

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200480022970.6

[51] Int. Cl.

C12N 15/10 (2006.01)

C07K 16/24 (2006.01)

[43] 公开日 2006 年 9 月 20 日

[11] 公开号 CN 1836041A

[22] 申请日 2004.6.25

[21] 申请号 200480022970.6

[30] 优先权

[32] 2003.6.27 [33] US [31] 60/483,282

[86] 国际申请 PCT/US2004/020306 2004.6.25

[87] 国际公布 WO2005/003345 英 2005.1.13

[85] 进入国家阶段日期 2006.2.10

[71] 申请人 比奥伦公司

地址 美国加利福尼亚

[72] 发明人 R·克里

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

代理人 罗菊华

权利要求书 6 页 说明书 30 页 序列表 7 页
附图 21 页

[54] 发明名称

浏览突变

[57] 摘要

本发明公开了一种突变方法，通过该方法可将预定的氨基酸导入到多肽的预选区域(或几个不同区域)的选定的一组位置的每个位置上，以此产生多肽类似物的文库。该方法是以某些氨基酸在蛋白质的结构和功能中具有重要作用为前提的。可以产生仅含期望的多肽类似物并具有对于筛选而言合理的大小的文库。该文库可以用于研究特定的氨基酸在多肽的结构和功能中的作用，以及用于开发新的或改进的多肽，如抗体、抗体片段、单链抗体、酶以及配体。

1. 生成多肽类似物文库的方法，在所述多肽类似物文库中，预定的氨基酸出现在该多肽指定区的每个位置处，该方法包括：

选择该多肽氨基酸序列的指定区；

确定将要在该指定区内每个氨基酸位置处进行置换的氨基酸残基；

按照下述标准合成各个编码指定区的多核苷酸，所述多核苷酸集合起来代表可能的变体多核苷酸，所述标准为：

i) 每个多核苷酸在指定区内每个密码子位置处或者含有该多肽的氨基酸残基所必需的密码子，或者含有预定氨基酸残基的密码子，以及

ii) 每个多核苷酸含有不超过 1 个预定氨基酸残基的密码子，由此生成多核苷酸文库，其中，预定的氨基酸残基出现在指定区内的每个氨基酸位置处。

2. 权利要求 1 的方法，其中，将多核苷酸合并在一起。

3. 权利要求 1 的方法，其中，突变所述多肽内的 2 个或者 2 个以上的指定区。

4. 权利要求 3 的方法，其中，选择相同的预定氨基酸用于在所述 2 个或者 2 个以上指定区的每一个内进行置换。

5. 权利要求 3 的方法，其中，选择不同的预定氨基酸用于分别在所述 2 个或者 2 个以上指定区的每一个内进行置换。

6. 上述任一权利要求的方法，其中，指定区包括多肽的功能域。

7. 权利要求 6 的方法，其中，功能域选自抗体结合位点、抗体构架区、抗体效应子区、受体结合位点以及催化位点。

8. 权利要求 7 的方法，其中，抗体结合位点或其部分包括选自 CDR1、CDR2、CDR3、CDR4、CDR5、CDR6 及它们的组合的 CDR 结构域。

9. 权利要求 7 的方法，其中，抗体构架区包括选自 FR1、FR2、FR3 以及它们的组合的结构域。

10. 权利要求 7 的方法，其中，抗体效应子区包括选自补体结合位

点以及 Fc 结合区的结构域。

11. 权利要求 1 的方法，其中，预定的氨基酸残基选自 Ser、Thr、Asn、Gln、Tyr、Cys、His、Glu、Asp、Lys、Arg、Ala、Gly、Ile、Leu、Met、Phe、Pro、Trp 和 Val。

12. 上述任一个权利要求的方法，其中，指定区包括至少约 3~40 个氨基酸。

13. 上述任一个权利要求的方法，其中，合成多核苷酸为表达文库。

14. 上述任一个权利要求的方法，其中，使用酶催化方法合成多核苷酸。

15. 上述任一个权利要求的方法，其中，使用多聚酶链反应合成多核苷酸。

16. 上述任一个权利要求的方法，其中，表达文库选自噬菌体展示文库、核糖体/多聚核糖体展示文库、酵母展示文库、细菌展示文库以及阵列文库。

17. 用权利要求 1 的方法制备的多肽类似物的文库。

18. 鉴定具有期望的结构或者功能的多肽的方法，该方法包括：

选择该多肽氨基酸序列的指定区；

确定将要在该指定区内每个氨基酸位置处进行置换的氨基酸残基；

根据下述标准合成编码指定区的多核苷酸，所述多核苷酸集合起来代表可能的变体多核苷酸，所述标准为：

i) 每个多核苷酸在指定区内每个密码子位置处或者含有合成该多肽的氨基酸残基所必需的密码子，或者含有 1 个预定氨基酸残基的密码子，以及

ii) 每个多核苷酸含有不超过 1 个预定氨基酸残基的密码子，

由此生成含有所述多核苷酸的表达文库；

使该表达文库表达以产生多肽类似物；和

筛选该多肽类似物来选择具有期望的结构或功能的多肽。

19. 权利要求 18 的方法，其中，该方法还包括鉴定编码所选择的多

肽类似物的多核苷酸的步骤。

20. 权利要求 18 的方法，其中，筛选包括：

用靶底物接触多肽，该多肽与编码该多肽的多核苷酸相关联，该多核苷酸还包括可检测的部分，以检测出能结合靶底物的变体多肽，并由此鉴定其为由该多核苷酸所编码。

21. 权利要求 20 的方法，其中，可检测的部分选自荧光部分、UV 部分以及可见光吸收部分。

22. 权利要求 20 的方法，其中，可检测的部分选自生物素部分、GST 部分以及 His 标签部分。

23. 权利要求 20 的方法，其中，使用核糖体展示使该多核苷酸与多肽类似物相关联。

24. 权利要求 18 的方法，其中，突变多肽内的 2 个或者 2 个以上的指定区。

25. 权利要求 24 的方法，其中，选择相同的预定氨基酸用于在所述 2 个或者 2 个以上指定区的每个指定区内进行置换。

26. 权利要求 24 的方法，其中，选择不同的预定氨基酸用于在所述 2 个或者 2 个以上指定区的每个指定区内进行置换。

27. 权利要求 18 的方法，其中，该多肽是单链抗体（sFVs）。

28. 权利要求 18 的方法，其中，该指定区包括该多肽的功能域。

29. 权利要求 18 的方法，其中，指定区包括选自 CDR1、CDR2、CDR3、CDR4、CDR5、CDR6 以及它们的组合的 CDR 或者其部分。

30. 权利要求 18 的方法，其中，指定区是包括选自 FR1、FR2、FR3、FR4 以及它们的组合的结构域的抗体构架区。

31. 权利要求 18 的方法，其中，指定区是包括选自补体结合位点和 Fc 结合区的结构域的抗体效应子区。

32. 权利要求 18 的方法，其中，预定氨基酸残基选自 Ser、Thr、Asn、Gln、Tyr、Cys、His、Glu、Asp、Lys、Arg、Ala、Gly、Ile、Leu、Met、Phe、Pro、Trp 以及 Val。

33. 编码多肽类似物的多核苷酸的文库，所述的多肽类似物包括 1

个或者1个以上指定区,其中在所述指定区内的每个氨基酸位置处进行预定氨基酸残基的置换,所述多核苷酸根据下述标准集合起来代表所有可能的变体,所述标准为:

- i) 每个多核苷酸在指定区的每个密码子位置处或者含有该多肽的氨基酸残基所必需的密码子,或者含有预定氨基酸残基的密码子,以及
- ii) 每个多核苷酸含有不超过1个预定氨基酸残基的密码子。

34. 权利要求33的文库,其中,突变在该多肽内的2个或者2个以上的指定区。

35. 权利要求33的文库,其中,选择相同的预定氨基酸用于在所述2个或者2个以上指定区的每一个内进行置换。

36. 权利要求33的文库,其中,选择不同的预定氨基酸用于在所述2个或者2个以上指定区的每一个内进行置换。

37. 权利要求33的文库,其中,该文库是表达文库。

38. 权利要求33的文库,其中,该表达文库选自噬菌体展示文库、核糖体/多聚核糖体展示文库、酵母展示文库、细菌展示文库和阵列文库。

39. 权利要求33的文库,其中,多核苷酸还包括1个或者更多个转录调节元件。

40. 权利要求39的文库,其中,当在体外进行转录和翻译时,该多核苷酸与由相应的多核苷酸编码的多肽相关联。

41. 权利要求40的文库,其中,使用核糖体/多聚核糖体展示使该多核苷酸与多肽相关联。

42. 权利要求41的文库,其中,多核苷酸包括RNA。

43. 权利要求42的文库,其中,多核苷酸还包含可检测部分。

44. 权利要求43的文库,其中,所述可检测部分包括荧光部分。

45. 权利要求33的文库,其中,该文库包括至少 10^6 个不同的多核苷酸。

46. 权利要求33的文库,其中,该文库包括至少 $45 \sim 10^{12}$ 个不同的多核苷酸。

47. 权利要求 33 的文库, 其中, 该多肽编码结合多肽。

48. 权利要求 47 的文库, 其中, 该结合多肽选自重链可变区 (V_H)、轻链可变区 (V_L) 以及单链抗体 (sFv)。

49. 权利要求 33 的文库, 其中, 该多核苷酸编码酶。

50. 权利要求 33 的文库, 其中, 该多核苷酸编码酶抑制剂。

51. 权利要求 33 的文库, 其中, 该多核苷酸编码催化性多肽。

52. 权利要求 33 的文库, 其中, 该文库固定在固体支持物上。

53. 权利要求 33 的文库, 其中, 该固体支持物是微芯片。

54. 权利要求 33 的文库, 其中, 该文库是阵列文库。

55. 一种微芯片, 其包括根据权利要求 33 的文库的固定化多核苷酸的阵列。

56. 根据权利要求 33 的文库鉴定得到的多肽类似物, 其中, 该多肽可与靶分子结合并且包括选自重链可变区 (V_H)、轻链可变区 (V_L) 以及单链抗体 (sFv) 的结合区。

57. 权利要求 56 的分子, 其中, 特定的靶分子是 $TNF\alpha$ 。

58. 权利要求 56 的分子, 其中, 该多肽与 SEQ ID NO: 1、2、3、4、5 或者 6 的氨基酸序列至少有 70% 一致性。

59. 鉴定具有期望的结构或者功能的多肽类似物的子集的方法, 该方法包括:

选择该多肽氨基酸序列的指定区;

确定将要在该指定区内的每个氨基酸位置处进行置换的氨基酸残基;

根据下述标准, 合成编码指定区的多核苷酸, 所述的多核苷酸集合起来代表可能的变体多核苷酸, 所述标准是:

i) 每个多核苷酸在指定区内每个密码子位置处或者含有合成该多肽的氨基酸残基所必需的密码子, 或者含有预定氨基酸残基的密码子, 以及

ii) 每个多核苷酸含有不超过 1 个预定氨基酸残基的密码子,

- 由此生成包括所述多核苷酸的表达文库；
将该表达文库置于使文库表达的条件下；
筛选该表达的文库以鉴定具有期望结构或者功能的多肽；
与对照标准比较该多肽的结构或者功能，其中，将符合或者超过对照标准的多肽分类为应答者，将不及对照标准的多肽分类为不应答者；
将应答者和不应答者在数据库中归类；和
查询该数据库以确定将要合成的子集多肽的序列。
60. 权利要求 59 的方法，其中，计算机辅助上述的 1 个或者更多个步骤。
61. 权利要求 59 的方法，其中，对照标准选自结合亲和性、稳定性和效应子功能。
62. 权利要求 59 的方法，其中，对照标准是对规定的底物的催化活性。
63. 权利要求 59 的方法，其中，该多肽选自重链可变区 (V_H)、轻链可变区 (V_L) 以及单链抗体 (sFv)。
64. 适于在电子装置中使用的、具有执行权利要求 59 的方法的 1 个或者更多个步骤的指令的介质。
65. 执行权利要求 59 的方法的 1 个或者更多个步骤的装置。

