Ala Leu Ser Ala Ala Val Ala Ala Gln Leu Pro Cys Arg Asp Ala Leu		
	740	745 750
Met Gln Glu Tyr Asp Asp Lys Trp His Gln Asn Gly Leu Val Met Asp		

755	760	765
Lys Trp Phe Ile Leu Gln Ala Thr Ser Pro Ala Ala Asn Val Leu Glu		
770	775	780
Thr Val Arg Gly Leu Leu Gln His Arg Ser Phe Thr Met Ser Asn Pro		
785	790	795
Asn Arg Ile Arg Ser Leu Ile Gly Ala Phe Ala Gly Ser Asn Pro Ala		
805	810	815
Ala Phe His Ala Glu Asp Gly Ser Gly Tyr Leu Phe Leu Val Glu Met		
820	825	830
Leu Thr Asp Leu Asn Ser Arg Asn Pro Gln Val Ala Ser Arg Leu Ile		
835	840	845
Glu Pro Leu Ile Arg Leu Lys Arg Tyr Asp Ala Lys Arg Gln Glu Lys		
850	855	860
Met Arg Ala Ala Leu Glu Gln Leu Lys Gly Leu Glu Asn Leu Ser Gly		
865	870	875
Asp Leu Tyr Glu Lys Ile Thr Lys Ala Leu Ala		
885	890	

<210> 12

<211> 889

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 12

Pro Lys Ile His Tyr Arg Lys Asp Tyr Lys Pro Ser Gly Phe Ile Ile		
1	5	10
Asn Gln Val Thr Leu Asn Ile Asn Ile His Asp Gln Glu Thr Ile Val		
20	25	30
Arg Ser Val Leu Asp Met Asp Ile Ser Lys His Asn Val Gly Glu Asp		
35	40	45
Leu Val Phe Asp Gly Val Gly Leu Lys Ile Asn Glu Ile Ser Ile Asn		
50	55	60
Asn Lys Lys Leu Val Glu Gly Glu Glu Tyr Thr Tyr Asp Asn Glu Phe		
65	70	75
Leu Thr Ile Phe Ser Lys Phe Val Pro Lys Ser Lys Phe Ala Phe Ser		
85	90	95
Ser Glu Val Ile Ile His Pro Glu Thr Asn Tyr Ala Leu Thr Gly Leu		
100	105	110
Tyr Lys Ser Lys Asn Ile Ile Val Ser Gln Cys Glu Ala Thr Gly Phe		

115	120	125
Arg Arg Ile Thr Phe Phe Ile Asp Arg Pro Asp Met Met Ala Lys Tyr		
130	135	140
Asp Val Thr Val Thr Ala Asp Lys Glu Lys Tyr Pro Val Leu Leu Ser		
145	150	155
Asn Gly Asp Lys Val Asn Glu Phe Glu Ile Pro Gly Gly Arg His Gly		
165	170	175
Ala Arg Phe Asn Asp Pro Pro Leu Lys Pro Cys Tyr Leu Phe Ala Val		
180	185	190
Val Ala Gly Asp Leu Lys His Leu Ser Ala Thr Tyr Ile Thr Lys Tyr		
195	200	205
Thr Lys Lys Lys Val Glu Leu Tyr Val Phe Ser Glu Glu Lys Tyr Val		
210	215	220
Ser Lys Leu Gln Trp Ala Leu Glu Cys Leu Lys Lys Ser Met Ala Phe		
225	230	235
Asp Glu Asp Tyr Phe Gly Leu Glu Tyr Asp Leu Ser Arg Leu Asn Leu		
245	250	255
Val Ala Val Ser Asp Phe Asn Val Gly Ala Met Glu Asn Lys Gly Leu		
260	265	270
Asn Ile Phe Asn Ala Asn Ser Leu Leu Ala Ser Lys Lys Asn Ser Ile		
275	280	285
Asp Phe Ser Tyr Ala Arg Ile Leu Thr Val Val Gly His Glu Tyr Phe		
290	295	300
His Gln Tyr Thr Gly Asn Arg Val Thr Leu Arg Asp Trp Phe Gln Leu		
305	310	315
Thr Leu Lys Glu Gly Leu Thr Val His Arg Glu Asn Leu Phe Ser Glu		
325	330	335
Glu Met Thr Lys Thr Val Thr Thr Arg Leu Ser His Val Asp Leu Leu		
340	345	350
Arg Ser Val Gln Phe Leu Glu Asp Ser Ser Pro Leu Ser His Pro Ile		
355	360	365
Arg Pro Glu Ser Tyr Val Ser Met Glu Asn Phe Tyr Thr Thr Thr Val		
370	375	380
Tyr Asp Lys Gly Ser Glu Val Met Arg Met Tyr Leu Thr Ile Leu Gly		
385	390	395
Glu Glu Tyr Tyr Lys Lys Gly Phe Asp Ile Tyr Ile Lys Lys Asn Asp		
405	410	415
Gly Asn Thr Ala Thr Cys Glu Asp Phe Asn Tyr Ala Met Glu Gln Ala		
420	425	430

Tyr	Lys	Met	Lys	Lys	Ala	Asp	Asn	Ser	Ala	Asn	Leu	Asn	Gln	Tyr	Leu
435				440				445							
Leu	Trp	Phe	Ser	Gln	Ser	Gly	Thr	Pro	His	Val	Ser	Phe	Lys	Tyr	Asn
450				455				460							
Tyr	Asp	Ala	Glu	Lys	Lys	Gln	Tyr	Ser	Ile	His	Val	Asn	Gln	Tyr	Thr
465				470				475				480			
Lys	Pro	Asp	Glu	Asn	Gln	Lys	Glu	Lys	Lys	Pro	Leu	Phe	Ile	Pro	Ile
				485				490				495			
Ser	Val	Gly	Leu	Ile	Asn	Pro	Glu	Asn	Gly	Lys	Glu	Met	Ile	Ser	Gln
500								505				510			
Thr	Thr	Leu	Glu	Leu	Thr	Lys	Glu	Ser	Asp	Thr	Phe	Val	Phe	Asn	Asn
515				520				525							
Ile	Ala	Val	Lys	Pro	Ile	Pro	Ser	Leu	Phe	Arg	Gly	Phe	Ser	Ala	Pro
530				535				540							
Val	Tyr	Ile	Glu	Asp	Gln	Leu	Thr	Asp	Glu	Glu	Arg	Ile	Leu	Leu	Leu
545				550				555				560			
Lys	Tyr	Asp	Ser	Asp	Ala	Phe	Val	Arg	Tyr	Asn	Ser	Cys	Thr	Asn	Ile
				565				570				575			
Tyr	Met	Lys	Gln	Ile	Leu	Met	Asn	Tyr	Asn	Glu	Phe	Leu	Lys	Ala	Lys
580				585				590							
Asn	Glu	Lys	Leu	Glu	Ser	Phe	Gln	Leu	Thr	Pro	Val	Asn	Ala	Gln	Phe
595				600				605							
Ile	Asp	Ala	Ile	Lys	Tyr	Leu	Leu	Glu	Asp	Pro	His	Ala	Asp	Ala	Gly
610				615				620							
Phe	Lys	Ser	Tyr	Ile	Val	Ser	Leu	Pro	Gln	Asp	Arg	Tyr	Ile	Ile	Asn
625				630				635				640			
Phe	Val	Ser	Asn	Leu	Asp	Thr	Asp	Val	Leu	Ala	Asp	Thr	Lys	Glu	Tyr
				645				650				655			
Ile	Tyr	Lys	Gln	Ile	Gly	Asp	Lys	Leu	Asn	Asp	Val	Tyr	Tyr	Lys	Met
660				665				670							
Phe	Lys	Ser	Leu	Glu	Ala	Lys	Ala	Asp	Asp	Leu	Thr	Tyr	Phe	Asn	Asp
675				680				685							
Glu	Ser	His	Val	Asp	Phe	Asp	Gln	Met	Asn	Met	Arg	Thr	Leu	Arg	Asn
690				695				700							
Thr	Leu	Leu	Ser	Leu	Leu	Ser	Lys	Ala	Gln	Tyr	Pro	Asn	Ile	Leu	Asn
705				710				715				720			
Glu	Ile	Ile	Glu	His	Ser	Lys	Ser	Pro	Tyr	Pro	Ser	Asn	Trp	Leu	Thr
				725				730				735			
Ser	Leu	Ser	Val	Ser	Ala	Tyr	Phe	Asp	Lys	Tyr	Phe	Glu	Leu	Tyr	Asp

$\langle 210 \rangle$	13
$\langle 211 \rangle$	932
$\langle 212 \rangle$	PRT
$\langle 213 \rangle$	人工序列
$\langle 220 \rangle$	
$\langle 223 \rangle$	合成的
$\langle 400 \rangle$	13

97

	100		105		110
Ile Asp Ile	Ile Thr Ala Ser Tyr	Ala Pro Glu Gly	Asp Glu Glu	Ile	
	115		120		125
His Ala Thr	Gly Phe Asn Tyr	Gln Asn Glu Asp	Glu Lys Val	Thr Leu	
	130		135		140
Ser Phe Pro	Ser Thr Leu Gln	Thr Gly Thr	Gly Thr Leu	Lys Ile Asp	
145		150		155	160
Phe Val Gly	Glu Leu Asn Asp	Lys Met Lys	Gly Phe Tyr	Arg Ser Lys	
	165		170		175
Tyr Thr Thr	Pro Ser Gly Glu	Val Arg Tyr	Ala Ala Val	Thr Gln Phe	
	180		185		190
Glu Ala Thr	Asp Ala Arg Arg	Ala Phe Pro	Cys Trp Asp	Glu Pro Ala	
	195		200		205
Ile Lys Ala	Thr Phe Asp Ile	Ser Leu Val	Val Pro Lys	Asp Arg Val	
	210		215		220
Ala Leu Ser	Asn Met Asn Val	Ile Asp Arg	Lys Pro Tyr	Pro Asp Asp	
225		230		235	240
Glu Asn Leu	Val Glu Val Lys	Phe Ala Arg	Thr Pro Val	Met Ser Thr	
	245		250		255
Tyr Leu Val	Ala Phe Val Val	Gly Glu Tyr	Asp Phe Val	Glu Thr Arg	
	260		265		270
Ser Lys Asp	Gly Val Cys Val	Arg Val Tyr	Thr Pro Val	Gly Lys Ala	
	275		280		285
Glu Gln Gly	Lys Phe Ala Leu	Glu Val Ala	Ala Lys Thr	Leu Pro Phe	
	290		295		300
Tyr Lys Asp	Tyr Phe Asn Val	Pro Tyr Pro	Leu Pro Lys	Ile Asp Leu	
305		310		315	320
Ile Ala Ile	Ala Asp Phe Ala	Ala Gly Ala	Met Glu Asn	Trp Gly Leu	
	325		330		335
Val Thr Tyr	Arg Glu Thr Ala	Leu Leu Ile	Asp Pro Lys	Asn Ser Cys	
	340		345		350
Ser Ser Ser	Arg Gln Trp Val	Ala Leu Val	Val Gly His	Glu Leu Ala	
	355		360		365
His Gln Trp	Phe Gly Asn Leu	Val Thr Met	Glu Trp Trp	Thr His Leu	
	370		375		380
Trp Leu Asn	Glu Gly Phe Ala	Ser Trp Ile	Glu Tyr Leu	Cys Val Asp	
385		390		395	400
His Cys Phe	Pro Glu Tyr Asp	Ile Trp Thr	Gln Phe Val	Ser Ala Asp	
	405		410		415