浏览突变

相关的申请和信息

本申请要求享有于 2003 年 6 月 27 日提交的美国临时申请 No. 60/483282 的优先权，其全部内容引入此处作为参考。下述说明书中引用的所有其它专利、专利申请以及参考文献的全部内容均引入此处作为参考。

背景技术

在蛋白质结构与功能的研究中，突变是一种非常有力的工具。可以在编码目的蛋白质的克隆基因的核苷酸序列上进行突变，并且可使该经过修饰的基因表达以产生蛋白质突变体。通过对野生型蛋白质与所产生的突变体进行比较，通常能够鉴定对蛋白质的结构完整性和/或生化功能具有重要作用的单个氨基酸或氨基酸区域，所述的功能有例如：结合和/或催化活性。但是，由单个蛋白质可生成的突变体的数量造成难以选择可提供信息或具有期望的特性的突变体，甚至所选的突变体仅在推定为蛋白质的重要区域（例如：构成蛋白质的活性位点的区域）中含有突变的情况下也很难选择。例如，特定氨基酸的置换、缺失或者插入可能会对蛋白质产生局部或者整体性的影响。

用于突变多肽的以往方法或者限制得过于严格，或者过于宽泛，或者局限于敲除蛋白质的功能而不是得到或改进功能。例如：一种具有很高限制性的方法是选择性突变或者定点突变，所述方法用于鉴定特定的功能位点的存在或者用以了解在功能位点进行非常具体的改变后的结果。定点突变常应用于磷酸化蛋白质的研究中，其中改变通常被磷酸化而使该多肽表现出其功能的氨基酸残基以证明磷酸化与功能活性间的关联。该方法对所研究的多肽与残基非常特异。

相反，高度宽泛的方法有饱和突变或者随机突变，这类方法被设

计成在基因或者蛋白质的指定区内产生包括所有可能的改变的大量突变。其依据的原理是通过产生相关蛋白质结构域的基本上所有可能的变体，氨基酸的适当排列有可能被产生为一个随机生成的突变体。不过，在实际应用中，由于存在过多不期望的候选物的所谓“噪音”，所生成的突变的大量的随机组合可能会妨碍有意义地筛选出期望的候选物的能力。

另一种方法称作“步行突变”(Walk Through mutagenesis) (参见，例如：美国专利号：5,830,650; 5,798,208)，该方法通过合成在统计学上包括目的突变组的简并寡核苷酸的混合物来突变多肽上的指定区。不过，由于使用了简并多核苷酸合成，导致步行突变除了产生期望的突变组之外，还产生大量的不期望的改变。例如：为了连续地将突变导入到仅有5个氨基酸位的整个指定区，必须产生一组100个以上的多核苷酸(然后筛选)(参见例如：图6)。相应地，制造并且筛选例如2个或者3个区域会变得越来越复杂，即，当有仅10到15个突变时，需要分别制备并且筛选200到300以上个多核苷酸。

其它的用于突变蛋白质的方法有丙氨酸扫描突变，该方法对蛋白质某部分的整个区域进行丙氨酸残基“扫描”，以此鉴定破坏蛋白质功能的位点。不过，该方法仅通过在给定位置替换成中性丙氨酸残基来检测蛋白质功能的丧失，不是得到或者改进其功能。所以在产生具有改进的结构或者功能的蛋白质方面，该方法不是一个有效的方法。

因此亟需可突变蛋白质以使之获得新的或者改进的功能的系统性方法。

发明内容

本发明涉及一种突变方法，该方法用于生成新型的或者改进的蛋白质(或者多肽)；本发明还涉及通过该方法生成的多肽类似物的文库以及具体的多肽。用于突变的目的多肽可以是天然的、合成的或者工程的多肽，包括其片段、类似物及突变体形式。

在一个实施方案中，该方法包括将预定的氨基酸导入到多肽氨基

酸序列的一个指定区(或者几个不同的区)内的基本上每一个位置中。生成这样的多肽文库,该文库包含多肽类似物,所述的类似物单个地具有不超过一个预定氨基酸,但是集合起来在所述指定区内的每个位置上具有预定的氨基酸。因为实际上是以单个的预定氨基酸(并且仅仅该预定的氨基酸)逐个位置地在多肽的一个或者多个指定区整个区域内进行置换,所以该方法被称为“浏览”突变(“look-through” mutagenesis)。所以,本发明允许人们浏览分别地以预定的氨基酸在多肽的指定区内每个氨基酸位置进行置换造成的结构上和功能上的后果,由此能够没有任何由于产生不期望的多肽类似物(即,包含有遵循“浏览”突变方法的氨基酸置换以外的氨基酸置换的类似物)而引起的干扰或者“噪音”地使指定区的特定蛋白质化学各自分离(参见,例如:图6)。

所以,本发明可以在多肽的一个或者多个指定区内对具体的氨基酸改变的作用进行高度有效且精确的系统性评价。当对两个或者更多指定区进行评价(通过突变)时,由于极大地增加了需要的多肽类似物的数量、并因此增加不期望的类似物的存在,这是尤其重要的。本发明通过完全地消除了不期望的类似物,并且因此消除了观察到的蛋白质结构或者功能上的任何变化并非预定氨基酸置换引起的结果的可能性,解决了这个问题。所以,能够以高度的准确性和有效性对蛋白质的多个区域的特定蛋白质化学分离的效果进行研究。重要的是,这包括研究突变能如何影响这些区域的相互作用,由此改进蛋白质的整体结构和功能。

在本发明的特定的实施方案中,通过首先合成编码多肽的指定区的各个多核苷酸,来生成和筛选多肽类似物的文库,其中,根据此处所述的浏览标准,该多核苷酸从总体上代表所有可能的变体多核苷酸。表达变体多核苷酸,例如:使用体外转录与翻译和/或使用展示技术诸如核糖体展示、噬菌体展示、细菌展示、酵母展示、阵列展示(arrayed display)或者任何其它在本领域内公知的适当的展示系统进行表达。

然后使用诸如结合分析或酶促/催化分析这样的功能性分析对表

达的多肽进行筛选和选择。在一个实施方案中，所表达的多肽以与编码该多肽的多核苷酸有关联的形式表达，以此可鉴定编码该多肽的多核苷酸序列。在其它实施方案中，使用蛋白质化学方法直接合成该多肽。

本发明提供了能够用于生成具有对于筛选而言切实可行的大小的多肽类似物文库的突变方法，这部分是因为该文库没有任何不期望的类似物多肽或者所谓的噪音。该方法可以用于研究特定氨基酸在多肽的结构和功能中的作用，以及用于开发新的或者改进的多肽，例如：抗体、其结合片段及类似物、单链抗体、催化性抗体、酶以及配体。另外，该方法可以利用如下先验信息，例如通过计算机模拟获得的信息来实施，所述信息可用于选择将要使用“浏览”突变来生成和研究的多肽类似物起始子集。

通过下述的描述和实施例能够显而易见地看出本发明的其它优点和方面。

附图说明

图 1 举例说明可以用浏览突变(Look-Through Mutagenesis, LTM)检测的示范性指定区(或者 D 区域)和用于从这种 D 区鉴定目的多肽类似物的功能性分析。

图 2 举例说明在抗体可变区的三个指定区(即，抗体重链可变区的 CDR1、CDR2 和 CDR3)中应用 LTM。可以以相似方式单独地或者与重链可变区组合对轻链可变区进行研究。为了便于后续的筛选分析，可以以所示的单链抗体(sFv)的形式对重链可变区进行研究。

图 3 举例说明在重链可变区的指定区(即，CDR1 的 31-25 位)内应用 LTM。

图 4 举例说明在重链可变区的指定区(即，CDR2 的 55-68 位)内应用 LTM。

图 5 举例说明在重链可变区的指定区(即，CDR3 的 101-111 位)内应用 LTM。

图 6 举例说明与 Walk-Through 突变相比, LTM 的优点。对具有代表性的指定区, 即, 抗体重链可变区的 CDR1, 进行 LTM, 结果导致顺序地改变了整个指定区内的每个氨基酸位置, 而且没有导入任何不期望的氨基酸残基或所谓的噪音。

图 7 举例说明在浏览突变后, 将蛋白质的三个指定区整合回到完整蛋白质环境中, 具体地, 在浏览突变后, 将抗体重链可变区的所有三个 CDR 区整合成单链抗体形式。

图 8 举例说明使用聚合酶链反应 (PCR) 将经过 LTM 的抗体重链和轻链的指定区构建到更大的基因环境中。

图 9 举例说明了为了获得包括例如丝氨酸、组氨酸、和/或天冬氨酸的催化位点, 抗体可变区的各 CDR 的示范性多样性形式, 以及它们可如何排列。

图 10 举例说明经过 LTM 的抗体结合区的所有六个 CDR 的整合, 以及如果使用 1 个预定氨基酸残基或者 20 个不同的预定氨基酸残基时最终产生的多样性。

图 11 举例说明经过 LTM 的抗 TNF 单链抗体 (sFv) 结合区的所有六个 CDR 的整合, 以及如果使用 1 个预定的氨基酸残基或者 3 个不同的预定的氨基酸残基时最终产生的多样性。

图 12 举例说明使用 LTM 可以获得的代表抗体结合区六个 CDR 的某些可能的多肽类似物的阵列文库。

图 13 举例说明使用无细胞的核糖体展示筛选阵列表达文库。

图 14 举例说明在对抗体可变区的结合区 (即: 所有 6 个 CDR) 进行 LTM 时所研究的组合化学。

图 15 举例说明可用于 LTM 的几个代表性的抗 TNF 结合分子的可变区 (以单链的形式) 序列。

图 16 显示为了筛选 LTM 的催化性候选物进行蛋白酶选择分析的示意图。

图 17 显示当在细菌细胞中进行蛋白酶选择分析时该分析的机制 (和优点) 的详细示意图。

图 18 显示为了获得催化活性例如催化性抗体活性,使用核糖体展示或者酵母展示筛选基因文库的机制的详细流程图。

图 19 显示的是实施 LTM 的示意图,其中,为了更有效地进行分子设计和开发,可以对先验的信息(例如:计算机模拟信息)和经验性信息(检测结果)进行协调。该有引导的 LTM 方法被称作“引导式”突变 (“guide-through” mutagenesis)。

图 20 显示的是假定的 V_H CDR3 野生型序列(最上面加有阴影的椭圆)以及在 LTM His 置换(无阴影的椭圆)的文库成员中所得到的序列。各个 LTM His 置换由各个寡核苷酸,例如以高通量的形式合成的寡核苷酸编码。以相似的方式构建在该 CDR 结构域中其它(LTM)氨基酸置换的子集文库(subset libraries)。

图 21 举例说明 scFv 文库的生成。在 X 轴的最上面的一行以及在 Y 轴的最左面的栏内用 3 个数字表示轻链和重链各自的 3 个 CDR。“0”表示野生型 CDR 序列;“1”表示 LTM 突变的 CDR。刻度上的数字表示子集文库的复杂性。例如,在矩阵的最左上角为“0”,其对应的 X 轴和 Y 轴为分别为“000”和“000”,这表示在 V_H 和 V_L 中的 CDR 均未经 LTM 突变。当从角落 0 处下移一行至邻近的网格位置“1”时,用 X 轴“100”、Y 轴“000”来指定,这表示突变了 V_H 的 CDR1,同时 V_L 的所有 CDR 保持野生型。这样以此类推,编号为“4”的网格指的是同时有 4 个 CDR 被突变。首先使用 SOE-PCR 制备 7 个 V_H 和 7 个 V_L 链(箭头所示)。然后用 mega 引物扩增、混合、配对,一步生成全部其余的 V_L - V_H 组合。

发明详述

为了清楚地理解本说明书和权利要求书,提供了下述定义。

定义

此处所用的术语“类似物”指的是具有 1 个或者更多个氨基酸置换的变体或者突变体多肽(或者编码这类多肽的核酸)。

术语“结合分子”指的是任何结合分子,包括蛋白质、多肽、肽

以及小分子，其可结合于底物或者靶。在一个实施方案中，结合分子为抗体、或其结合片段（如 Fab 片段）、单结构域抗体、单链抗体（例如：sFv）或者能结合配体的肽。

术语“指定区”指的是多肽上选定的区域。典型地，指定区包括功能性位点的全部或者一部分，所述功能性位点例如有配体的结合位点、结合分子或者受体的结合位点、或者催化位点。指定区还可以包括功能性位点的多个部分。例如：指定区可以包括抗体的互补决定区（CDR）或者完整的重和/或轻链可变区（VR）的所有部分、一部分、多个部分。从而功能性位点可以包括对该分子的功能性活性有贡献的单个或者多个指定区。

术语“文库”指的是根据本发明的方法突变的 2 个或者更多个分子。该文库分子可以是下述形式：多核苷酸、多肽、多核苷酸和多肽、无细胞的抽提物中的多核苷酸和多肽，或者是以下述形式：在噬菌体、原核细胞环境中或者在真核细胞中的多肽和/或多核苷酸。

术语“突变”指的是氨基酸序列的改变。这可以通过改变或者生产能编码该改变的氨基酸序列的核酸（多核苷酸），或者通过使用蛋白质化学直接合成改变的多肽来实现。

术语“多核苷酸”指的是核酸诸如 DNA 分子和 RNA 分子及其类似物（例如：使用核苷酸类似物或者使用核酸化学生成的 DNA 或者 RNA）。如果需要，可以使用合成的方法，例如使用本领域公知的核酸化学，或者使用酶催化的方法，例如使用聚合酶，制备该多核苷酸。有代表性的修饰包括：甲基化、生物素化和其它本领域已知的修饰。另外，该核酸分子可以是单链也可以是双链，而且如果需要可以连接或者结合（例如：共价或者非共价地）到可检测部分上。

术语“变体多核苷酸”指的是编码本发明的相应的多肽类似物（或者其部分）的多核苷酸。变体多核苷酸含有 1 个或者更多个已经改变、导致表达不同氨基酸的密码子。

术语“多肽”指的是以肽键连接的 2 个或者更多个氨基酸，例如：肽（例如：2 到约 50 个氨基酸残基）、以及更长的肽序列，例如：典

型地含有从少至 50 个氨基酸残基到超过 1000 个氨基酸残基的氨基酸序列的蛋白质序列。

术语“合并 (pooling)”指的是组合多核苷酸变体或者多肽类似物以形成代表整个多肽区域的浏览突变的文库。这些分子可以是多核苷酸和/或多肽的形式,也可以以子文库 (sublibrary) 的形式、作为在固体支持物上的分子、作为在溶液中的分子、和/或作为在一个或者更多个生物体 (例如: 噬菌体、原核细胞或者真核细胞) 中的分子而共存。

术语“预定的氨基酸”指的是选定用于在将要突变的多肽的指定区域内每个位置处进行置换的氨基酸残基。这不包括在该区域内已经 (例如, 天然地) 含有该预定氨基酸、进而不需要用预定氨基酸置换的位置。所以根据本发明生成的每个多肽类似物在给定的指定区域中含有不超过 1 个“预定氨基酸”残基。不过, 在总体上, 所生成的多肽类似物文库在被突变的区域内的每个位置都含有预定的氨基酸。典型地, 根据与氨基酸侧链基团有关的特定的大小或化学性质选择预定的氨基酸。适当的预定氨基酸包括例如: 甘氨酸和丙氨酸 (空间上小); 丝氨酸、苏氨酸和半胱氨酸 (亲核性); 缬氨酸、亮氨酸、异亮氨酸、蛋氨酸和脯氨酸 (疏水性); 苯丙氨酸、酪氨酸和色氨酸 (芳香族); 天冬氨酸和谷氨酸 (酸性); 天冬酰胺、谷氨酰胺以及组氨酸 (酰胺 (amide)); 和赖氨酸和精氨酸 (碱性)。

详细说明

蛋白质研究已经揭示了某些氨基酸在其结构和功能上起重要作用。例如: 已经发现仅仅是不连续的多个氨基酸参与抗体对抗原的结合或者参与酶的催化事件。

尽管清楚某些氨基酸对蛋白质的活性或者功能是重要的, 但是难以鉴定哪些氨基酸参与其中, 它们是如何参与的, 以及怎样的置换能够改进该蛋白质的结构和功能。部分地, 上述困难是由于多肽中氨基酸侧链的空间构象的复杂性和对形成功能位点有贡献的该多肽不同部

分的相互作用造成的。例如：抗体的重链可变区和轻链可变区的 6 个 CDR 之间的相互作用对抗原或者配体结合口袋有贡献。

先前的突变方法如选择性（定点）突变和饱和突变由于复杂多肽的可能变异数量巨大而在进行蛋白质结构和功能的研究时用途有限。在希望的组合常常伴随有大量的不期望的组合或者所谓的噪音的情况下，尤其是如此。

本发明的方法提供了系统性的、可实际应用的以及具有高度精确性的、用于评价在多肽的指定区内特定氨基酸和它们的位置对多肽的结构或者功能的作用以及进而用于生成改进的多肽的方法。

1. 选择指定区

根据本发明，选择蛋白质内的指定区用于突变。典型地，认为该区域是对蛋白质的结构或者功能很重要的区域（参见例如：图 1）。这可例如通过已知的结构或者功能方面，或者通过将指定区与在其它蛋白质的研究中得知的信息相比较推测出来，并且可得到模拟信息的辅助。例如：该指定区可以是一个在诸如结合、催化或者其它功能的功能位点中具有作用的区域。在一个实施方案中，指定区为抗原结合分子的高变区或者互补决定区（CDR）。在另一个实施方案中，指定区为互补决定区（CDR）的一部分。在其它的实施方案中，选择两个或者多个指定区例如其 CDR 或者 CDR 的一部分用于突变。

2. 选择预定氨基酸残基

所选的用于在指定区内进行置换的氨基酸残基通常从已知可参与目的结构或者功能的氨基酸残基中进行选择。20 种天然存在的氨基酸在它们的侧链上有所不同。每个侧链决定了其化学特性，使每个氨基酸独一无二。为了实现改变结合或者创造新的结合亲和性的目的，可以选择 20 种天然存在的氨基酸中的任一氨基酸。先前的突变方法对每种置换产生了大量的类似物，这造成实际上无法进行对 20 种氨基酸的每个置换对蛋白质结合的影响的评价。与此相反，本发明的方法对每个氨基酸置换产生了切实可行数量的类似物，因此可以在蛋白质的隔开的区域内对更多不同的蛋白质化学进行评价。

与蛋白质结合相反，典型地仅有一小类氨基酸残基参与酶催化或者催化事件。例如：从侧链的化学特性出发，仅有选定数量的天然氨基酸偏好参与催化事件。这些氨基酸属于极性和中性氨基酸组，诸如：丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、酪氨酸和半胱氨酸；带电荷氨基酸组诸如天冬氨酸和谷氨酸、赖氨酸和精氨酸，特别是组氨酸。其它极性和中性侧链是半胱氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺和酪氨酸的侧链。认为甘氨酸是这一组的边界成员。丝氨酸和苏氨酸对于氢键形成具有重要作用。苏氨酸还在 β 碳上具有不对称性，所以仅使用一个立体异构体。酰胺谷氨酰胺和天冬酰胺也能形成氢键，酰氨基起氢供体的作用，羧基起受体的作用。谷氨酰胺比天冬酰胺多一个 CH_2 基团，其使该极性基团更灵活，同时减少了它与主链的相互作用。酪氨酸具有很强极性的羟基基团（酚羟基），后者可在高 pH 值下解离。酪氨酸在行为上与带电荷侧链有些相似，其氢键非常强。

在蛋白质分子的表面和内部发现中性极性酸。作为内部残基，其通常互相或者与其它的多肽骨架形成氢键。半胱氨酸能形成二硫键。

组氨酸(His)带有 pK 值为 6.0 的芳香杂环侧链。在生理 pH 范围内，它的咪唑环可以不带有电荷或者从溶液中获取氢离子后带有电荷。因为这两种状态都很容易实现，所以组氨酸非常适于用于催化化学反应。已经在大多数酶例如丝氨酸蛋白酶的活性中心发现含有组氨酸。

天冬氨酸和谷氨酸在生理 pH 下带有负电荷。由于它们的侧链短，天冬氨酸的羧基对主链而言具有很大的刚性。这可能是在很多催化位点中的羧基基团是由天冬氨酸而不是谷氨酸提供的原因。带电氨基酸通常发现于多肽的表面。

另外，在表面发现赖氨酸和精氨酸。它们具有呈现相似能量的多种旋转异构体的长而且柔性的侧链。在几种情况下，赖氨酸和精氨酸参与内部的盐桥的形成或者帮助催化。由于其暴露于多肽的表面，所以赖氨酸是一种更经常被可修饰其侧链或者在赖氨酸残基的羧基端切断肽链的酶所识别的残基。

虽然氨基酸侧链基团化学特性可以指导预定氨基酸残基的选择，

缺乏目的侧链基团化学特性也可作为排除将氨基酸用作预定氨基酸的标准。例如：空间上小以及化学中性的氨基酸诸如丙氨酸因为缺少所期望的化学特性而可被排除在浏览突变之外。

3. 合成多肽类似物文库

在一个实施方案中，通过合成编码该多肽的指定区的、且具有不超过 1 个编码预定氨基酸的密码子的各个寡核苷酸来生成用于筛选的多肽类似物文库。这可通过在寡核苷酸内每个密码子位置处掺入或者野生型多肽合成所需要的密码子或者预定氨基酸的密码子来完成。由此制备的寡核苷酸与用饱和突变、随机突变或者 walk-through 突变的不同之处在于：对于每个寡核苷酸，制造仅仅一个突变不是多个突变。

可以单个地产生这种寡核苷酸然后按照所期望的进行混合或者合并。当野生型序列的密码子与预定氨基酸的密码子相同时，不进行置换。

所以，指定区内的氨基酸位置的数量决定所制备的寡核苷酸的最大数量。例如：如果用预定氨基酸改变 5 个密码子位置，则合成 5 个多核苷酸加上 1 个代表野生型氨基酸序列的多核苷酸。可以同时改变 2 个或者更多个区域。