Tyr Thr Arg Ala Gln Glu Leu Asp Ala Leu Asp Asn Ser His Pro Ile		
420	425	430
Glu Val Ser Val Gly His Pro Ser Glu Val Asp Glu Ile Phe Asp Ala		
435	440	445
Ile Ser Tyr Ser Lys Gly Ala Ser Val Ile Arg Met Leu His Asp Tyr		
450	455	460
Ile Gly Asp Lys Asp Phe Lys Lys Gly Met Asn Met Tyr Leu Thr Lys		
465	470	475
Phe Gln Gln Lys Asn Ala Ala Thr Glu Asp Leu Trp Glu Ser Leu Glu		
485	490	495
Asn Ala Ser Gly Lys Pro Ile Ala Ala Val Met Asn Thr Trp Thr Lys		
500	505	510
Gln Met Gly Phe Pro Leu Ile Tyr Val Glu Ala Glu Gln Val Glu Asp		
515	520	525
Asp Arg Leu Leu Arg Leu Ser Gln Lys Lys Phe Cys Ala Gly Gly Ser		
530	535	540
Tyr Val Gly Glu Asp Cys Pro Gln Trp Met Val Pro Ile Thr Ile Ser		
545	550	555
Thr Ser Glu Asp Pro Asn Gln Ala Lys Leu Lys Ile Leu Met Asp Lys		
565	570	575
Pro Glu Met Asn Val Val Leu Lys Asn Val Lys Pro Asp Gln Trp Val		
580	585	590
Lys Leu Asn Leu Gly Thr Val Gly Phe Tyr Arg Thr Gln Tyr Ser Ser		
595	600	605
Ala Met Leu Glu Ser Leu Leu Pro Gly Ile Arg Asp Leu Ser Leu Pro		
610	615	620
Pro Val Asp Arg Leu Gly Leu Gln Asn Asp Leu Phe Ser Leu Ala Arg		
625	630	635
Ala Gly Ile Ile Ser Thr Val Glu Val Leu Lys Val Met Glu Ala Phe		
645	650	655
Val Asn Glu Pro Asn Tyr Thr Val Trp Ser Asp Leu Ser Cys Asn Leu		
660	665	670
Gly Ile Leu Ser Thr Leu Leu Ser His Thr Asp Phe Tyr Glu Glu Ile		
675	680	685
Gln Glu Phe Val Lys Asp Val Phe Ser Pro Ile Gly Glu Arg Leu Gly		
690	695	700
Trp Asp Pro Lys Pro Gly Glu Gly His Leu Asp Ala Leu Leu Arg Gly		
705	710	715
Leu Val Leu Gly Lys Leu Gly Lys Ala Gly His Lys Ala Thr Leu Glu		

	725		730		735
Glu Ala Arg Arg Arg Phe Lys Asp His Val Glu Gly Lys Gln Ile Leu					
	740		745		750
Ser Ala Asp Leu Arg Ser Pro Val Tyr Leu Thr Val Leu Lys His Gly					
	755		760		765
Asp Gly Thr Thr Leu Asp Ile Met Leu Lys Leu His Lys Gln Ala Asp					
	770		775		780
Met Gln Glu Glu Lys Asn Arg Ile Glu Arg Val Leu Gly Ala Thr Leu					
785		790		795	800
Leu Pro Asp Leu Ile Gln Lys Val Leu Thr Phe Ala Leu Ser Glu Glu					
	805		810		815
Val Arg Pro Gln Asp Thr Val Ser Val Ile Gly Gly Val Ala Gly Gly					
	820		825		830
Ser Lys His Gly Arg Lys Ala Ala Trp Lys Phe Ile Lys Asp Asn Trp					
	835		840		845
Glu Glu Leu Tyr Asn Arg Tyr Gln Gly Gly Phe Leu Ile Ser Arg Leu					
	850		855		860
Ile Lys Leu Ser Val Glu Gly Phe Ala Val Asp Lys Met Ala Gly Glu					
865		870		875	880
Val Lys Ala Phe Phe Glu Ser His Pro Ala Pro Ser Ala Glu Arg Thr					
	885		890		895
Ile Gln Gln Cys Cys Glu Asn Ile Leu Leu Asn Ala Ala Trp Leu Lys					
	900		905		910
Arg Asp Ala Glu Ser Ile His Gln Tyr Leu Leu Gln Arg Lys Ala Ser					
	915		920		925
Pro Pro Thr Val					
	930				
<210>	14				
<211>	932				
<212>	PRT				
<213>	人工序列				
<220>					
<223>	合成的				
<400>	14				
Met Gly Ser Ser His His His His His Ser Ser Gly Met Trp Leu					
1		5		10	15
Ala Ala Ala Ala Pro Ser Leu Ala Arg Arg Leu Leu Phe Leu Gly Pro					
	20		25		30
Pro Pro Pro Pro Leu Leu Leu Leu Val Phe Ser Arg Ser Ser Arg Arg					

35	40	45
Arg Leu His Ser Leu Gly Leu Ala Ala Met Pro Glu Lys Arg Pro Phe		
50	55	60
Glu Arg Leu Pro Ala Asp Val Ser Pro Ile Asn Tyr Ser Leu Cys Leu		
65	70	75
Lys Pro Asp Leu Leu Asp Phe Thr Phe Glu Gly Lys Leu Glu Ala Ala		
85	90	95
Ala Gln Val Arg Gln Ala Thr Asn Gln Ile Val Met Asn Cys Ala Asp		
100	105	110
Ile Asp Ile Ile Thr Ala Ser Tyr Ala Pro Glu Gly Asp Glu Glu Ile		
115	120	125
His Ala Thr Gly Phe Asn Tyr Gln Asn Glu Asp Glu Lys Val Thr Leu		
130	135	140
Ser Phe Pro Ser Thr Leu Gln Thr Gly Thr Gly Thr Leu Lys Ile Asp		
145	150	155
Phe Val Gly Glu Leu Asn Asp Lys Met Lys Gly Phe Tyr Arg Ser Lys		
165	170	175
Tyr Thr Thr Pro Ser Gly Glu Val Arg Tyr Ala Ala Val Thr Gln Phe		
180	185	190
Glu Ala Thr Asp Ala Arg Arg Ala Phe Pro Cys Trp Asp Glu Pro Ala		
195	200	205
Ile Lys Ala Thr Phe Asp Ile Ser Leu Val Val Pro Lys Asp Arg Val		
210	215	220
Ala Leu Ser Asn Met Asn Val Ile Asp Arg Lys Pro Tyr Pro Asp Asp		
225	230	235
Glu Asn Leu Val Glu Val Lys Phe Ala Arg Thr Pro Val Met Ser Thr		
245	250	255
Tyr Leu Val Ala Phe Val Val Gly Glu Tyr Asp Phe Val Glu Thr Arg		
260	265	270
Ser Lys Asp Gly Val Cys Val Arg Val Tyr Thr Pro Val Gly Lys Ala		
275	280	285
Glu Gln Gly Lys Phe Ala Leu Glu Val Ala Ala Lys Thr Leu Pro Phe		
290	295	300
Tyr Lys Asp Tyr Phe Asn Val Pro Tyr Pro Leu Pro Lys Ile Asp Leu		
305	310	315
Ile Ala Ile Ala Asp Phe Ala Ala Gly Ala Met Glu Asn Trp Gly Leu		
325	330	335
Val Thr Tyr Arg Glu Thr Ala Leu Leu Ile Asp Pro Lys Asn Ser Cys		
340	345	350

Ser Ser Ser Arg Gln Trp Val Ala Leu Val Val Gly His Val Leu Ala		
355	360	365
His Gln Trp Phe Gly Asn Leu Val Thr Met Glu Trp Trp Thr His Leu		
370	375	380
Trp Leu Asn Glu Gly Phe Ala Ser Trp Ile Glu Tyr Leu Cys Val Asp		
385	390	400
His Cys Phe Pro Glu Tyr Asp Ile Trp Thr Gln Phe Val Ser Ala Asp		
405	410	415
Tyr Thr Arg Ala Gln Glu Leu Asp Ala Leu Asp Asn Ser His Pro Ile		
420	425	430
Glu Val Ser Val Gly His Pro Ser Glu Val Asp Glu Ile Phe Asp Ala		
435	440	445
Ile Ser Tyr Ser Lys Gly Ala Ser Val Ile Arg Met Leu His Asp Tyr		
450	455	460
Ile Gly Asp Lys Asp Phe Lys Lys Gly Met Asn Met Tyr Leu Thr Lys		
465	470	480
Phe Gln Gln Lys Asn Ala Ala Thr Glu Asp Leu Trp Glu Ser Leu Glu		
485	490	495
Asn Ala Ser Gly Lys Pro Ile Ala Ala Val Met Asn Thr Trp Thr Lys		
500	505	510
Gln Met Gly Phe Pro Leu Ile Tyr Val Glu Ala Glu Gln Val Glu Asp		
515	520	525
Asp Arg Leu Leu Arg Leu Ser Gln Lys Lys Phe Cys Ala Gly Gly Ser		
530	535	540
Tyr Val Gly Glu Asp Cys Pro Gln Trp Met Val Pro Ile Thr Ile Ser		
545	550	560
Thr Ser Glu Asp Pro Asn Gln Ala Lys Leu Lys Ile Leu Met Asp Lys		
565	570	575
Pro Glu Met Asn Val Val Leu Lys Asn Val Lys Pro Asp Gln Trp Val		
580	585	590
Lys Leu Asn Leu Gly Thr Val Gly Phe Tyr Arg Thr Gln Tyr Ser Ser		
595	600	605
Ala Met Leu Glu Ser Leu Leu Pro Gly Ile Arg Asp Leu Ser Leu Pro		
610	615	620
Pro Val Asp Arg Leu Gly Leu Gln Asn Asp Leu Phe Ser Leu Ala Arg		
625	630	640
Ala Gly Ile Ile Ser Thr Val Glu Val Leu Lys Val Met Glu Ala Phe		
645	650	655
Val Asn Glu Pro Asn Tyr Thr Val Trp Ser Asp Leu Ser Cys Asn Leu		

660	665	670
Gly Ile Leu Ser Thr Leu Leu Ser His Thr Asp Phe Tyr Glu Glu Ile		
675	680	685
Gln Glu Phe Val Lys Asp Val Phe Ser Pro Ile Gly Glu Arg Leu Gly		
690	695	700
Trp Asp Pro Lys Pro Gly Glu Gly His Leu Asp Ala Leu Leu Arg Gly		
705	710	715
Leu Val Leu Gly Lys Leu Gly Lys Ala Gly His Lys Ala Thr Leu Glu		
725	730	735
Glu Ala Arg Arg Arg Phe Lys Asp His Val Glu Gly Lys Gln Ile Leu		
740	745	750
Ser Ala Asp Leu Arg Ser Pro Val Tyr Leu Thr Val Leu Lys His Gly		
755	760	765
Asp Gly Thr Thr Leu Asp Ile Met Leu Lys Leu His Lys Gln Ala Asp		
770	775	780
Met Gln Glu Glu Lys Asn Arg Ile Glu Arg Val Leu Gly Ala Thr Leu		
785	790	795
Leu Pro Asp Leu Ile Gln Lys Val Leu Thr Phe Ala Leu Ser Glu Glu		
805	810	815
Val Arg Pro Gln Asp Thr Val Ser Val Ile Gly Gly Val Ala Gly Gly		
820	825	830
Ser Lys His Gly Arg Lys Ala Ala Trp Lys Phe Ile Lys Asp Asn Trp		
835	840	845
Glu Glu Leu Tyr Asn Arg Tyr Gln Gly Gly Phe Leu Ile Ser Arg Leu		
850	855	860
Ile Lys Leu Ser Val Glu Gly Phe Ala Val Asp Lys Met Ala Gly Glu		
865	870	875
Val Lys Ala Phe Phe Glu Ser His Pro Ala Pro Ser Ala Glu Arg Thr		
885	890	895
Ile Gln Gln Cys Cys Glu Asn Ile Leu Leu Asn Ala Ala Trp Leu Lys		
900	905	910
Arg Asp Ala Glu Ser Ile His Gln Tyr Leu Leu Gln Arg Lys Ala Ser		
915	920	925
Pro Pro Thr Val		
930		
<210> 15		
<211> 864		
<212> PRT		
<213> 人工序列		