可以用已知的 DNA 合成方法容易地合成用于生成文库的寡核苷酸混合物。优选的方法包括使用固相 β -氰乙基亚磷酰胺化学法。参见美国专利号 4725677。为方便起见，可以使用含有核苷酸的特定试剂容器的用于自动进行 DNA 合成的仪器。还可以合成该多核苷酸使之含有限制性位点或者引物杂交位点，以便于将代表诸如指定区的多核苷酸导入或者组装到更大的基因环境 (gene context) 中。

该合成的多核苷酸可通过标准基因工程技术被插入到突变的该多肽的更大的基因环境中。例如：该多核苷酸可以制成含有限制性酶识别的侧翼识别位点。参见 Crea, R., 美国专利号 4888286。该识别位点设计成与天然存在的或被导入的、最接近编码该区域的 DNA 的基因中识别位点相对应。转化为双链形式后，通过标准技术将该多核苷酸连接至基因中。使用适当的载体（包括例如：噬菌体载体、质粒）作为

工具可将该基因导入到适于表达突变多肽的无细胞抽提物、噬菌体、原核细胞或者真核细胞中。

当将被突变的多肽的氨基酸序列或者 DNA 序列已知时，可以使用基因合成方法。例如：可以设计部分重叠多核苷酸，典型地，长约 20-60 核苷酸。然后将内部的多核苷酸磷酸化并与其互补的伙伴退火以形成带有可用于进一步退火的单链延伸部分的双链 DNA 分子。然后将退火对混合在一起并且连接形成全长双链分子（参见例如：图 8）。可以将适当的限制性位点设计到接近于合成基因的末端处用于克隆到适当的载体中。可以用那些限制性酶切割全长分子，然后连接到适当的载体中。也可以将适当的限制性位点掺入到合成基因的序列中，便于突变盒的导入。

作为合成代表全长双链基因的多核苷酸的替代方案，可以将 3' 末端部分重叠（即带有互补的 3' 末端）的多核苷酸组装成带有缺口的结构，然后用适当的聚合酶填充，制成全长双链基因。典型地，重叠的多核苷酸长 40~90 个核苷酸。然后连接延长后的多核苷酸。为了克隆可以将适当的限制性位点导入到末端和/或中部。用适当的限制性酶消化后，将该基因片段连接到适当的载体中。或者，基因片段可以平端连接至适当载体中。

在这些方法中，如果在基因组装之后可得到方便的限制性位点（天然存在的或者人工引入的），简并的多核苷酸可以通过克隆该盒到适当的载体中而随后引入。或者，可以在基因组装阶段掺入该简并多核苷酸。例如：当该基因的两条链是全部化学合成的情况下，可以生成重叠和互补的简并多核苷酸。互补的对可以彼此退火。

当用部分重叠的多核苷酸进行组装时，可以将一组简并核苷酸替代该多核苷酸之一直接掺入。在延伸反应过程中通过用聚合酶进行酶促延伸由另外的链从部分互补的多核苷酸合成适当的互补链。当突变基因的 1 个以上的区域或者指定区时，在合成阶段掺入简并多核苷酸还可以简化克隆。

在另一方法中，目的基因存在于单链的质粒上。例如：可以将基

因克隆到噬菌体载体或者带有丝状噬菌体复制起点的载体中，后者允许应用辅助噬菌体增殖单链分子。可以用代表目的突变的一组简并引物与该单链模板退火，然后延长和连接，以此将各个类似物链掺入到可被导入到适当宿主中的一群分子中（Sayers, J. R. 等人, *Nucleic Acids Res.* 16: 791-802 (1988)）。当选择多个区域进行突变时，该方法可以避免多个克隆步骤。

也可以使用聚合酶链反应（PCR）方法将多核苷酸掺入到基因中。例如：可以将多核苷酸本身作为延伸引物。在该方法中，编码对应于指定区（或者其一部分）的突变盒的多核苷酸彼此互补（至少部分互补），可以用聚合酶，例如使用 PCR 扩增，将其延伸形成大基因盒。

文库的大小将根据所突变的区域的长度和数目及一个区域内突变的氨基酸的数量而不同。优选将该文库设计成含有少于 10^{15} 、 10^{14} 、 10^{13} 、 10^{12} 、 10^{11} 、 10^{10} 、 10^9 、 10^8 、 10^7 ，更优选 10^6 个多肽类似物或者更少。

上述内容主要着眼于通过改变编码相应多肽的多核苷酸而突变多肽和多肽文库。但是，可以理解本发明还包括通过使用蛋白质化学法直接合成目的多肽类似物来突变多肽的方法。实施该方法时，除了不使用多核苷酸中间体之外，最终得到的多肽仍具有本发明的特点。

对于上述的文库，无论是多核苷酸和/或对应的多肽的形式，应当理解，该文库也可以附着于固体支持物上，如使用本领域已知技术的微芯片，优选阵列。

4. 表达和筛选系统

可以对由上述任一技术或者其它适当的技术所生成的多核苷酸文库进行表达和筛选，以鉴定具有期望的结构和/或者活性的多肽类似物。可使用本领域已知的任何适当的表达展示系统进行该多肽类似物的表达，所述系统包括但是不局限于无细胞抽提物展示系统（例如：核糖体展示系统和阵列式（例如：微阵列或者大阵列式）展示系统）、细菌展示系统、噬菌体展示系统、原核细胞和/或真核细胞（例如：酵母展示系统）。

在一个实施方案中，可以将该多核苷酸人为改造用作为模板，其

可在无细胞的抽提物中表达。可以使用例如美国专利号 5324637、5492817、5665563 中介绍的载体和抽提物，许多种是可以商购的。可以使用将多核苷酸（基因型）与多肽（表型）相连的核糖体展示和其它无细胞技术，例如：Profusion™（参见美国专利号 6348315、6261804、6258558 和 6214553）。

或者，可以将本发明的多核苷酸在诸如 Pluckthum 和 Skerra（Pluckthum, A. and Skerra, A. Meth. Enzymol. 178: 476-515 (1989); Skerra, A. et al., Biotechnology 9: 273-278 (1991)）所介绍的等适当的大肠杆菌表达系统中进行表达。如 M. Better and A. Horwitz. Meth. Enzymol. 178: 476 (1989) 所述，可将突变蛋白质分泌表达达到培养基中和/或在细菌的细胞质中表达。在一个实施方案中，编码 VH 和 VL 的单结构域各自附在编码信号序列的序列的 3' 末端，所述信号序列诸如：ompA、phoA 或者 pelB 信号序列（Lei, S. P. et al., J. Bacteriol. 169: 4379 (1987)）。将这些基因融合物组装到双顺反子构建体中，使之能够由单个载体进行表达，而且分泌到大肠杆菌的周质间隙中，在周质间隙中蛋白重折叠并且能以活性形式回收（Skerra, A. et al., Biotechnology 9: 273-278 (1991)）。例如：可以同时表达抗体重链基因和抗体轻链基因来生成抗体或者抗体片段。

在另一实施方案中，可以在真核细胞诸如酵母中表达该多核苷酸，例如：如在例如美国专利号 6423538、6331391 以及 6300065 中介绍的酵母展示。在该方法中，文库的多肽类似物融合到在酵母的表面表达和展示的多肽上。可以使用其它真核细胞诸如哺乳动物细胞表达本发明的多肽，例如骨髓瘤细胞、杂交瘤细胞或者中国仓鼠卵巢（CHO）细胞。典型地，当在哺乳动物细胞中表达时，该多肽类似物设计成表达达到培养基中或者表达达到这类细胞的表面上。抗体或者抗体片段可以例如作为完整抗体分子或者作为单独的 VH 与 VL 片段，Fab 片段、单结构域或者作为单链（sFv）来产生（参见 Huston, J. S. et al., Proc. Natl. Acad. Sci. USA 85: 5879-5883 (1988)）。

可以使用任何适当的方式对所表达的多肽类似物（或者通过直接合成生产的多肽）进行筛选。例如：可以用标准的免疫分析和/或亲和层析对结合活性进行评价，以及可以用适当的对底物转化的分析来评价催化活性。可以使用例如美国专利号 5798208 中介绍的标准血红蛋白分析对本发明的多肽类似物进行蛋白水解功能筛选。

5. 计算机模拟 (Computer Modeling) 辅助的浏览突变

也可以利用与要产生的多肽类似物有关的结构信息或者模拟信息进行本发明的浏览突变，这样可以增加生成具有所期望的改进功能的类似物的可能性。结构信息或者模拟信息也可以用于指导用于导入到指定区的预定氨基酸的选择。而且，用本发明的多肽类似物获得的实际结果可以以重复的形式指导将要制备和筛选的后续多肽的选择（或者排除）。相应地，可以使用结构或者模拟信息生成用于本发明的多肽类似物的起始子集 (subset)，由此进一步提高生成改进的多肽的效率。

在具体的实施方案中，用计算机模拟 (in silico modeling) 消除任何预知具有差的或者不期望的结构和/或者功能的多肽类似物的产生。通过该方法能够显著地降低所产生的多肽类似物的数量，进而在随后的分析中增加信噪比。在另一具体的实施方案中，用其它来自任何相关资源诸如基因和蛋白质序列和三维结构数据库和/或来自之前检测的类似物的结果等的模拟信息不断更新该计算机模拟，以致使计算机数据库 (in silico database) 在预测能力方面变得更为精准 (图 19)。

在再一的实施方案中，向计算机数据库提供以前经过检测的多肽类似物的分析结果并且在该分析标准的基础上，将类似物分类为应答者或者不应答者，例如：分为可以很好地结合的多肽类似物或者不能很好地结合的多肽类似物或者分为具有酶催化性/催化性的或者没有酶催化性/催化性的。以这种方式，本发明的指导突变 (guide through mutagenesis) 可以使一系列功能性应答与特定的结构信息对应起来并使用这类信息来指导待测试的未来多肽的生产。因此该方法特别适合

用于筛选具有特定功能的抗体或者抗体片段，所述的特定功能可列举结合亲和性（例如：特异性）、稳定性（例如：半寿期）、和/或效应子功能（例如：补体激活以及 ADCC）。所以如果已知（例如：通过计算机模拟获知的）一个区域中某些残基不参与期望的功能，则可以期望在该区域上进行非邻接的残基突变。可以考虑并模拟指定区之间的协调的结构和空间相互关系，所述的指定区例如：多肽的指定区内的功能性氨基酸残基，例如：已经导入的预定氨基酸。这种模拟的原则包括例如：氨基酸残基侧链基团化学、原子距离、结晶学数据等。所以可以智能性地最小化将要生成的多肽类似物的数量。

在优选的实施方案中，计算机辅助上述的 1 个或者多个步骤。也可以通过装置如计算机驱动装置部分地或者全部地实施该方法。所以，可以部分地或者全部地将实施该方法的指令赋予给适合于在电子装置中执行该指令的介质。总之，本发明的方法适于包含软件（例如：计算机可读的指令）和硬件（例如：计算机、机器人以及芯片）的高通量方法。

6. 对多个指定区的组合化学的研究

本发明的重要的优点在于允许通过同时突变多肽的几个不同的区域或者结构域以进行评价。这可以通过在每个区域内使用相同或者不同的预定氨基酸来进行，由此使得可以对在构象上的相关区域上的氨基酸置换进行评价，所述区域诸如多肽折叠时缔合形成功能位点（例如：抗体的结合位点或者酶的催化位点）的区域。进而提供了创造新的或者改进的功能位点的有效的方法。

例如，如图 14 所示，可以同时突变构成抗原结合位点（Fv 区）的独特方面的抗体的 6 个 CDR，或者可以分别地在 VH 或 VL 链中对这些 CDR 进行突变，来研究在该位点中选定的氨基酸之间的三维相互关系。在一个实施方案中，使用 look-through 突变系统性地对 3 个或者更多指定区的组合化学进行研究，并且优选 6 个指定区，例如：抗体重链和轻链可变区的 6 个 CDR。为了在一个 CDR 上执行 look-through 突变，通常改变 3、4、5、6、7、8、9、10、11、12、13、14、15、16、

17、18、19、20、21、22、23、24、25、或者更多氨基酸位置。

所以本发明为设计很多不同类型的新型的改进多肽开创了新的途径。可以将该方法用于改进蛋白质的现有结构或功能，例如：可以向抗体或者抗体片段导入结合位点，或者可以改进针对预先存在的抗原的亲合性，或者改进效应子功能和/或稳定性。或者，可以向酶的催化域中导入另外的“具有催化重要性”的氨基酸，使之修饰或者提高针对底物的催化活性。或者，可以向多肽中导入全新的结构、特异性或者活性。也可以从头合成酶活性。可以通过使用本发明的方法仅仅突变相关区域的方式在现有蛋白的天然或者共有“支架”（consensus “scaffold”）上构建新的结构。

7. 为了制造新的或者改进的抗体而进行的 look-through 突变

本发明的方法对修饰抗体分子特别有用。此处所用的抗体分子或者抗体指的是抗体或者抗体的一部分，如：全长抗体、Fv 分子或者其它的抗体片段、单链或者它们的片段（例如：Fv 的单链）、单链抗体和嵌合抗体。可以向抗体的可变区和/或构架（恒定）区导入改变。对可变区的修饰可以使抗体具有更好的抗原结合特性，以及如果希望的话，催化特性。对构架区的修饰也能使其产生理化性质，如：可溶性或者稳定性（例如：半衰期）的改善，这些理化性质对例如商业生产、生物利用度、效应子功能（例如：补体激活和/或者 ADCC）以及对抗原的结合亲和性（例如：特异性）是非常有用的。典型地，该突变会以抗体分子的 Fv 区为靶，所述的 Fv 区为由一个来自重链（VH）和一个来自轻链（VL）的两个链可变区构成的结构，该结构负责抗原结合活性。一旦鉴定得到期望的抗原结合特性，可以将可变区人工改造成适当的抗体类型，如：IgG、IgM、IgA、IgD 或者 IgE。

8. 用于制备/改进具有催化性/酶催化性多肽的 look-through 突变

本发明的方法也特别适用于设计催化性蛋白质，特别是催化性抗体。目前，可以使用改进的标准体细胞融合技术制备催化性抗体。在该方法中，用类似于目的底物的过渡态的抗原免疫动物，以诱导能结

合该过渡态并催化该反应的抗体的产生。从动物中收获抗体产生细胞并且与永生化细胞融合制备杂交细胞。然后对能够分泌催化该反应的抗体的细胞进行筛选。该方法取决于是否能够得到底物过渡态的类似物。由于在大多数情况下很可能难以鉴定或者合成这种类似物，所以该方法可能受到限制。

本发明的方法提供了不同的方法，该方法不需要过渡态类似物。通过本发明的方法，可以向免疫球蛋白的结合位点（Fv 区域）导入适当的氨基酸使该抗体具有催化性。抗原结合位点（Fv）是由 6 个高变（CDR）环构成的，其中 3 个源自免疫球蛋白重链（H），3 个源自轻链（L），其在每个亚基内连接 β 链。CDR 环的氨基酸残基几乎全部贡献于每个特定单克隆抗体的结合特性。例如：可以在对底物分子具有已知亲和性的抗体 Fv 区域的高变片段上制造出模拟丝氨酸蛋白酶的催化三联体（catalytic triads）（由丝氨酸、组氨酸和天冬氨酸残基组成），并且对其针对该底物的蛋白水解活性进行筛选。

特别是，本发明的方法可以用于产生很多不同的酶或者催化性抗体，包括：氧化还原酶、转移酶、水解酶、裂合酶、异构酶以及连接酶。在这些类酶中，改进的蛋白酶、糖酶、脂肪酶、双加氧酶以及过氧化物酶的生产特别重要。可通过本发明的方法制备的这些酶以及其它酶具有重要的用于酶催化转化的商业应用性，如：在卫生保健、化妆品、食品、酿造、清洁剂、环境（例如废水处理）、农业、制革、纺织和其它化学工艺中。这些应用包括但不限于诊断和治疗应用、转化脂肪、糖和蛋白质、降解有机污染物以及化学品合成。例如：可以工程改造得到具有纤维蛋白溶解活性或者具有抗病毒感染所必需的结构如：病毒外壳蛋白的活性的治疗上有效的蛋白酶。这些蛋白酶可以用作抗血栓制剂或者针对病毒的抗病毒制剂，所述的病毒例如：HIV、鼻病毒、流感病毒或者肝炎病毒。对于加氧酶（例如：双加氧酶），即需要辅因子用于氧化芳香环和其它双键的一类酶的情况，工业应用于生物制浆工艺、将生物质转化成燃料或者其它化学物质、转化废水污染物、煤炭的生物加工以及危险性有机化合物的解毒是新蛋白的可

能用途。

通过下面的实施例进一步阐述本发明，不过本发明不局限于此。

实施例

除非另外说明，全部实施例中使用了下述的材料和方法。

材料和方法

一般地，除非另外说明，本发明在实施中采用均属于本领域技术人员技能范围内的并在文献中有说明的常规化学技术、分子生物学技术、重组 DNA 技术、PCR 技术、免疫学技术（特别是例如：抗体技术）、表达体系（例如：无细胞表达、噬菌体展示、核糖体展示以及 Profusion™）以及任何必需的细胞培养技术。参见文章例如：Sambrook, Fritsch and Maniatis, *Molecular Cloning* : Cold Spring Harbor Laboratory Press (1989); *DNA Cloning*, Vols.1 and 2, (D. N. Glover, Ed. 1985); *Oligonucleotide Synthesis* (M. J. Gait, Ed. 1984); *PCR Handbook Current Protocols in Nucleic Acid Chemistry*, Beaucage, Ed. John Wiley & Sons (1999) (Editor); *Oxford Handbook of Nucleic Acid Structure*, Neidle, Ed. , Oxford Univ Press (1999); *PCR Protocols : A Guide to Methods and Applications*, Innis 等, Academic Press (1990); *PCR Essential Techniques : Essential Techniques*, Burke, Ed., John Wiley & Son Ltd (1996); *The PCR Technique : RT-PCR*, Siebert, Ed. , Eaton Pub. Co. (1998); *Antibody Engineering Protocols (Methods in Molecular Biology)*, 510, Paul, S., Humana Pr (1996); *Antibody Engineering: A Practical Approach (Practical Approach Series, 169)*, McCafferty, Ed. , Irl Pr (1996); *Antibodies : A Laboratory Manual*, Harlow 等 , C. S. H. L. Press, Pub. (1999); *Current Protocols in Molecular Biology*, eds. Ausubel 等 , John Wiley & Sons (1992); *Large-Scale Mammalian Cell Culture Technology*, Lubiniecki, A. , Ed. , Marcel Dekker, Pub., (1990). Phage

Display : A Laboratory Manual, C. Barbas (Ed.), CSHL Press, (2001); Antibody Phage Display, P O'Brien (Ed.), Humana Press (2001); Border 等, Yeast surface display for screening combinatorial polypeptide libraries, Nature Biotechnology, 15 (6): 553-7 (1997); Border 等, Yeast surface display for directed evolution of protein expression, affinity, and stability, Methods Enzymol., 328: 430-44 (2000); 由 Pluckthun 等在 U. S. Patent No. 6, 348, 315 中描述的核糖体展示, 和由 Szostak 等在 U. S. Patent Nos. 6, 258, 558; 6, 261, 804; 和 6, 214, 553 中描述的 Profusion™。

实施例 1

look-through 突变抗原结合分子的 3 个指定区

在该实施例中, 介绍了 look-through 突变抗体的 3 个 CDR 区来改进对底物的结合和蛋白水解作用。

具体地, 对单克隆抗体的 3 个互补决定区 (CDR) 进行了 “look-through” 突变。选作 look-through 突变的指定区的为重链可变区 (VH) 的 CDR1、CDR2 和 CDR3。在该实施方案中, 选作预定氨基酸的是丝氨酸蛋白酶的催化三联体 (天冬氨酸。组氨酸以及丝氨酸) 的三个残基。为 VH 的 CDR1 选择天冬氨酸, 为 VL 的 CDR2 选择组氨酸, 为 VH 的 CDR3 选择丝氨酸。选择这三个预定氨基酸使得可以使用常规蛋白酶试验检测何时这三个残基正确地定位于能展现功能性活性, 即, 对待试底物的蛋白水解作用的位置。

MCPC603 是示例性抗体, 因为已经很好地表征了该抗体的结合区域, 所以认为该抗体是进行结合和催化研究的很好的模型。该 MCPC 603 VH 和 VL 区域的氨基酸序列和 DNA 序列均是可以公开获得的 (参见例如: Rudikoff, S. and Potter, M., Biochemistry 13: 4033(1974); Pluckthum, A. et al., Cold Spring Harbor Symp. Quant. Biol., Vol. LII: 105-112(1987))。如图 2 所示已经对 MCPC 603 的 CDR 区

进行了鉴定。在重链上，CDR1 跨第 31 - 35 位氨基酸残基，CDR2 跨第 50 - 69 位，CDR3 跨第 101 - 111 位。在轻链上，CDR1 的氨基酸残基位于第 24 - 40 位处，CDR2 跨第 55 - 62 位氨基酸，CDR3 跨第 95 - 103 位氨基酸。

对在 MCPC 603 的 CDR 区上用于 look-through 突变的寡核苷酸是这样设计的：该设计导致结果得到具有如图 3-5 所示的 CDR1、CDR2 和 CDR3 的氨基酸序列的多肽类似物。可以理解，为了方便插入到如图 7 所示的靶构建体中，所合成的寡核苷酸可以比将要改变的 CDR 区大。为了便于后续的表达和筛选步骤，选用了单链抗体形式。如图 8 所示，可以通过酶催化技术将寡核苷酸转化成双链（参见，例如：Oliphant, A. R. et al., 1986, 如前所述）然后连接到限制性酶切割后的质粒中。该限制性位点可以是天然存在的位点或者是人工改造的限制性位点。

可以在此处所述的任何方便的表达系统中表达编码相应多肽类似物的多核苷酸，然后根据例如：美国专利号 5798208 中介绍的丝氨酸蛋白酶分析法进行筛选（也可以参见图 16-17）。简而言之，将表达的多肽类似物接触待测底物，并且检测对待测底物的蛋白质水解。由底物清除的区揭示蛋白质水解的量，该量指示多肽类似物具有期望的功能性活性。

实施例 2

look-through 突变（浏览突变）抗原结合分子的 6 个指定区

在该实施例中，介绍了浏览突变抗体的所有 6 个 CDR 区来改进对底物的结合以及蛋白水解特性。

具体地，对上述的模型抗体（MCPC 603）的所有 6 个高变区或者互补决定区（CDR）进行了“浏览”突变。在该实施例中，用不同的氨基酸对给定的区域或者结构域进行了 2 到 3 次“浏览”突变。例如：如图 10 所示，使天冬氨酸、丝氨酸和组氨酸相继地步行通过（walked-through）重链和轻链。