<220>

<223> 合成的

<400> 15

Met	Ile	Tyr	Glu	Phe	Val	Met	Thr	Asp	Pro	Lys	Ile	Lys	Tyr	Leu	Lys
1				5					10					15	
Asp	Tyr	Lys	Pro	Ser	Asn	Tyr	Leu	Ile	Asp	Glu	Thr	His	Leu	Ile	Phe
			20					25					30		
Glu	Leu	Asp	Glu	Ser	Lys	Thr	Arg	Val	Thr	Ala	Asn	Leu	Tyr	Ile	Val
		35					40					45			
Ala	Asn	Arg	Glu	Asn	Arg	Glu	Asn	Asn	Thr	Leu	Val	Leu	Asp	Gly	Val
	50					55				60					
Glu	Leu	Lys	Leu	Leu	Ser	Ile	Lys	Leu	Asn	Asn	Lys	His	Leu	Ser	Pro
65				70				75					80		
Ala	Glu	Phe	Ala	Val	Asn	Glu	Asn	Gln	Leu	Ile	Ile	Asn	Asn	Val	Pro
			85					90					95		
Glu	Lys	Phe	Val	Leu	Gln	Thr	Val	Val	Glu	Ile	Asn	Pro	Ser	Ala	Asn
			100					105					110		
Thr	Ser	Leu	Glu	Gly	Leu	Tyr	Lys	Ser	Gly	Asp	Val	Phe	Ser	Thr	Gln
		115					120					125			
Cys	Glu	Ala	Thr	Gly	Phe	Arg	Lys	Ile	Thr	Tyr	Tyr	Leu	Asp	Arg	Pro
	130					135						140			
Asp	Val	Met	Ala	Ala	Phe	Thr	Val	Lys	Ile	Ile	Ala	Asp	Lys	Lys	Lys
145				150				155					160		
Tyr	Pro	Ile	Ile	Leu	Ser	Asn	Gly	Asp	Lys	Ile	Asp	Ser	Gly	Asp	Ile
			165					170					175		
Ser	Asp	Asn	Gln	His	Phe	Ala	Val	Trp	Lys	Asp	Pro	Phe	Lys	Lys	Pro
		180						185					190		
Cys	Tyr	Leu	Phe	Ala	Leu	Val	Ala	Gly	Asp	Leu	Ala	Ser	Ile	Lys	Asp
	195					200						205			
Thr	Tyr	Ile	Thr	Lys	Ser	Gln	Arg	Lys	Val	Ser	Leu	Glu	Ile	Tyr	Ala
	210					215						220			
Phe	Lys	Gln	Asp	Ile	Asp	Lys	Cys	His	Tyr	Ala	Met	Gln	Ala	Val	Lys
225				230				235					240		
Asp	Ser	Met	Lys	Trp	Asp	Glu	Asp	Arg	Phe	Gly	Leu	Glu	Tyr	Asp	Leu
			245					250					255		
Asp	Thr	Phe	Met	Ile	Val	Ala	Val	Pro	Asp	Phe	Asn	Ala	Gly	Ala	Met
		260						265					270		
Glu	Asn	Lys	Gly	Leu	Asn	Ile	Phe	Asn	Thr	Lys	Tyr	Ile	Met	Ala	Ser
	275						280						285		

Asn Lys Thr Ala Thr Asp Lys Asp Phe Glu Leu Val Gln Ser Val Val		
290	295	300
Gly His Glu Tyr Phe His Asn Trp Thr Gly Asp Arg Val Thr Cys Arg		
305	310	315 320
Asp Trp Phe Gln Leu Ser Leu Lys Glu Gly Leu Thr Val Phe Arg Asp		
	325	330 335
Gln Glu Phe Thr Ser Asp Leu Asn Ser Arg Asp Val Lys Arg Ile Asp		
	340	345 350
Asp Val Arg Ile Ile Arg Ser Ala Gln Phe Ala Glu Asp Ala Ser Pro		
	355	360 365
Met Ser His Pro Ile Arg Pro Glu Ser Tyr Ile Glu Met Asn Asn Phe		
	370	375 380
Tyr Thr Val Thr Val Tyr Asn Lys Gly Ala Glu Ile Ile Arg Met Ile		
385	390	395 400
His Thr Leu Leu Gly Glu Glu Gly Phe Gln Lys Gly Met Lys Leu Tyr		
	405	410 415
Phe Glu Arg His Asp Gly Gln Ala Val Thr Cys Asp Asp Phe Val Asn		
	420	425 430
Ala Met Ala Asp Ala Asn Asn Arg Asp Phe Ser Leu Phe Lys Arg Trp		
	435	440 445
Tyr Ala Gln Ser Gly Thr Pro Asn Ile Lys Val Ser Glu Asn Tyr Asp		
450	455	460
Ala Ser Ser Gln Thr Tyr Ser Leu Thr Leu Glu Gln Thr Thr Leu Pro		
465	470	475 480
Thr Ala Asp Gln Lys Glu Lys Gln Ala Leu His Ile Pro Val Lys Met		
	485	490 495
Gly Leu Ile Asn Pro Glu Gly Lys Asn Ile Ala Glu Gln Val Ile Glu		
	500	505 510
Leu Lys Glu Gln Lys Gln Thr Tyr Thr Phe Glu Asn Ile Ala Ala Lys		
	515	520 525
Pro Val Ala Ser Leu Phe Arg Asp Phe Ser Ala Pro Val Lys Val Glu		
	530	535 540
His Lys Arg Ser Glu Lys Asp Leu Leu His Ile Val Lys Tyr Asp Asn		
545	550	555 560
Asn Ala Phe Asn Arg Trp Asp Ser Leu Gln Gln Ile Ala Thr Asn Ile		
	565	570 575
Ile Leu Asn Asn Ala Asp Leu Asn Asp Glu Phe Leu Asn Ala Phe Lys		
	580	585 590
Ser Ile Leu His Asp Lys Asp Leu Asp Lys Ala Leu Ile Ser Asn Ala		

595	600	605
Leu Leu Ile Pro Ile Glu Ser Thr Ile Ala Glu Ala Met Arg Val Ile		
610	615	620
Met Val Asp Asp Ile Val Leu Ser Arg Lys Asn Val Val Asn Gln Leu		
625	630	635
Ala Asp Lys Leu Lys Asp Asp Trp Leu Ala Val Tyr Gln Gln Cys Asn		
645	650	655
Asp Asn Lys Pro Tyr Ser Leu Ser Ala Glu Gln Ile Ala Lys Arg Lys		
660	665	670
Leu Lys Gly Val Cys Leu Ser Tyr Leu Met Asn Ala Ser Asp Gln Lys		
675	680	685
Val Gly Thr Asp Leu Ala Gln Gln Leu Phe Asp Asn Ala Asp Asn Met		
690	695	700
Thr Asp Gln Gln Thr Ala Phe Thr Glu Leu Leu Lys Ser Asn Asp Lys		
705	710	715
Gln Val Arg Asp Asn Ala Ile Asn Glu Phe Tyr Asn Arg Trp Arg His		
725	730	735
Glu Asp Leu Val Val Asn Lys Trp Leu Leu Ser Gln Ala Gln Ile Ser		
740	745	750
His Glu Ser Ala Leu Asp Ile Val Lys Gly Leu Val Asn His Pro Ala		
755	760	765
Tyr Asn Pro Lys Asn Pro Asn Lys Val Tyr Ser Leu Ile Gly Gly Phe		
770	775	780
Gly Ala Asn Phe Leu Gln Tyr His Cys Lys Asp Gly Leu Gly Tyr Ala		
785	790	795
Phe Met Ala Asp Thr Val Leu Ala Leu Asp Lys Phe Asn His Gln Val		
805	810	815
Ala Ala Arg Met Ala Arg Asn Leu Met Ser Trp Lys Arg Tyr Asp Ser		
820	825	830
Asp Arg Gln Ala Met Met Lys Asn Ala Leu Glu Lys Ile Lys Ala Ser		
835	840	845
Asn Pro Ser Lys Asn Val Phe Glu Ile Val Ser Lys Ser Leu Glu Ser		
850	855	860

<210> 16

<211> 366

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 16

Met	Gly	Ser	Ser	His	His	His	His	His	His	Ser	Ser	Gly	Met	Glu	Val
1				5					10					15	
Arg	Asn	Met	Val	Asp	Tyr	Glu	Leu	Leu	Lys	Lys	Val	Val	Glu	Ala	Pro
			20					25					30		
Gly	Val	Ser	Gly	Tyr	Glu	Phe	Leu	Gly	Ile	Arg	Asp	Val	Val	Ile	Glu
		35					40					45			
Glu	Ile	Lys	Asp	Tyr	Val	Asp	Glu	Val	Lys	Val	Asp	Lys	Leu	Gly	Asn
	50					55					60				
Val	Ile	Ala	His	Lys	Lys	Gly	Glu	Gly	Pro	Lys	Val	Met	Ile	Ala	Ala
65				70					75					80	
His	Met	Asp	Gln	Ile	Gly	Leu	Met	Val	Thr	His	Ile	Glu	Lys	Asn	Gly
			85					90					95		
Phe	Leu	Arg	Val	Ala	Pro	Ile	Gly	Gly	Val	Asp	Pro	Lys	Thr	Leu	Ile
			100					105					110		
Ala	Gln	Arg	Phe	Lys	Val	Trp	Ile	Asp	Lys	Gly	Lys	Phe	Ile	Tyr	Gly
			115				120					125			
Val	Gly	Ala	Ser	Val	Pro	Pro	His	Ile	Gln	Lys	Pro	Glu	Asp	Arg	Lys
	130					135					140				
Lys	Ala	Pro	Asp	Trp	Asp	Gln	Ile	Phe	Ile	Asp	Ile	Gly	Ala	Glu	Ser
145				150					155					160	
Lys	Glu	Glu	Ala	Glu	Asp	Met	Gly	Val	Lys	Ile	Gly	Thr	Val	Ile	Thr
			165					170					175		
Trp	Asp	Gly	Arg	Leu	Glu	Arg	Leu	Gly	Lys	His	Arg	Phe	Val	Ser	Ile
			180					185					190		
Ala	Phe	Asp	Asp	Arg	Ile	Ala	Val	Tyr	Thr	Ile	Leu	Glu	Val	Ala	Lys
		195				200						205			
Gln	Leu	Lys	Asp	Ala	Lys	Ala	Asp	Val	Tyr	Phe	Val	Ala	Thr	Val	Gln
	210					215					220				
Glu	Glu	Val	Gly	Leu	Arg	Gly	Ala	Arg	Thr	Ser	Ala	Phe	Gly	Ile	Glu
225				230					235					240	
Pro	Asp	Tyr	Gly	Phe	Ala	Ile	Asp	Val	Thr	Ile	Ala	Ala	Asp	Ile	Pro
			245					250					255		
Gly	Thr	Pro	Glu	His	Lys	Gln	Val	Thr	His	Leu	Gly	Lys	Gly	Thr	Ala
		260						265					270		
Ile	Lys	Ile	Met	Asp	Arg	Ser	Val	Ile	Cys	His	Pro	Thr	Ile	Val	Arg
	275					280						285			
Trp	Leu	Glu	Glu	Leu	Ala	Lys	Lys	His	Glu	Ile	Pro	Tyr	Gln	Leu	Glu
	290					295						300			