当已知或者可以推论出位于一个区域中的某些残基不参与期望的功能时，可以在该区域进行非连续性残基突变。另外，可以将类似物

的数量最小化。在选择预定氨基酸和将要改变的具体位置时还需要考虑这些残基必需能够相互形成氢键。这种考虑可以对所生成的变体加上邻近度限制 (proximity constraint)。这样, 仅仅在该 CDR 的某些位置可允许催化三联体的氨基酸发生适当相互作用。由此可以用分子模拟或者其它的结构信息富集功能变体。

在本例中, 用已知的结构信息鉴定所述区域中位于可以足够接近进而允许天冬氨酸、组氨酸以及丝氨酸之间形成氢键的残基以及将要突变的残基的范围。Roberts 等人鉴定了 CDR 区的各部分间紧密接触的区域 (Roberts, V.A. et al., Proc. Natl. Acad. Sci. USA 87: 6654-6658 (1990))。将该信息与来自 MCPC 603 X 射线结构的数据一起用于选择用作突变靶的 CDR 中有潜力的紧密接触的区域。这种由结构/模拟信息指导的浏览突变可以称之为“指导”突变 (“guide-through” mutagenesis)。

如图 10 所示, 以每个 CDR 区接受 1 个预定氨基酸进行浏览突变, 结果产生 8×10^5 个多肽类似物, 如果对所有的 20 种氨基酸进行研究将产生 5×10^{13} 个多肽类似物。

可以在此处所述的任何方便的表达体系中表达多肽, 并且可以使用例如: 美国专利号 5798208 中介绍的丝氨酸蛋白酶进行筛选。简而言之, 将表达的多肽类似物接触待测底物, 并且检测对待测底物的蛋白质水解。底物清除的区域揭示蛋白质水解的量, 该量指示多肽类似物具有期望的功能性活性。

实施例 3

浏览突变抗 TNF 结合分子以改进其功能

在该实施例中, 介绍了浏览突变抗 TNF 抗体以改进其结合性。

具体地, 对 2 个不同的抗 TNF 抗体的所有 6 个高变区或者互补决定区 (CDR) 进行了“浏览”突变。通常将抗 TNF 抗体用于治疗 TNF (肿瘤坏死因子) 配体水平异常的患者的免疫疾病。存在两种商业可获得的抗 TNF 抗体。为了便于进行浏览突变和随后的筛选, 使用多聚 Gly-Ser 接头使这些抗体的重链和轻链的可变区 (参见: SEQ ID NOs:

2-4) 转变成单链形式(参见图 15)。如图 15 所示, 以黑线条的存在标识选出的用预定氨基酸浏览突变的指定区。这些指定区对应于单链抗体的 CDR。

如此处所述, 合成代表 6 个 CDR 区和足够的侧翼区的多核苷酸以允许将该多核苷酸组装成如图 15 所示的完整的单链序列。为各个 CDR 区选出预定氨基酸残基, 然后独立地相继导入到 6 个 CDR 区的每一个氨基酸位置。使用核糖体展示进一步人工改造该多核苷酸使之成为能支持转录成相应的 RNA 转录物的模板, 所述的 RNA 转录物可以随后被翻译成多肽。

使用无细胞转录和翻译提取物(extracts)表达编码相应多肽类似物的多核苷酸。使 RNA 转录物共价连接可以检测的部分, 如: 荧光部分。或者, 可以将目的序列以符合读框的形式融合到荧光部分, 如: 绿色荧光蛋白(GFP)上, 以便于表达的检测以及结合者相对不结合者的标准化。优选, 编码多肽类似物的多核苷酸至少部分按阵列排列, 例如: 表达在具有多个多肽类似物但可以容易地去卷积(deconvoluted)的孔中。因此, 每个孔含有一个多肽类物子集, 这些类似物使用核糖体展示现在与连接有荧光部分的对应转录物相连。用诸如 TNF 的靶配体探测该孔, 然后与野生型多肽相比较, 对多肽类似物的结合性以及结合亲和性进行分析。

如果多肽类似物的结合优于野生型多肽, 之后可以将之再次工程改造成全长 IgG 抗体形式, 使用标准的技术针对改进的结合性与类似的商品化抗体进行平行检测。

实施例 4

浏览突变抗肉毒杆菌神经毒素血清型 B (BoNT/B) 的抗体和抗肉毒杆菌神经毒素血清型 A (BoNT/A) 的抗体以改进其功能

在该实施例中, 使用浏览突变产生改进的抗肉毒杆菌神经毒素血清型 B (BoNT/B) 的抗体和抗肉毒杆菌神经毒素血清型 A (BoNT/A) 的抗体以改进其功能。LTM 方法的基础是以 20 种氨基酸的子集(LTM 集)为基础在整个结合口袋中每个 CDR 构建单突变, 所述子集研究大小、

电荷、疏水性和氢键键合特性。在 LTM 集中选择氨基酸的标准讨论如下。

1. 抗体纯化、基因测序以及单链 (scFv) 设计

可以根据 Emanuele et al. (1996) *Journal of Immunological Methods* 193:189-197 中所述的内容获得具有对 BoNT/B 的结合亲和性的鼠源抗体片段 (Fab)。可以使用抗体 *BotFab5* (SEQ ID NOs: 10 和 11, 分别为轻链和重链多肽)、*BotFab20* (SEQ ID NOs: 12 和 13, 分别为轻链和重链多肽) 或者 *BotFab22* (SEQ ID NOs: 14 和 15, 分别为轻链和重链多肽)。在美国专利号 5932449 中介绍了前述序列, 将其内容引入此处作为参考。可以从 Metabionics, Inc., WI 得到该鼠源 BoNT/B 抗原 (全长毒素、轻链和/或重链)。

在 Pless et al (2001) *Infect. Immun.* 69: 570 中介绍的抗 BoNT/A 抗体可以用于使亲和性至少增加 1 个数量级的目的。可以从 Metabionics, Inc., WI 获得 BoNT/A 重链结合结构域 (BoNT/A Hc) 抗原。

使用标准分子生物学技术克隆并且测定抗体的 V_L 和 V_H 片段的序列。使用聚合酶链反应 (PCR) 扩增这些分子的可变区, 然后与 poly-Gly-Ser 接头 (典型地为 SGGGSGGGGSGGGGS (SEQ ID NO: 7)) 连接生成单链抗体 (scFv)。也可将多聚 His 标签 (HHHHHH (SEQ ID NO: 8)) 和 myc 标签 (EQKLISEEDL (SEQ ID NO: 9)) 添加到基因的 C 末端以便于纯化和检测。使用任何本领域的已知技术 (例如: 基于酵母、细菌或者噬菌体的技术) 展示这些分子, 然后检查其结合 BoNT/B 或者 BoNT/A 的能力。

完整抗体的单价 scFv 形式是用于进行突变研究的很好的形式, 已知除了多价分子的亲和力作用外, 其可以再现完整分子的结合机制。

2. 使用浏览突变 (LTM) 和基因设计改进抗体

为了进行浏览突变 (LTM), 选择下述 9 个氨基酸和其代表性功能特性: 丙氨酸和亮氨酸 (脂肪族)、丝氨酸 (羟基基团)、天冬氨酸 (酸性) 和谷氨酰胺 (酰胺)、赖氨酸和组氨酸 (碱性)、酪氨酸 (芳

香族)、脯氨酸(疏水性)。这些氨基酸在大小、电荷、疏水性以及氢键成键能力方面表现出足够的化学多样性,可以为改进抗体特性所需要的化学功能性方面提供有意义的初始信息。该选择还以这些氨基酸在抗体 CDR 中的出现频率为依据。例如:假定在酪氨酸和苯丙氨酸之间进行选择以代表具有芳香侧链的氨基酸,将选择前者,因为其在抗体 CDR 中明显更高的优势以及其氢键的形成能力。最初使用 LTM 鉴定对最终抗体中期望的结合、中和和/或任何其它特性有益的特定的氨基酸和化学特性。不过 LTM 不局限于这 9 个氨基酸或者总共 9 种氨基酸。可以以氨基酸的任意组合实施 LTM 分析,而且 LTM 子集可以高达 18 种氨基酸(一种简化的饱和突变)。

2.1 LTM 寡核苷酸合成

在最初的 LTM 分析中,其目的是调查在每个 CDR 中每个氨基酸的侧链对整体结合亲和性的贡献。为了有效地生成有意义的多样性,LTM 子集中的每个氨基酸按照每个 CDR 仅有 1 个置换的方式靶向 CDR 序列中的每一单个位置。这样,每个单独的寡核苷酸仅编码单一一个 CDR 突变。例如:为了对假设的 11 个氨基酸的 V_H CDR3 结构域进行 LTM 组氨酸突变,将合成 11 种寡核苷酸以编码所有 11 种可能的单组氨酸突变(参见图 20)。该分析在于检查在 CDR 的各位点具有大体积酰胺的作用。因此为了生成 V_H CDR3 的 LTM 文库,仅合成 99 种寡核苷酸进行该 LTM 分析(9 种 LTM 氨基酸 $\times$ 11 个 V_H CDR3 位置)。使用公开可获得的软件对寡核苷酸序列进行无意出现的终止密码子、野生型副本、低效密码子使用、发夹、环以及其它二级结构的检测。

2.2 为了组合有益的 LTM 突变合成简并寡核苷酸

用系统性 LTM 置换 CDR 上的各氨基酸揭示了各个位点上优选的化学官能度(chemical functionalities)。为了合并来自这些 LTM 选择的所有突变以及探索它们之间可能的叠加性和能量上的协同,合成了编码这些突变的简并寡核苷酸以及野生型序列,用具有 1.6×10^4 个变体的此寡聚核苷酸简并集合(degenerate pool)随后生成第二代文库,用该文库研究这些替代的叠加性质。

2.3 计算机辅助的寡聚核苷酸设计、文库和结果数据库

软件与定制的自动 DNA 合成仪相结合能够迅速合成寡核苷酸。第一步涉及决定将哪些靶氨基酸并入到 CDR 中。由软件确定导入这些靶氨基酸所需的密码子偏倚性（例如：酵母、细菌或者噬菌体的密码子偏向性，其取决于所选择的系统），并且消除可能在该设计过程中产生的野生型序列的副本。然后对潜在的终止密码子、发夹和环结构或者对其它问题性序列进行分析，随后在合成前加以校正。然后将该完整的 LTM 设计计划发送到 DNA 合成仪，由该仪器自动合成该寡核苷酸。以这种方式能够迅速合成用来创建文库所需的寡核苷酸。

可以在电子数据库中存储所有合成的 LTM 寡核苷酸序列的信息、scFv 文库 CDR 置换的详细情况以及对靶抗原的结合分析结果。该寡核苷酸数据档案允许设计策略的合理重复以及寡核苷酸试剂的重复使用。

3. 全面的 LTM 策略

LTM 最初用于鉴定对最终抗体中期望的结合、中和和/或任何其它特性有益处的特定的氨基酸和化学特性。LTM 也可以很快地鉴定在不显著损失亲和性（或者任何选定的物理特性）的情况下不耐受任何突变的区域。所以，LTM 分析不仅是研究抗体的所有 CDR 的每个位置上的化学需求的好方法，而且其还可以迅速地确定抗原结合所绝对必需的氨基酸。用 LTM 鉴定有益的突变后，可以使用组合突变策略，该策略首先插入所有这些完全不同的氨基酸突变以生成多重突变的 CDR。另外或者作为可替换的方案，可以用“walk through”突变（WTM）探测在 CDR 中同一氨基酸的多种突变的效果（例如参见美国专利号 5830650，5798208 所述）