Ile	Leu	Leu	Gly	Gly	Gly	Thr	Asp	Ala	Gly	Ala	Ile	His	Leu	Thr	Lys
305					310					315					320
Ala	Gly	Val	Pro	Thr	Gly	Ala	Leu	Ser	Val	Pro	Ala	Arg	Tyr	Ile	His
					325					330					335
Ser	Asn	Thr	Glu	Val	Val	Asp	Glu	Arg	Asp	Val	Asp	Ala	Thr	Val	Glu
					340					345					350
Leu	Met	Thr	Lys	Ala	Leu	Glu	Asn	Ile	His	Glu	Leu	Lys	Ile		
					355					360					365
<210> 17															
<211> 408															
<212> PRT															
<213> 人工序列															
<220>															
<223> 合成的															
<400> 17															
Met	Asp	Ala	Phe	Thr	Glu	Asn	Leu	Asn	Lys	Leu	Ala	Glu	Leu	Ala	Ile
1				5					10					15	
Arg	Val	Gly	Leu	Asn	Leu	Glu	Glu	Gly	Gln	Glu	Ile	Val	Ala	Thr	Ala
				20					25					30	
Pro	Ile	Glu	Ala	Val	Asp	Phe	Val	Arg	Leu	Leu	Ala	Glu	Lys	Ala	Tyr
				35					40					45	
Glu	Asn	Gly	Ala	Ser	Leu	Phe	Thr	Val	Leu	Tyr	Gly	Asp	Asn	Leu	Ile
				50					55					60	
Ala	Arg	Lys	Arg	Leu	Ala	Leu	Val	Pro	Glu	Ala	His	Leu	Asp	Arg	Ala
65					70					75					80
Pro	Ala	Trp	Leu	Tyr	Glu	Gly	Met	Ala	Lys	Ala	Phe	His	Glu	Gly	Ala
					85					90					95
Ala	Arg	Leu	Ala	Val	Ser	Gly	Asn	Asp	Pro	Lys	Ala	Leu	Glu	Gly	Leu
				100					105					110	
Pro	Pro	Glu	Arg	Val	Gly	Arg	Ala	Gln	Gln	Ala	Gln	Ser	Arg	Ala	Tyr
				115					120					125	
Arg	Pro	Thr	Leu	Ser	Ala	Ile	Thr	Glu	Phe	Val	Thr	Asn	Trp	Thr	Ile
					130				135					140	
Val	Pro	Phe	Ala	His	Pro	Gly	Trp	Ala	Lys	Ala	Val	Phe	Pro	Gly	Leu
145					150					155					160
Pro	Glu	Glu	Glu	Ala	Val	Gln	Arg	Leu	Trp	Gln	Ala	Ile	Phe	Gln	Ala
					165					170					175
Thr	Arg	Val	Asp	Gln	Glu	Asp	Pro	Val	Ala	Ala	Trp	Glu	Ala	His	Asn
					180					185					190

Arg Val Leu His Ala Lys Val Ala Phe Leu Asn Glu Lys Arg Phe His		
195	200	205
Ala Leu His Phe Gln Gly Pro Gly Thr Asp Leu Thr Val Gly Leu Ala		
210	215	220
Glu Gly His Leu Trp Gln Gly Gly Ala Thr Pro Thr Lys Lys Gly Arg		
225	230	235 240
Leu Cys Asn Pro Asn Leu Pro Thr Glu Glu Val Phe Thr Ala Pro His		
	245	250 255
Arg Glu Arg Val Glu Gly Val Val Arg Ala Ser Arg Pro Leu Ala Leu		
	260	265 270
Ser Gly Gln Leu Val Glu Gly Leu Trp Ala Arg Phe Glu Gly Gly Val		
	275	280 285
Ala Val Glu Val Gly Ala Glu Lys Gly Glu Glu Val Leu Lys Lys Leu		
	290	295 300
Leu Asp Thr Asp Glu Gly Ala Arg Arg Leu Gly Glu Val Ala Leu Val		
305	310	315 320
Pro Ala Asp Asn Pro Ile Ala Lys Thr Gly Leu Val Phe Phe Asp Thr		
	325	330 335
Leu Phe Asp Glu Asn Ala Ala Ser His Ile Ala Phe Gly Gln Ala Tyr		
	340	345 350
Ala Glu Asn Leu Glu Gly Arg Pro Ser Gly Glu Glu Phe Arg Arg Arg		
	355	360 365
Gly Gly Asn Glu Ser Met Val His Val Asp Trp Met Ile Gly Ser Glu		
	370	375 380
Glu Val Asp Val Asp Gly Leu Leu Glu Asp Gly Thr Arg Val Pro Leu		
385	390	395 400
Met Arg Arg Gly Arg Trp Val Ile		
	405	

<210> 18

<211> 362

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 18

Met Ala Lys Leu Asp Glu Thr Leu Thr Met Leu Lys Ala Leu Thr Asp		
1	5	10 15
Ala Lys Gly Val Pro Gly Asn Glu Arg Glu Ala Arg Asp Val Met Lys		
	20	25 30

Thr Tyr Ile Ala Pro Tyr Ala Asp Glu Val Thr Thr Asp Gly Leu Gly		
35	40	45
Ser Leu Ile Ala Lys Lys Glu Gly Lys Ser Gly Gly Pro Lys Val Met		
50	55	60
Ile Ala Gly His Leu Asp Glu Val Gly Phe Met Val Thr Gln Ile Asp		
65	70	75
Asp Lys Gly Phe Ile Arg Phe Gln Thr Leu Gly Gly Trp Trp Ser Gln		
85	90	95
Val Met Leu Ala Gln Arg Val Thr Ile Val Thr Lys Lys Gly Asp Ile		
100	105	110
Thr Gly Val Ile Gly Ser Lys Pro Pro His Ile Leu Pro Ser Glu Ala		
115	120	125
Arg Lys Lys Pro Val Glu Ile Lys Asp Met Phe Ile Asp Ile Gly Ala		
130	135	140
Thr Ser Arg Glu Glu Ala Met Glu Trp Gly Val Arg Pro Gly Asp Met		
145	150	155
Ile Val Pro Tyr Phe Glu Phe Thr Val Leu Asn Asn Glu Lys Met Leu		
165	170	175
Leu Ala Lys Ala Trp Asp Asn Arg Ile Gly Cys Ala Val Ala Ile Asp		
180	185	190
Val Leu Lys Gln Leu Lys Gly Val Asp His Pro Asn Thr Val Tyr Gly		
195	200	205
Val Gly Thr Val Gln Glu Glu Val Gly Leu Arg Gly Ala Arg Thr Ala		
210	215	220
Ala Gln Phe Ile Gln Pro Asp Ile Ala Phe Ala Val Asp Val Gly Ile		
225	230	235
Ala Gly Asp Thr Pro Gly Val Ser Glu Lys Glu Ala Met Gly Lys Leu		
245	250	255
Gly Ala Gly Pro His Ile Val Leu Tyr Asp Ala Thr Met Val Ser His		
260	265	270
Arg Gly Leu Arg Glu Phe Val Ile Glu Val Ala Glu Glu Leu Asn Ile		
275	280	285
Pro His His Phe Asp Ala Met Pro Gly Val Gly Thr Asp Ala Gly Ala		
290	295	300
Ile His Leu Thr Gly Ile Gly Val Pro Ser Leu Thr Ile Ala Ile Pro		
305	310	315
Thr Arg Tyr Ile His Ser His Ala Ala Ile Leu His Arg Asp Asp Tyr		
325	330	335
Glu Asn Thr Val Lys Leu Leu Val Glu Val Ile Lys Arg Leu Asp Ala		

	340					345					350				
Asp	Lys	Val	Lys	Gln	Leu	Thr	Phe	Asp	Glu						
	355					360									
<210>	19														
<211>	490														
<212>	PRT														
<213>	人工序列														
<220>															
<223>	合成的														
<400>	19														
Met	Glu	Asp	Lys	Val	Trp	Ile	Ser	Met	Gly	Ala	Asp	Ala	Val	Gly	Ser
1				5					10					15	
Leu	Asn	Pro	Ala	Leu	Ser	Glu	Ser	Leu	Leu	Pro	His	Ser	Phe	Ala	Ser
	20					25					30				
Gly	Ser	Gln	Val	Trp	Ile	Gly	Glu	Val	Ala	Ile	Asp	Glu	Leu	Ala	Glu
	35					40					45				
Leu	Ser	His	Thr	Met	His	Glu	Gln	His	Asn	Arg	Cys	Gly	Gly	Tyr	Met
	50					55					60				
Val	His	Thr	Ser	Ala	Gln	Gly	Ala	Met	Ala	Ala	Leu	Met	Met	Pro	Glu
65					70					75					80
Ser	Ile	Ala	Asn	Phe	Thr	Ile	Pro	Ala	Pro	Ser	Gln	Gln	Asp	Leu	Val
	85					90					95				
Asn	Ala	Trp	Leu	Pro	Gln	Val	Ser	Ala	Asp	Gln	Ile	Thr	Asn	Thr	Ile
	100					105					110				
Arg	Ala	Leu	Ser	Ser	Phe	Asn	Asn	Arg	Phe	Tyr	Thr	Thr	Thr	Ser	Gly
	115					120					125				
Ala	Gln	Ala	Ser	Asp	Trp	Leu	Ala	Asn	Glu	Trp	Arg	Ser	Leu	Ile	Ser
	130					135					140				
Ser	Leu	Pro	Gly	Ser	Arg	Ile	Glu	Gln	Ile	Lys	His	Ser	Gly	Tyr	Asn
145					150					155					160
Gln	Lys	Ser	Val	Val	Leu	Thr	Ile	Gln	Gly	Ser	Glu	Lys	Pro	Asp	Glu
	165					170					175				
Trp	Val	Ile	Val	Gly	Gly	His	Leu	Asp	Ser	Thr	Leu	Gly	Ser	His	Thr
	180					185					190				
Asn	Glu	Gln	Ser	Ile	Ala	Pro	Gly	Ala	Asp	Asp	Asp	Ala	Ser	Gly	Ile
	195					200					205				
Ala	Ser	Leu	Ser	Glu	Ile	Ile	Arg	Val	Leu	Arg	Asp	Asn	Asn	Phe	Arg
	210					215					220				
Pro	Lys	Arg	Ser	Val	Ala	Leu	Met	Ala	Tyr	Ala	Ala	Glu	Glu	Val	Gly

225	230	235	240
Leu Arg Gly Ser Gln Asp Leu Ala Asn Gln Tyr Lys Ala Gln Gly Lys			
	245	250	255
Lys Val Val Ser Val Leu Gln Leu Asp Met Thr Asn Tyr Arg Gly Ser			
	260	265	270
Ala Glu Asp Ile Val Phe Ile Thr Asp Tyr Thr Asp Ser Asn Leu Thr			
	275	280	285
Gln Phe Leu Thr Thr Leu Ile Asp Glu Tyr Leu Pro Glu Leu Thr Tyr			
	290	295	300
Gly Tyr Asp Arg Cys Gly Tyr Ala Cys Ser Asp His Ala Ser Trp His			
305	310	315	320
Lys Ala Gly Phe Ser Ala Ala Met Pro Phe Glu Ser Lys Phe Lys Asp			
	325	330	335
Tyr Asn Pro Lys Ile His Thr Ser Gln Asp Thr Leu Ala Asn Ser Asp			
	340	345	350
Pro Thr Gly Asn His Ala Val Lys Phe Thr Lys Leu Gly Leu Ala Tyr			
	355	360	365
Val Ile Glu Met Ala Asn Ala Gly Ser Ser Gln Val Pro Asp Asp Ser			
	370	375	380
Val Leu Gln Asp Gly Thr Ala Lys Ile Asn Leu Ser Gly Ala Arg Gly			
385	390	395	400
Thr Gln Lys Arg Phe Thr Phe Glu Leu Ser Gln Ser Lys Pro Leu Thr			
	405	410	415
Ile Gln Thr Tyr Gly Gly Ser Gly Asp Val Asp Leu Tyr Val Lys Tyr			
	420	425	430
Gly Ser Ala Pro Ser Lys Ser Asn Trp Asp Cys Arg Pro Tyr Gln Asn			
	435	440	445
Gly Asn Arg Glu Thr Cys Ser Phe Asn Asn Ala Gln Pro Gly Ile Tyr			
	450	455	460
His Val Met Leu Asp Gly Tyr Thr Asn Tyr Asn Asp Val Ala Leu Lys			
465	470	475	480
Ala Ser Thr Gln His His His His His His			
	485	490	

<210> 20

<211> 494

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 20

Met	Glu	Asp	Lys	Val	Trp	Ile	Ser	Ile	Gly	Ser	Asp	Ala	Ser	Gln	Thr
1				5					10					15	
Val	Lys	Ser	Val	Met	Gln	Ser	Asn	Ala	Arg	Ser	Leu	Leu	Pro	Glu	Ser
			20					25						30	
Leu	Ala	Ser	Asn	Gly	Pro	Val	Trp	Val	Gly	Gln	Val	Asp	Tyr	Ser	Gln
			35					40						45	
Leu	Ala	Glu	Leu	Ser	His	His	Met	His	Glu	Asp	His	Gln	Arg	Cys	Gly
			50					55						60	
Gly	Tyr	Met	Val	His	Ser	Ser	Pro	Glu	Ser	Ala	Ile	Ala	Ala	Ser	Asn
65					70					75					80
Met	Pro	Gln	Ser	Leu	Val	Ala	Phe	Ser	Ile	Pro	Glu	Ile	Ser	Gln	Gln
					85					90					95
Asp	Thr	Val	Asn	Ala	Trp	Leu	Pro	Gln	Val	Asn	Ser	Gln	Ala	Ile	Thr
			100						105						110
Gly	Thr	Ile	Thr	Ser	Leu	Thr	Ser	Phe	Ile	Asn	Arg	Phe	Tyr	Thr	Thr
			115						120						125
Thr	Ser	Gly	Ala	Gln	Ala	Ser	Asp	Trp	Leu	Ala	Asn	Glu	Trp	Arg	Ser
			130						135						140
Leu	Ser	Ala	Ser	Leu	Pro	Asn	Ala	Ser	Val	Arg	Gln	Val	Ser	His	Phe
145					150					155					160
Gly	Tyr	Asn	Gln	Lys	Ser	Val	Val	Leu	Thr	Ile	Thr	Gly	Ser	Glu	Lys
					165					170					175
Pro	Asp	Glu	Trp	Ile	Val	Leu	Gly	Gly	His	Leu	Asp	Ser	Thr	Ile	Gly
			180						185						190
Ser	His	Thr	Asn	Glu	Gln	Ser	Val	Ala	Pro	Gly	Ala	Asp	Asp	Asp	Ala
			195						200						205
Ser	Gly	Ile	Ala	Ser	Val	Thr	Glu	Ile	Ile	Arg	Val	Leu	Ser	Glu	Asn
			210						215						220
Asn	Phe	Gln	Pro	Lys	Arg	Ser	Ile	Ala	Phe	Met	Ala	Tyr	Ala	Ala	Glu
225					230					235					240
Glu	Val	Gly	Leu	Arg	Gly	Ser	Gln	Asp	Leu	Ala	Asn	Gln	Tyr	Lys	Ala
					245					250					255
Glu	Gly	Lys	Gln	Val	Ile	Ser	Ala	Leu	Gln	Leu	Asp	Met	Thr	Asn	Tyr
			260						265						270
Lys	Gly	Ser	Val	Glu	Asp	Ile	Val	Phe	Ile	Thr	Asp	Tyr	Thr	Asp	Ser
			275						280						285
Asn	Leu	Thr	Thr	Phe	Leu	Ser	Gln	Leu	Val	Asp	Glu	Tyr	Leu	Pro	Ser
			290						295						300

Leu Thr Tyr Gly Phe Asp Thr Cys Gly Tyr Ala Cys Ser Asp His Ala			
305	310	315	320
Ser Trp His Lys Ala Gly Phe Ser Ala Ala Met Pro Phe Glu Ala Lys			
	325	330	335
Phe Asn Asp Tyr Asn Pro Met Ile His Thr Pro Asn Asp Thr Leu Gln			
	340	345	350
Asn Ser Asp Pro Thr Ala Ser His Ala Val Lys Phe Thr Lys Leu Gly			
	355	360	365
Leu Ala Tyr Ala Ile Glu Met Ala Ser Thr Thr Gly Gly Thr Pro Pro			
	370	375	380
Pro Thr Gly Asn Val Leu Lys Asp Gly Val Pro Val Asn Gly Leu Ser			
385	390	395	400
Gly Ala Thr Gly Ser Gln Val His Tyr Ser Phe Glu Leu Pro Ala Gln			
	405	410	415
Lys Asn Leu Gln Ile Ser Thr Ala Gly Gly Ser Gly Asp Val Asp Leu			
	420	425	430
Tyr Val Ser Phe Gly Ser Glu Ala Thr Lys Gln Asn Trp Asp Cys Arg			
	435	440	445
Pro Tyr Arg Asn Gly Asn Asn Glu Val Cys Thr Phe Ala Gly Ala Thr			
	450	455	460
Pro Gly Thr Tyr Ser Ile Met Leu Asp Gly Tyr Arg Gln Phe Ser Gly			
465	470	475	480
Val Thr Leu Lys Ala Ser Thr Gln His His His His His His			
	485	490	

<210> 21

<211> 877

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 21

Met Thr Gln Gln Pro Gln Ala Lys Tyr Arg His Asp Tyr Arg Ala Pro			
1	5	10	15
Asp Tyr Thr Ile Thr Asp Ile Asp Leu Asp Phe Ala Leu Asp Ala Gln			
	20	25	30
Lys Thr Thr Val Thr Ala Val Ser Lys Val Lys Arg Gln Gly Thr Asp			
	35	40	45
Val Thr Pro Leu Ile Leu Asn Gly Glu Asp Leu Thr Leu Ile Ser Val			
50	55	60	

Ser	Val	Asp	Gly	Gln	Ala	Trp	Pro	His	Tyr	Arg	Gln	Gln	Asp	Asn	Thr
65					70					75					80
Leu	Val	Ile	Glu	Gln	Leu	Pro	Ala	Asp	Phe	Thr	Leu	Thr	Ile	Val	Asn
				85					90					95	
Asp	Ile	His	Pro	Ala	Thr	Asn	Ser	Ala	Leu	Glu	Gly	Leu	Tyr	Leu	Ser
			100					105					110		
Gly	Glu	Ala	Leu	Cys	Thr	Gln	Cys	Glu	Ala	Glu	Gly	Phe	Arg	His	Ile
			115					120					125		
Thr	Tyr	Tyr	Leu	Asp	Arg	Pro	Asp	Val	Leu	Ala	Arg	Phe	Thr	Thr	Arg
			130					135				140			
Ile	Val	Ala	Asp	Lys	Ser	Arg	Tyr	Pro	Tyr	Leu	Leu	Ser	Asn	Gly	Asn
145					150					155					160
Arg	Val	Gly	Gln	Gly	Glu	Leu	Asp	Asp	Gly	Arg	His	Trp	Val	Lys	Trp
					165					170					175
Glu	Asp	Pro	Phe	Pro	Lys	Pro	Ser	Tyr	Leu	Phe	Ala	Leu	Val	Ala	Gly
			180						185					190	
Asp	Phe	Asp	Val	Leu	Gln	Asp	Lys	Phe	Ile	Thr	Arg	Ser	Gly	Arg	Glu
			195						200					205	
Val	Ala	Leu	Glu	Ile	Phe	Val	Asp	Arg	Gly	Asn	Leu	Asp	Arg	Ala	Asp
			210						215					220	
Trp	Ala	Met	Thr	Ser	Leu	Lys	Asn	Ser	Met	Lys	Trp	Asp	Glu	Thr	Arg
225					230					235					240
Phe	Gly	Leu	Glu	Tyr	Asp	Leu	Asp	Ile	Tyr	Met	Ile	Val	Ala	Val	Asp
					245					250					255
Phe	Phe	Asn	Met	Gly	Ala	Met	Glu	Asn	Lys	Gly	Leu	Asn	Val	Phe	Asn
			260						265					270	
Ser	Lys	Tyr	Val	Leu	Ala	Lys	Ala	Glu	Thr	Ala	Thr	Asp	Lys	Asp	Tyr
			275						280					285	
Leu	Asn	Ile	Glu	Ala	Val	Ile	Gly	His	Glu	Tyr	Phe	His	Asn	Trp	Thr
			290						295					300	
Gly	Asn	Arg	Val	Thr	Cys	Arg	Asp	Trp	Phe	Gln	Leu	Ser	Leu	Lys	Glu
305					310					315					320
Gly	Leu	Thr	Val	Phe	Arg	Asp	Gln	Glu	Phe	Ser	Ser	Asp	Leu	Gly	Ser
					325					330					335
Arg	Ser	Val	Asn	Arg	Ile	Glu	Asn	Val	Arg	Val	Met	Arg	Ala	Ala	Gln
			340						345					350	
Phe	Ala	Glu	Asp	Ala	Ser	Pro	Met	Ala	His	Ala	Ile	Arg	Pro	Asp	Lys
			355						360					365	
Val	Ile	Glu	Met	Asn	Asn	Phe	Tyr	Thr	Leu	Thr	Val	Tyr	Glu	Lys	Gly

370	375	380
Ser Glu Val Ile Arg Met Met His Thr Leu Leu Gly Glu Gln Gln Phe		
385	390	395
Gln Ala Gly Met Arg Leu Tyr Phe Glu Arg His Asp Gly Ser Ala Ala		400
405	410	415
Thr Cys Asp Asp Phe Val Gln Ala Met Glu Asp Val Ser Asn Val Asp		
420	425	430
Leu Ser Leu Phe Arg Arg Trp Tyr Ser Gln Ser Gly Thr Pro Leu Leu		
435	440	445
Thr Val His Asp Asp Tyr Asp Val Glu Lys Gln Gln Tyr His Leu Phe		
450	455	460
Val Ser Gln Lys Thr Leu Pro Thr Ala Asp Gln Pro Glu Lys Leu Pro		
465	470	475
Leu His Ile Pro Leu Asp Ile Glu Leu Tyr Asp Ser Lys Gly Asn Val		
485	490	495
Ile Pro Leu Gln His Asn Gly Leu Pro Val His His Val Leu Asn Val		
500	505	510
Thr Glu Ala Glu Gln Thr Phe Thr Phe Asp Asn Val Ala Gln Lys Pro		
515	520	525
Ile Pro Ser Leu Leu Arg Glu Phe Ser Ala Pro Val Lys Leu Asp Tyr		
530	535	540
Pro Tyr Ser Asp Gln Gln Leu Thr Phe Leu Met Gln His Ala Arg Asn		
545	550	555
Glu Phe Ser Arg Trp Asp Ala Ala Gln Ser Leu Leu Ala Thr Tyr Ile		
565	570	575
Lys Leu Asn Val Ala Lys Tyr Gln Gln Gln Gln Pro Leu Ser Leu Pro		
580	585	590
Ala His Val Ala Asp Ala Phe Arg Ala Ile Leu Leu Asp Glu His Leu		
595	600	605
Asp Pro Ala Leu Ala Ala Gln Ile Leu Thr Leu Pro Ser Glu Asn Glu		
610	615	620
Met Ala Glu Leu Phe Thr Thr Ile Asp Pro Gln Ala Ile Ser Thr Val		
625	630	635
His Glu Ala Ile Thr Arg Cys Leu Ala Gln Glu Leu Ser Asp Glu Leu		
645	650	655
Leu Ala Val Tyr Val Ala Asn Met Thr Pro Val Tyr Arg Ile Glu His		
660	665	670
Gly Asp Ile Ala Lys Arg Ala Leu Arg Asn Thr Cys Leu Asn Tyr Leu		
675	680	685