4. LTM scFv 文库

上述的 LTM 技术用于创建在轻链或者重链的单个 CDR 上带有突变的寡聚核苷酸的集合。合成这些寡核苷酸使其包括部分的周边构架，以此便于在 PCR 中进行重叠和杂交。使用单重叠延伸 PCR（SOE-PCR）（如 Horton et al. (1989) Gene 77: 61-68 中所述）将这些寡核苷酸

集合用于产生所有可能的 V_L 和 V_H 链，在所述的 V_L 和 V_H 链中的单个、两个以及三个 CDR (CDR1、2 和 3 中的一个、两个以及三个的组合) 上有突变。SOE-PCR 是一种快速而且简单的用于组合 DNA 片段的方法，该方法不需要限制性位点、限制性内切酶或者 DNA 连接酶。在 SOE-PCR 中，首先使用设计成可使 PCR 产物在一端共有互补序列的引物通过 PCR 扩增基因上的 2 个区域。在 PCR 条件下，互补序列杂交、形成重叠。然后该互补序列作为引物，允许 DNA 聚合酶的延伸作用以产生重组分子。

例如：为了创建其中的 CDR-H1 和 CDR-H2 二者均被突变而 CDR3-H3 是野生型的 V_H 链的集合（其表示为“110”（1 表示突变的 CDR，0 表示野生型的 CDR）），该 CDR-H1 突变基因用作模板，进行 SOE-PCR 以连接 CDR-H2 寡核苷酸从而生成双突变的集合（图 21）。考虑到每个 CDR 可能或者是野生型或者是突变型，对于 V_L 和 V_H 链的集合（不包括野生型分子“000”）而言，各有 7 种可能的组合（如图 21 的箭头所示）。组合这 7 种 V_L 和 8 种 V_H 集合创造出 63 种 V_L - V_H 的非野生型组合（scFv）（图 21）。这 64 种 V_L - V_H 组合中的每一种（包括野生型序列）被称作整个 LTM scFv 文库总集（ensemble）的“子集”。针对每个选定用于置换的氨基酸，创建 scFv 文库总集。每个子集文库所提供的氨基酸序列的数量取决于 CDR 的长度、在 CDR 内的氨基酸序列以及 LTM 寡核苷酸设计策略。

5. 高通量文库筛选和改进的抗体的选择

可以获得多种方法用于表达和展示抗体。这些方法包括噬菌体、大肠杆菌和酵母。尽管这些方法均已经用于抗体改进，但酵母展示系统具有几个优点（Boder and Wittrup (1997) Nat. Biotechnol. 15: 553-557）。酵母可以容易地容纳高达 10^7 的文库大小，其中在每个细胞表面展示的每种抗体各 $10^3 \sim 10^5$ 个拷贝。使用流式细胞术和荧光激活细胞分选（FACS）或者磁珠可以很容易地进行酵母细胞的筛选和分离。酵母也可以迅速地选择和再生。与原核展示系统相比，酵母的真核分泌系统和糖化通路允许大得多的 scFv 分子子集正确折叠以及展

示到细胞的表面。已经用酵母展示与定向进化相结合的方式使 scFv 抗体片段对荧光素的 K_D 增加到 48fM, 比以前报道的任何单价配体 (Border et al. (2000)PNAS97:10701-10705) 的结合性增强了 2 个数量级。该展示系统使用 α -凝集素酵母粘着受体将蛋白质展示到细胞表面。本例中, 目的蛋白质抗 BoNT/B scFv LTM 文库或者抗 BoNT/A scFv LTM 文库均作为融合伴侣和 Aga2 蛋白质一起表达。这些融合蛋白质从细胞中分泌出来后以二硫键连接 Aga1 蛋白质, Aga1 蛋白质附着于酵母细胞的细胞壁上 (参见 Invitrogen, pYD1 Yeast Display product literature)。另外, 还可以包括羧基末端标签, 该标签可以用于监测表达水平和/或标准化结合亲和性测定结果。

链霉亲和素包被的磁珠 (Spherotech) 用于筛选和选择以高亲和性结合 BoNT/B 或者 BoNT/A 的抗体。该方法用具有高亲和力的抗体结合生物素化的抗原, 生物素化的抗原然后结合到链霉亲和素包被的珠子上, 以此对酵母克隆进行选择 (Yeung and Wittrup(2002)Biotechnol.Prog 18:212-220) 和 Feldhaus et al (2003)Nature Biotech. 21:163-170)。使用标准的操作流程(Pierce)生物素化 BoNT/B 或者 BoNT/A 多肽 (Metabionics), 然后用本领域已知的方法进行筛选。使用基于平衡和动力学的选择, 从这些文库中选出亲和性改善的抗体。用分析性 FACS (FACSscan) 监测每轮选择的功效。另外, 通过滴定抗原测定展示在酵母表面上的每个分子的相对结合亲和性。这可以迅速地鉴定出亲和力改善的分子。然后测定该 scFv 克隆的序列以鉴定有益的突变。

6. 产生可溶性抗体以进行 BLAcore 亲和力测定

将目的抗体亚克隆到可溶性表达系统 (巴斯德毕赤酵母 (*Pichia pastoris*) 和/或者大肠杆菌) 中, 然后产生可溶性蛋白质。有几种用于抗体的可溶性表达的商品化的载体和细胞系, 包括来自 Invitrogen 公司 (例如: 用于巴斯德毕赤酵母的 pPIC9) 和 Novagen 公司 (例如: 用于在大肠杆菌的周质中表达的 pET20b) 的那些。这些系统常规地用于生产可溶性单链抗体或者全长抗体。通常巴斯德毕赤

酵母表达系统 (Invitrogen) 能生产 1~5mg/L 的可溶的纯化的 scFv。当为单链抗体时, 使分子的 C 末端存在 His 标签便于蛋白质的纯化, 当为全长抗体时可以使用蛋白 A 或蛋白 G 柱纯化。生成可溶性单链和全长抗体以进行 BIAcore 亲和力动力学速率测定。此步骤是必需的, 因为位于酵母克隆细胞表面的高亲和力的 scFv 分子必须被验证为无细胞的可溶性分子的形式。

可以将以上述方式产生的可溶性单链和全长抗体用于进行 BIAcore 亲和力测定和如下所述的神经突突起分析或者小鼠致死分析。

7. 根据体内或者体外中和作用筛选选定的抗体

使用鸡的原代神经元的神经突突起作为 BoNT 中毒的指针, 并且由此作为定量毒素中和作用的手段, 利用基于细胞的分析对选定的抗体进行筛选。使用来自受精的鸡蛋的背根神经节外植块培养物进行初步的试验, 结果表明与对照相比, 来自用 BoNT/A 处理的外植块的轴突突出较少。或者, 可以使用如 Lomneth et al. (1990) Neuroscience letters 113: 211-216 中所述的鸡睫状神经节-虹膜肌神经肌肉接头分析。

小鼠致死分析 (MLA, 例如在 Schantz 和 Kautter (1978) J. Assoc. Off. Anal. Chem. 61: 96-99 中所述的) 是另外一个已知的所接受的检测对 BoNT 的中和的体内方法。该检测涉及将含有或者不含有抗体的近 0.5ml 的 BoNT/B 或者 BoNT/A 样品制剂腹膜内注射到 20~30g 的白色 ICR 品系小鼠中。经 1~4 天观察由呼吸衰竭造成的死亡。对中和进行定量时需要具有不同水平的小鼠 BoTN/B 或者 BoTN/A LD₅₀ 的系列稀释 MAb。可以将 MLA 用于确定通过体外筛选鉴定到的优化的抗体的中和能力。

还可以通过小鼠保护分析测定抗体的中和能力 (MPA, Goeschel et al. (1997) Exp. Neurol. 147: 96-102)。在 MPA 中, 从小鼠中将左膈神经与左半膈一起切取出来。然后在组织浴中连续地电刺激该膈神经。将纯化的抗体与 BoNT/B 或者 BoNT/A 温育, 然后加到该组织浴中。毒

素诱导的麻痹被定义为初始肌肉颤搐减低 50 %。

等同方案

本领域的技术人员将意识到或者仅使用常规实验即能够确定与此处所述的本发明的特定的实施方案等同的多种方案，这些等同方案也将包括在下面的权利要求中。

<110> R. Crea & Co., et al.

<120> Look-through 突变

<130> BRI-001PC

<150> 60/483,282

<151> 2003-06-27

<160> 13

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 121

<212> PRT

<213> 人工序列

<220>

<223> 合成的构建体

<400> 1

Gln	Val	Gln	Leu	Val	Gln	Ser	Gly	Ala	Glu	Val	Val	Lys	Pro	Gly	Ser
1			5						10					15	
Ser	Val	Lys	Val	Ser	Cys	Lys	Ala	Ser	Gly	Tyr	Thr	Phe	Thr	Asp	Tyr
			20					25					30		
Asn	Val	Asp	Trp	Val	Lys	Gln	Ala	Pro	Gly	Gln	Gly	Leu	Gln	Trp	Ile
		35					40					45			
Gly	Asn	Ile	Asn	Pro	Asn	Asn	Gly	Gly	Thr	Ile	Tyr	Asn	Gln	Lys	Phe
	50					55				60					
Lys	Gly	Lys	Gly	Thr	Leu	Thr	Val	Asp	Lys	Ser	Thr	Ser	Thr	Ala	Tyr
65					70				75					80	
Met	Glu	Leu	Ser	Ser	Leu	Thr	Ser	Glu	Asp	Thr	Ala	Val	Tyr	Tyr	Cys
			85					90					95		
Ala	Arg	Ser	Ala	Phe	Tyr	Asn	Asn	Tyr	Glu	Tyr	Phe	Asp	Val	Trp	Gly
			100					105					110		
Gln	Gly	Thr	Thr	Val	Thr	Val	Ser	Ser							
		115					120								

<210> 2

<211> 119

<212> PRT

<213> 人工序列

<220>

<223> 合成的构建体

<400> 2

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

```

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ser Lys Ser Ala
65                               70                               75                               80

Val Tyr Leu Gln Met Thr Asp Leu Arg Thr Glu Asp Thr Gly Val Tyr
                               85                               90                               95
Tyr Cys Ser Arg Asn Tyr Tyr Gly Ser Thr Tyr Asp Tyr Trp Gly Gln
                               100                              105                              110
Gly Thr Thr Leu Thr Val Ser
                               115

```

<210> 3
 <211> 114
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

```

<400> 3
Asp Ile Met Met Thr Gln Ser Pro Ser Thr Leu Ser Ala Ser Val Gly
1                               5                               10                               15
Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
                               20                               25                               30
Asn Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
                               35                               40                               45
Ala Pro Lys Leu Leu Ile Ser Trp Ala Ser Thr Arg Glu Ser Gly Val
                               50                               55                               60
Pro Ser Arg Phe Ile Gly Ser Gly Ser Gly Thr Glu Phe Thr Leu Thr
65                               70                               75                               80
Ile Ser Ser Leu Gln Pro Asp Asp Val Ala Thr Tyr Tyr Cys Gln Gln
                               85                               90                               95
Tyr Tyr Asp Tyr Pro Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
                               100                              105                              110
Lys Arg