Ala Phe Gly Asp Glu Glu Phe Ala Asn Lys Leu Val Ser Leu Gln Tyr			
690	695	700	
His Gln Ala Asp Asn Met Thr Asp Ser Leu Ala Ala Leu Ala Ala Ala			
705	710	715	720
Val Ala Ala Gln Leu Pro Cys Arg Asp Glu Leu Leu Ala Ala Phe Asp			
	725	730	735
Val Arg Trp Asn His Asp Gly Leu Val Met Asp Lys Trp Phe Ala Leu			
	740	745	750
Gln Ala Thr Ser Pro Ala Ala Asn Val Leu Val Gln Val Arg Thr Leu			
	755	760	765
Leu Lys His Pro Ala Phe Ser Leu Ser Asn Pro Asn Arg Thr Arg Ser			
	770	775	780
Leu Ile Gly Ser Phe Ala Ser Gly Asn Pro Ala Ala Phe His Ala Ala			
785	790	795	800
Asp Gly Ser Gly Tyr Gln Phe Leu Val Glu Ile Leu Ser Asp Leu Asn			
	805	810	815
Thr Arg Asn Pro Gln Val Ala Ala Arg Leu Ile Glu Pro Leu Ile Arg			
	820	825	830
Leu Lys Arg Tyr Asp Ala Gly Arg Gln Ala Leu Met Arg Lys Ala Leu			
	835	840	845
Glu Gln Leu Lys Thr Leu Asp Asn Leu Ser Gly Asp Leu Tyr Glu Lys			
	850	855	860
Ile Thr Lys Ala Leu Ala Ala His His His His His His			
865	870	875	
<210>	22		
<211>	489		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	合成的		
<400>	22		
Met Glu Glu Lys Val Trp Ile Ser Ile Gly Gly Asp Ala Thr Gln Thr			
1	5	10	15
Ala Leu Arg Ser Gly Ala Gln Ser Leu Leu Pro Glu Asn Leu Ile Asn			
	20	25	30
Gln Thr Ser Val Trp Val Gly Gln Val Pro Val Ser Glu Leu Ala Thr			
	35	40	45
Leu Ser His Glu Met His Glu Asn His Gln Arg Cys Gly Gly Tyr Met			
50	55	60	

Val His Pro Ser Ala Gln Ser Ala Met Ser Val Ser Ala Met Pro Leu			
65	70	75	80
Asn Leu Asn Ala Phe Ser Ala Pro Glu Ile Thr Gln Gln Thr Thr Val			
	85	90	95
Asn Ala Trp Leu Pro Ser Val Ser Ala Gln Gln Ile Thr Ser Thr Ile			
	100	105	110
Thr Thr Leu Thr Gln Phe Lys Asn Arg Phe Tyr Thr Thr Ser Thr Gly			
	115	120	125
Ala Gln Ala Ser Asn Trp Ile Ala Asp His Trp Arg Ser Leu Ser Ala			
	130	135	140
Ser Leu Pro Ala Ser Lys Val Glu Gln Ile Thr His Ser Gly Tyr Asn			
145	150	155	160
Gln Lys Ser Val Met Leu Thr Ile Thr Gly Ser Glu Lys Pro Asp Glu			
	165	170	175
Trp Val Val Ile Gly Gly His Leu Asp Ser Thr Leu Gly Ser Arg Thr			
	180	185	190
Asn Glu Ser Ser Ile Ala Pro Gly Ala Asp Asp Ala Ser Gly Ile			
	195	200	205
Ala Gly Val Thr Glu Ile Ile Arg Leu Leu Ser Glu Gln Asn Phe Arg			
	210	215	220
Pro Lys Arg Ser Ile Ala Phe Met Ala Tyr Ala Ala Glu Glu Val Gly			
225	230	235	240
Leu Arg Gly Ser Gln Asp Leu Ala Asn Arg Phe Lys Ala Glu Gly Lys			
	245	250	255
Lys Val Met Ser Val Met Gln Leu Asp Met Thr Asn Tyr Gln Gly Ser			
	260	265	270
Arg Glu Asp Ile Val Phe Ile Thr Asp Tyr Thr Asp Ser Asn Phe Thr			
	275	280	285
Gln Tyr Leu Thr Gln Leu Leu Asp Glu Tyr Leu Pro Ser Leu Thr Tyr			
	290	295	300
Gly Phe Asp Thr Cys Gly Tyr Ala Cys Ser Asp His Ala Ser Trp His			
305	310	315	320
Ala Val Gly Tyr Pro Ala Ala Met Pro Phe Glu Ser Lys Phe Asn Asp			
	325	330	335
Tyr Asn Pro Asn Ile His Ser Pro Gln Asp Thr Leu Gln Asn Ser Asp			
	340	345	350
Pro Thr Gly Phe His Ala Val Lys Phe Thr Lys Leu Gly Leu Ala Tyr			
	355	360	365
Val Val Glu Met Gly Asn Ala Ser Thr Pro Pro Thr Pro Ser Asn Gln			

370	375	380
Leu Lys Asn Gly Val Pro Val Asn Gly Leu Ser Ala Ser Arg Asn Ser		
385	390	395
Lys Thr Trp Tyr Gln Phe Glu Leu Gln Glu Ala Gly Asn Leu Ser Ile		400
405	410	415
Val Leu Ser Gly Gly Ser Gly Asp Ala Asp Leu Tyr Val Lys Tyr Gln		
420	425	430
Thr Asp Ala Asp Leu Gln Gln Tyr Asp Cys Arg Pro Tyr Arg Ser Gly		
435	440	445
Asn Asn Glu Thr Cys Gln Phe Ser Asn Ala Gln Pro Gly Arg Tyr Ser		
450	455	460
Ile Leu Leu His Gly Tyr Asn Asn Tyr Ser Asn Ala Ser Leu Val Ala		
465	470	475
Asn Ala Gln His His His His His His		480
485		
<210> 23		
<211> 488		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成的		
<400> 23		
Met Glu Asp Lys Lys Val Trp Ile Ser Ile Gly Ala Asp Ala Gln Gln		
1	5	10
Thr Ala Leu Ser Ser Gly Ala Gln Pro Leu Leu Ala Gln Ser Val Ala		15
20	25	30
His Asn Gly Gln Ala Trp Ile Gly Glu Val Ser Glu Ser Glu Leu Ala		
35	40	45
Ala Leu Ser His Glu Met His Glu Asn His His Arg Cys Gly Gly Tyr		
50	55	60
Ile Val His Ser Ser Ala Gln Ser Ala Met Ala Ala Ser Asn Met Pro		
65	70	75
Leu Ser Arg Ala Ser Phe Ile Ala Pro Ala Ile Ser Gln Gln Ala Leu		
85	90	95
Val Thr Pro Trp Ile Ser Gln Ile Asp Ser Ala Leu Ile Val Asn Thr		
100	105	110
Ile Asp Arg Leu Thr Asp Phe Pro Asn Arg Phe Tyr Thr Thr Thr Ser		
115	120	125
Gly Ala Gln Ala Ser Asp Trp Ile Lys Gln Arg Trp Gln Ser Leu Ser		

130	135	140
Ala Gly Leu Ala Gly Ala Ser Val Thr Gln Ile Ser His Ser Gly Tyr		
145	150	155
Asn Gln Ala Ser Val Met Leu Thr Ile Glu Gly Ser Glu Ser Pro Asp		
	165	170
Glu Trp Val Val Val Gly Gly His Leu Asp Ser Thr Ile Gly Ser Arg		
	180	185
Thr Asn Glu Gln Ser Ile Ala Pro Gly Ala Asp Asp Asp Ala Ser Gly		
	195	200
Ile Ala Ala Val Thr Glu Val Ile Arg Val Leu Ala Gln Asn Asn Phe		
	210	220
Gln Pro Lys Arg Ser Ile Ala Phe Val Ala Tyr Ala Ala Glu Glu Val		
225	230	235
Gly Leu Arg Gly Ser Gln Asp Val Ala Asn Gln Phe Lys Gln Ala Gly		
	245	250
Lys Asp Val Arg Gly Val Leu Gln Leu Asp Met Thr Asn Tyr Gln Gly		
	260	265
Ser Ala Glu Asp Ile Val Phe Ile Thr Asp Tyr Thr Asp Asn Gln Leu		
	275	280
Thr Gln Tyr Leu Thr Gln Leu Leu Asp Glu Tyr Leu Pro Thr Leu Asn		
	290	295
Tyr Gly Phe Asp Thr Cys Gly Tyr Ala Cys Ser Asp His Ala Ser Trp		
305	310	315
His Gln Val Gly Tyr Pro Ala Ala Met Pro Phe Glu Ala Lys Phe Asn		
	325	330
Asp Tyr Asn Pro Asn Ile His Thr Pro Gln Asp Thr Leu Ala Asn Ser		
	340	345
Asp Ser Glu Gly Ala His Ala Ala Lys Phe Thr Lys Leu Gly Leu Ala		
	355	360
Tyr Thr Val Glu Leu Ala Asn Ala Asp Ser Ser Pro Asn Pro Gly Asn		
	370	375
Glu Leu Lys Leu Gly Glu Pro Ile Asn Gly Leu Ser Gly Ala Arg Gly		
385	390	395
Asn Glu Lys Tyr Phe Asn Tyr Arg Leu Asp Gln Ser Gly Glu Leu Val		
	405	410
Ile Arg Thr Tyr Gly Gly Ser Gly Asp Val Asp Leu Tyr Val Lys Ala		
	420	425
Asn Gly Asp Val Ser Thr Gly Asn Trp Asp Cys Arg Pro Tyr Arg Ser		
	435	440
		445

Gly Asn Asp Glu Val Cys Arg Phe Asp Asn Ala Thr Pro Gly Asn Tyr
 450 455 460
 Ala Val Met Leu Arg Gly Tyr Arg Thr Tyr Asp Asn Val Ser Leu Ile
 465 470 475 480
 Val Glu His His His His His His
 485
 <210> 24
 <211> 308
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成的
 <400> 24
 Gly Met Pro Pro Ile Thr Gln Gln Ala Thr Val Thr Ala Trp Leu Pro
 1 5 10 15
 Gln Val Asp Ala Ser Gln Ile Thr Gly Thr Ile Ser Ser Leu Glu Ser
 20 25 30
 Phe Thr Asn Arg Phe Tyr Thr Thr Thr Ser Gly Ala Gln Ala Ser Asp
 35 40 45
 Trp Ile Ala Ser Glu Trp Gln Phe Leu Ser Ala Ser Leu Pro Asn Ala
 50 55 60
 Ser Val Lys Gln Val Ser His Ser Gly Tyr Asn Gln Lys Ser Val Val
 65 70 75 80
 Met Thr Ile Thr Gly Ser Glu Ala Pro Asp Glu Trp Ile Val Ile Gly
 85 90 95
 Gly His Leu Asp Ser Thr Ile Gly Ser His Thr Asn Glu Gln Ser Val
 100 105 110
 Ala Pro Gly Ala Asp Asp Asp Ala Ser Gly Ile Ala Ala Val Thr Glu
 115 120 125
 Val Ile Arg Val Leu Ser Glu Asn Asn Phe Gln Pro Lys Arg Ser Ile
 130 135 140
 Ala Phe Met Ala Tyr Ala Ala Glu Glu Val Gly Leu Arg Gly Ser Gln
 145 150 155 160
 Asp Leu Ala Asn Gln Tyr Lys Ser Glu Gly Lys Asn Val Val Ser Ala
 165 170 175
 Leu Gln Leu Asp Met Thr Asn Tyr Lys Gly Ser Ala Gln Asp Val Val
 180 185 190
 Phe Ile Thr Asp Tyr Thr Asp Ser Asn Phe Thr Gln Tyr Leu Thr Gln
 195 200 205

Leu Met Asp Glu Tyr Leu Pro Ser Leu Thr Tyr Gly Phe Asp Thr Cys		
210	215	220
Gly Tyr Ala Cys Ser Asp His Ala Ser Trp His Asn Ala Gly Tyr Pro		
225	230	235 240
Ala Ala Met Pro Phe Glu Ser Lys Phe Asn Asp Tyr Asn Pro Arg Ile		
	245	250 255
His Thr Thr Gln Asp Thr Leu Ala Asn Ser Asp Pro Thr Gly Ser His		
	260	265 270
Ala Lys Lys Phe Thr Gln Leu Gly Leu Ala Tyr Ala Ile Glu Met Gly		
	275	280 285
Ser Ala Thr Gly Asp Thr Pro Thr Pro Gly Asn Gln Leu Glu His His		
	290	295 300
His His His His		
305		
<210>	25	
<211>	354	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	25	
Met Val Asp Trp Glu Leu Met Lys Lys Ile Ile Glu Ser Pro Gly Val		
1	5	10 15
Ser Gly Tyr Glu His Leu Gly Ile Arg Asp Leu Val Val Asp Ile Leu		
	20	25 30
Lys Asp Val Ala Asp Glu Val Lys Ile Asp Lys Leu Gly Asn Val Ile		
	35	40 45
Ala His Phe Lys Gly Ser Ala Pro Lys Val Met Val Ala Ala His Met		
	50	55 60
Asp Lys Ile Gly Leu Met Val Asn His Ile Asp Lys Asp Gly Tyr Leu		
65	70	75 80
Arg Val Val Pro Ile Gly Gly Val Leu Pro Glu Thr Leu Ile Ala Gln		
	85	90 95
Lys Ile Arg Phe Phe Thr Glu Lys Gly Glu Arg Tyr Gly Val Val Gly		
	100	105 110
Val Leu Pro Pro His Leu Arg Arg Glu Ala Lys Asp Gln Gly Gly Lys		
	115	120 125
Ile Asp Trp Asp Ser Ile Ile Val Asp Val Gly Ala Ser Ser Arg Glu		
	130	135 140

Glu	Ala	Glu	Glu	Met	Gly	Phe	Arg	Ile	Gly	Thr	Ile	Gly	Glu	Phe	Ala
145					150				155						160
Pro	Asn	Phe	Thr	Arg	Leu	Ser	Glu	His	Arg	Phe	Ala	Thr	Pro	Tyr	Leu
				165					170						175
Asp	Asp	Arg	Ile	Cys	Leu	Tyr	Ala	Met	Ile	Glu	Ala	Ala	Arg	Gln	Leu
			180					185						190	
Gly	Glu	His	Glu	Ala	Asp	Ile	Tyr	Ile	Val	Ala	Ser	Val	Gln	Glu	Glu
		195					200						205		
Ile	Gly	Leu	Arg	Gly	Ala	Arg	Val	Ala	Ser	Phe	Ala	Ile	Asp	Pro	Glu
	210					215					220				
Val	Gly	Ile	Ala	Met	Asp	Val	Thr	Phe	Ala	Lys	Gln	Pro	Asn	Asp	Lys
225					230					235					240
Gly	Lys	Ile	Val	Pro	Glu	Leu	Gly	Lys	Gly	Pro	Val	Met	Asp	Val	Gly
				245					250						255
Pro	Asn	Ile	Asn	Pro	Lys	Leu	Arg	Gln	Phe	Ala	Asp	Glu	Val	Ala	Lys
			260					265						270	
Lys	Tyr	Glu	Ile	Pro	Leu	Gln	Val	Glu	Pro	Ser	Pro	Arg	Pro	Thr	Gly
		275					280							285	
Thr	Asp	Ala	Asn	Val	Met	Gln	Ile	Asn	Arg	Glu	Gly	Val	Ala	Thr	Ala
	290					295					300				
Val	Leu	Ser	Ile	Pro	Ile	Arg	Tyr	Met	His	Ser	Gln	Val	Glu	Leu	Ala
305					310					315					320
Asp	Ala	Arg	Asp	Val	Asp	Asn	Thr	Ile	Lys	Leu	Ala	Lys	Ala	Leu	Leu
				325					330					335	
Glu	Glu	Leu	Lys	Pro	Met	Asp	Phe	Thr	Pro	Leu	Glu	His	His	His	His
			340					345						350	
His	His														

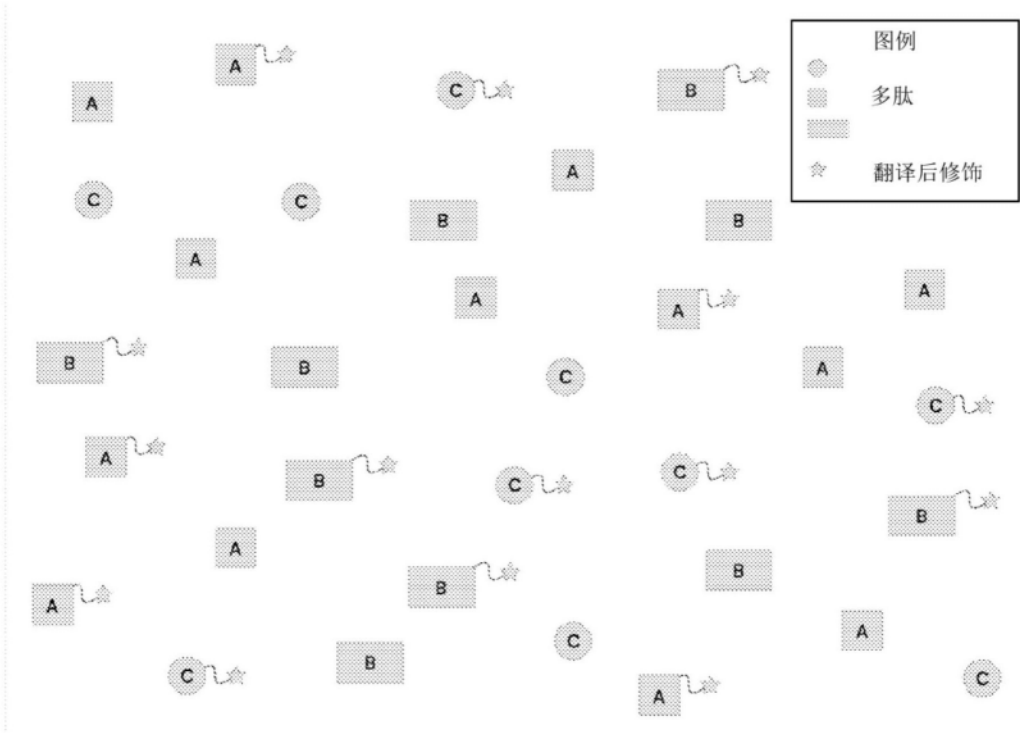


图1

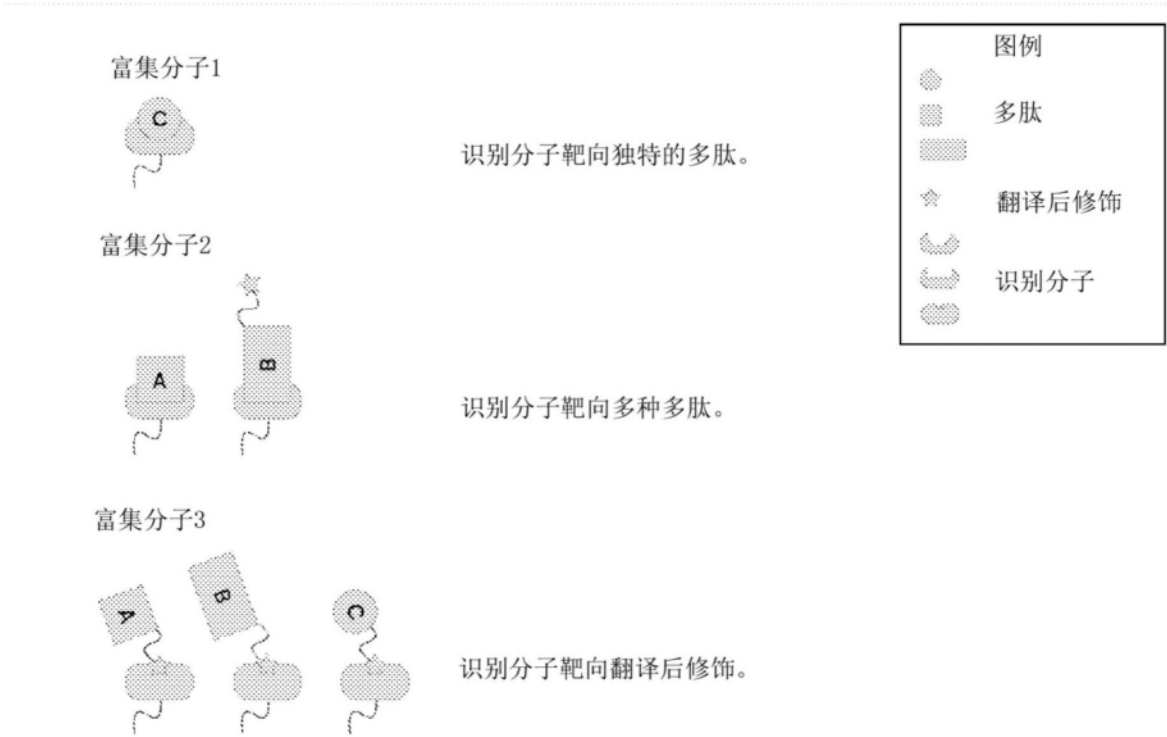


图2

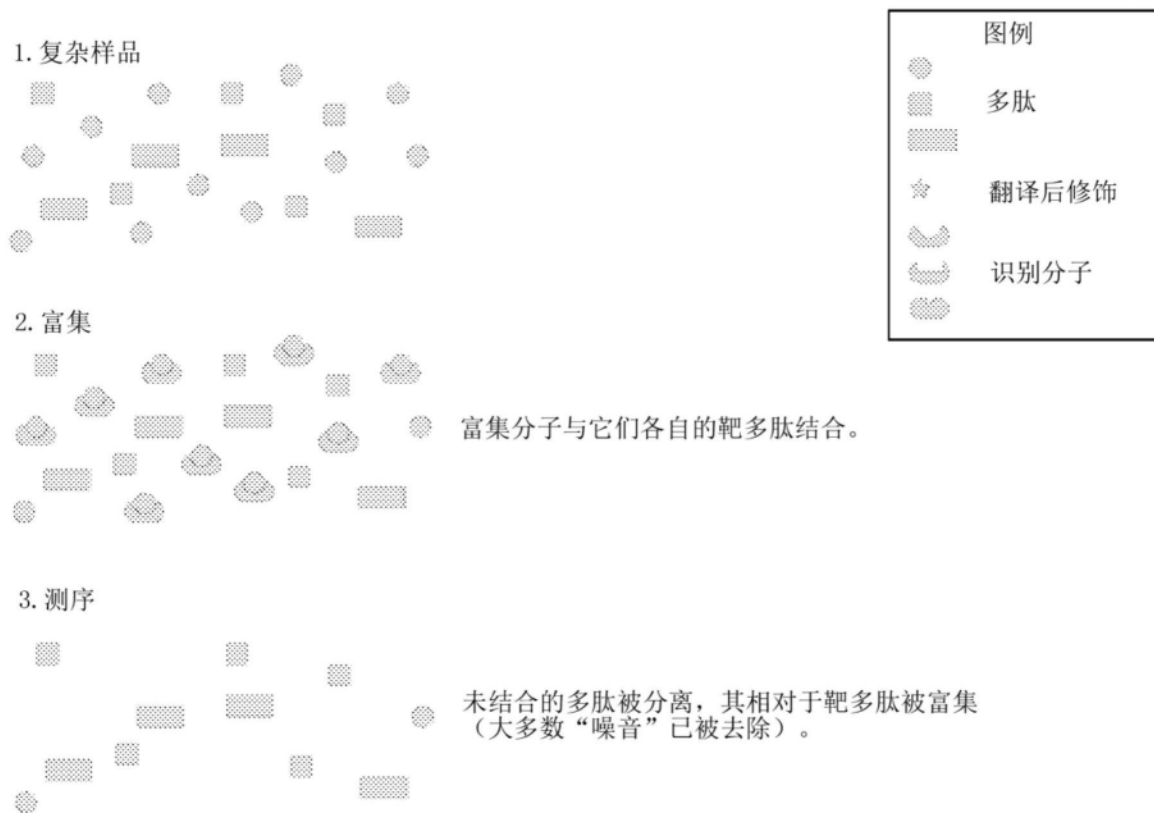
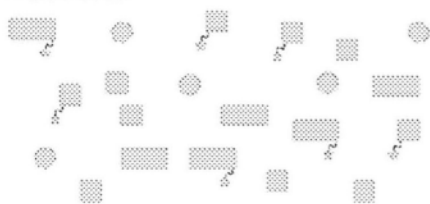
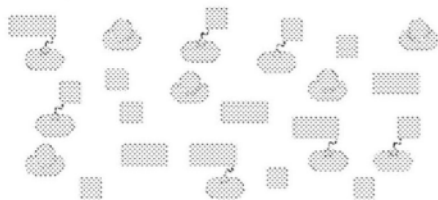


图3

1. 复杂样品

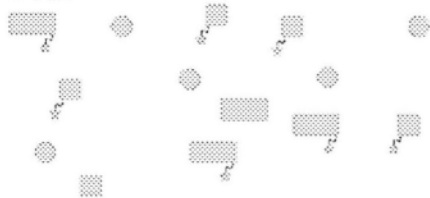


2. 富集



富集分子与它们各自的靶多肽结合。

3. 测序



结合的多肽被分离，其相对于非靶多肽被富集（大多数“噪音”已被去除）。



图4

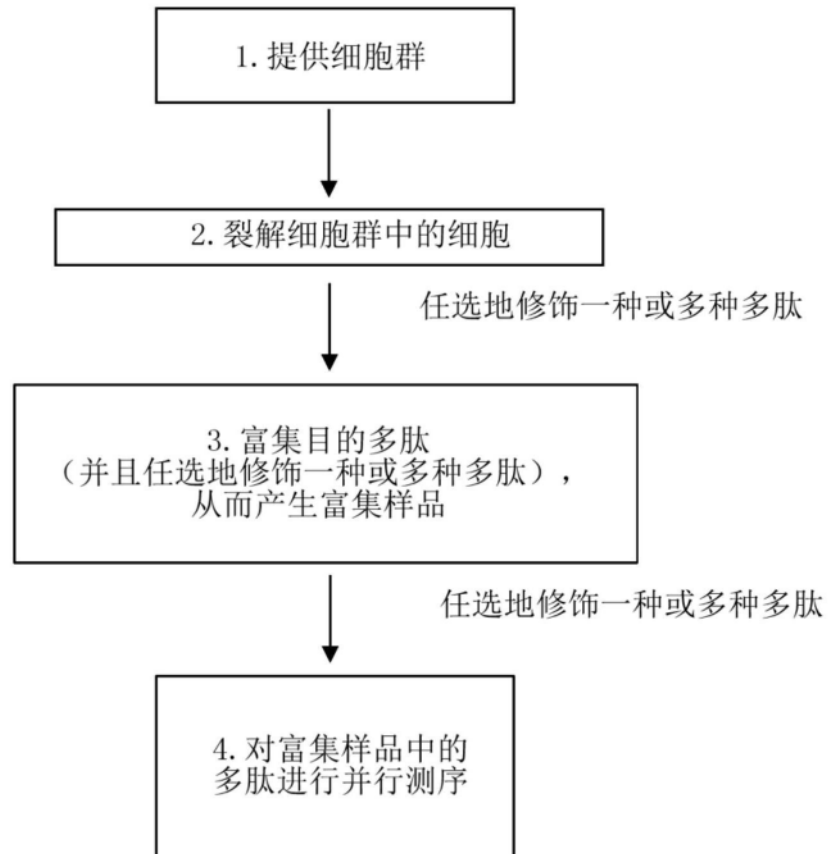


图5

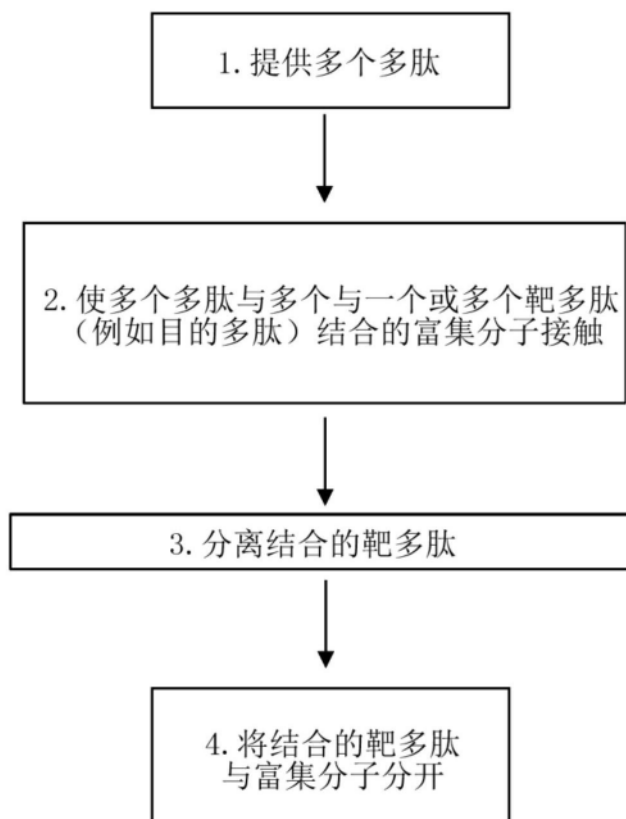


图6

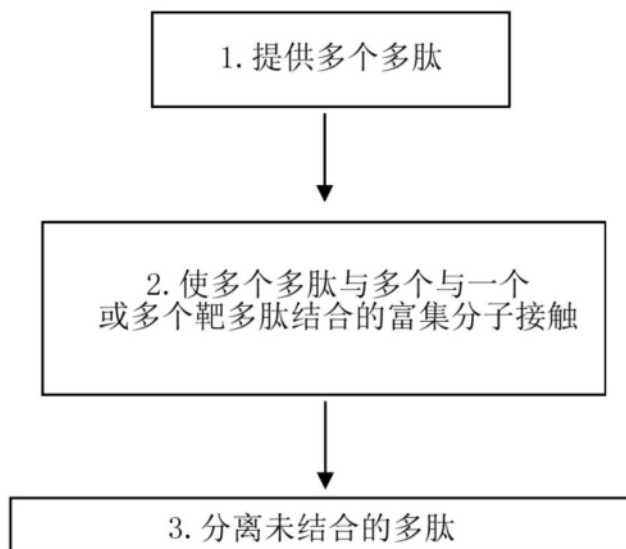


图7

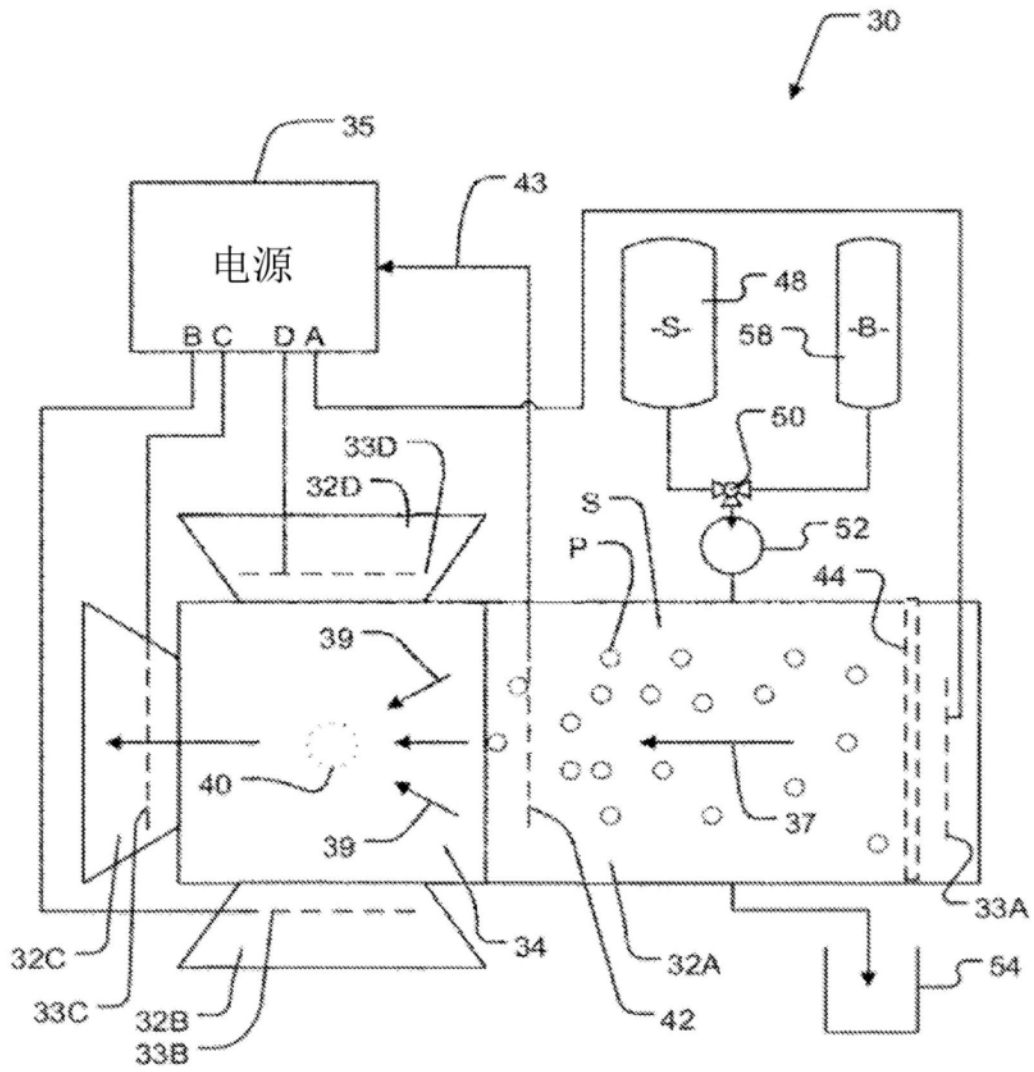


图8