```

<210> 4
 <211> 107
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

```

<400> 4
Asp Ile Leu Leu Thr Gln Ser Pro Ala Ile Leu Ser Val Ser Pro Gly
1                               5                               10                               15
Glu Arg Val Ser Phe Ser Cys Arg Ala Ser Gln Phe Val Gly Ser Ser
                               20                               25                               30
Ile His Trp Tyr Gln Gln Arg Thr Asn Gly Ser Pro Arg Leu Leu Ile
                               35                               40                               45
Lys Tyr Ala Ser Glu Ser Met Ser Gly Ile Pro Ser Arg Phe Ser Gly
                               50                               55                               60
Ser Gly Ser Gly Thr Asp Phe Thr Leu Ser Ile Asn Thr Val Glu Ser
65                               70                               75                               80
Glu Asp Ile Ala Asp Tyr Tyr Cys Gln Gln Ser His Ser Trp Pro Phe
                               85                               90                               95
Thr Phe Gly Ser Gly Thr Asn Leu Glu Val Lys
                               100                              105

```

<210> 5
 <211> 16
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

<400> 5
 Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15

<210> 6
 <211> 6
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

<400> 6
 His His His His His His
 1 5

<210> 7
 <211> 10
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

<400> 7
 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
 1 5 10

<210> 8
 <211> 236
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

<400> 8
 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1 5 10 15
 Ala Gln Pro Ala Met Ala Asp Ile Gln Met Thr Gln Ser Pro Ala Ser
 20 25 30
 Leu Ser Ala Ser Val Gly Glu Thr Val Thr Ile Thr Cys Arg Ala Ser
 35 40 45
 Gly Asn Ile His Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys
 50 55 60
 Ser Pro Gln Leu Leu Val Tyr Asn Ala Lys Thr Leu Ala Asp Gly Val
 65 70 75 80
 Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys
 85 90 95

```

Ile Asn Ser Leu Gln Pro Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His
      100                      105                      110
Phe Trp Ser Thr Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
      115                      120                      125
Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser
      130                      135                      140
Glu Gln Leu Thr Ser Gly Gly Ala Ser Val Val Cys Phe Leu Asn Asn
      145                      150                      155                      160
Phe Tyr Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu
      165                      170                      175
Arg Gln Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp
      180                      185                      190
Ser Thr Tyr Ser Met Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr
      195                      200                      205
Glu Arg His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr
      210                      215                      220
Ser Pro Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
      225                      230                      235

```

<210> 9
 <211> 254
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

```

<400> 9
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
  1      5      10      15
Ala Gln Pro Ala Met Ala Glu Val Gln Leu Gln Gln Ser Gly Ala Glu
      20      25      30
Leu Val Lys Pro Gly Ala Ser Val Lys Leu Ser Cys Thr Ala Ser Gly
      35      40      45
Phe Asn Ile Lys Asp Thr Phe Met His Trp Val Lys Gln Arg Pro Glu
      50      55      60
Gln Gly Leu Glu Trp Ile Gly Arg Ile Asp Pro Ala Asn Gly Asn Thr
      65      70      75      80
Glu Tyr Asp Pro Lys Phe Gln Gly Lys Ala Thr Ile Thr Ala Asp Thr
      85      90      95
Ser Ser Asn Thr Val Asn Leu Gln Leu Ser Ser Leu Thr Ser Glu Asp
      100     105     110
Thr Ala Val Tyr Tyr Cys Ala Ser Gly Gly Glu Leu Gly Phe Pro Tyr
      115     120     125
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala Ala Lys Thr Thr Pro
      130     135     140
Pro Ser Val Tyr Pro Leu Ala Pro Gly Ser Ala Ala Gln Thr Asn Ser
      145     150     155     160
Met Val Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu Pro Val
      165     170     175
Thr Val Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe
      180     185     190
Pro Ala Val Leu Gln Phe Asp Leu Tyr Thr Leu Ser Ser Ser Val Thr
      195     200     205
Val Pro Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala
      210     215     220
His Pro Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp
      225     230     235     240
Cys Thr Ser Gly Gly Gly Gly Ser His His His His His
      245     250

```

<210> 10
 <211> 236
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

<400> 10
 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1 5 10 15
 Ala Gln Pro Ala Met Ala Asp Ile Gln Met Thr Gln Ser Pro Ala Ser
 20 25 30
 Leu Ser Ala Ser Val Gly Glu Thr Val Thr Ile Thr Cys Arg Ala Ser
 35 40 45
 Gly Asn Ile His Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys
 50 55 60
 Ser Pro Gln Leu Leu Val Tyr Asn Ala Lys Thr Leu Ala Asp Gly Val
 65 70 75 80
 Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys
 85 90 95
 Ile Asn Ser Leu Gln Pro Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His
 100 105 110
 Phe Trp Ser Thr Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
 115 120 125
 Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser
 130 135 140
 Glu Gln Leu Thr Ser Gly Gly Ala Ser Val Val Cys Phe Leu Asn Asn
 145 150 155 160
 Phe Tyr Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu
 165 170 175
 Arg Gln Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp
 180 185 190
 Ser Thr Tyr Ser Met Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr
 195 200 205
 Glu Arg His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr
 210 215 220
 Ser Pro Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
 225 230 235

<210> 11
 <211> 254
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

<400> 11
 Met Lys Tyr Leu Leu Pro Thr Ala Ala Val Gly Leu Leu Leu Leu Ala
 1 5 10 15
 Ala Gln Pro Ala Met Ala Glu Val Gln Leu Gln Gln Ser Gly Ala Glu
 20 25 30
 Leu Val Lys Pro Gly Ala Ser Val Lys Leu Ser Cys Thr Ala Ser Gly
 35 40 45
 Phe Asn Ile Lys Asp Thr Phe Met His Trp Val Lys Gln Arg Pro Glu
 50 55 60

```

Gln Gly Leu Glu Trp Ile Gly Arg Ile Asp Pro Ala Asn Gly Asn Thr
65          70          75          80
Glu Tyr Asp Pro Lys Phe Gln Gly Lys Ala Thr Ile Thr Ala Asp Thr
85          90          95
Ser Ser Asn Thr Val Asn Leu Gln Leu Ser Ser Leu Thr Ser Glu Asp
100        105        110
Thr Ala Val Tyr Tyr Cys Ala Ser Gly Gly Glu Leu Gly Phe Pro Tyr
115        120        125
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala Ala Lys Thr Thr Pro
130        135        140
Pro Ser Val Tyr Pro Leu Ala Pro Gly Ser Ala Ala Gln Thr Asn Ser
145        150        155        160
Met Val Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu Pro Val
165        170        175
Thr Val Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe
180        185        190
Pro Ala Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Ser Ser Val Thr
195        200        205
Val Pro Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala
210        215        220
His Pro Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp
225        230        235        240
Cys Thr Ser Gly Gly Gly Gly Ser His His His His His His
245        250

```

<210> 12
 <211> 236
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成的构建体

```

<400> 12
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1      5      10      15
Ala Gln Pro Ala Met Ala Asp Ile Gln Met Thr Gln Ser Pro Ala Ser
20      25      30
Leu Ser Ala Ser Val Gly Glu Thr Val Thr Ile Thr Cys Arg Ala Ser
35      40      45
Gly Asn Ile His Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Gln Gly Lys
50      55      60
Ser Pro Gln Leu Leu Val Tyr Asn Ala Lys Thr Leu Ala Asp Gly Val
65      70      75      80
Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Gln Tyr Ser Leu Lys
85      90      95
Ile Asn Ser Leu Gln Pro Glu Asp Phe Gly Ser Tyr Tyr Cys Gln His
100     105     110
Phe Trp Ser Thr Pro Trp Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
115     120     125
Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser
130     135     140
Glu Gln Leu Thr Ser Gly Gly Ala Ser Val Val Cys Phe Leu Asn Asn
145     150     155     160
Phe Tyr Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu
165     170     175
Arg Gln Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp
180     185     190
Ser Thr Tyr Ser Met Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr
195     200     205

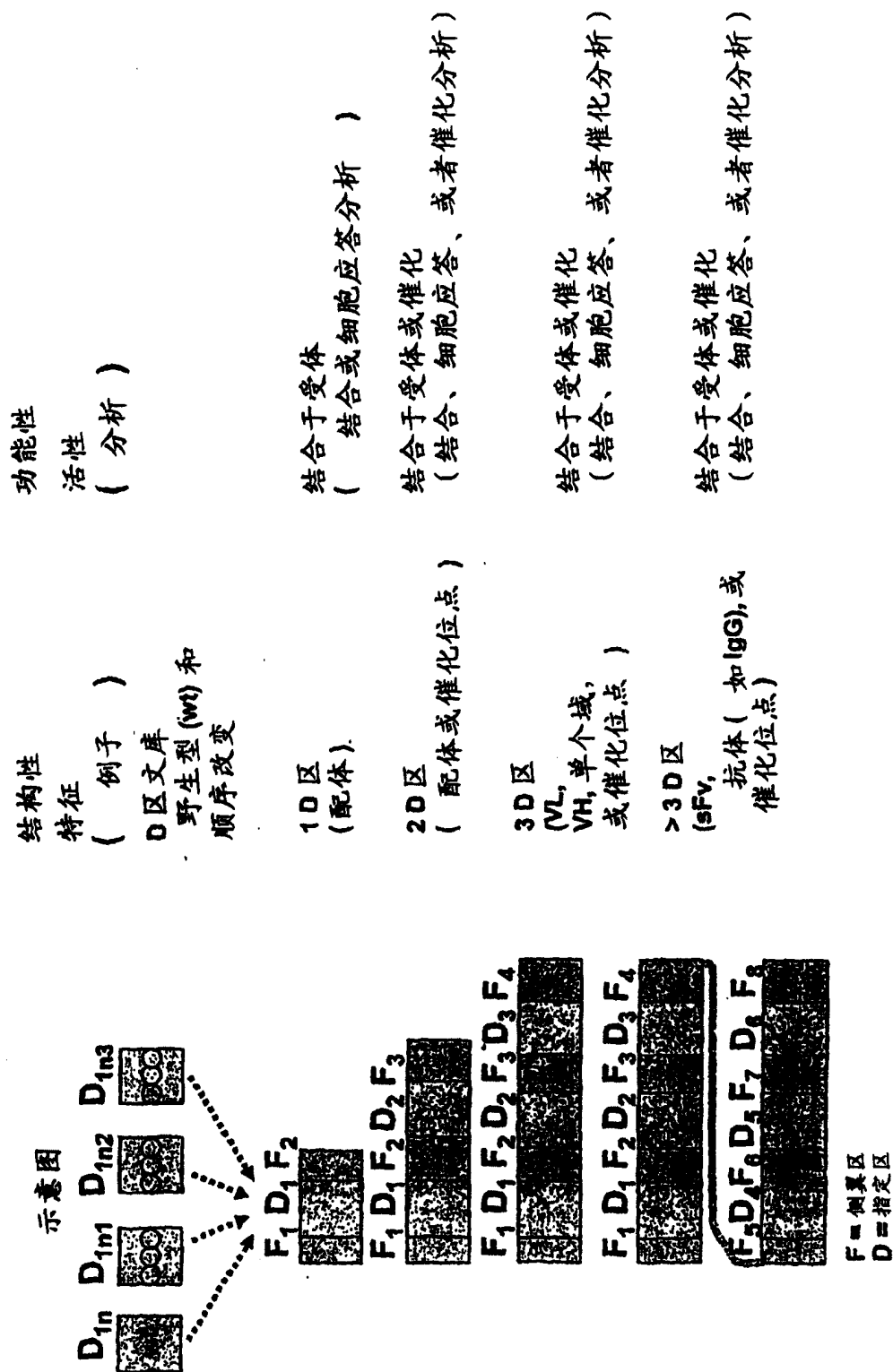
```

Glu Arg His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr
 210 215 220
 Ser Pro Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
 225 230 235

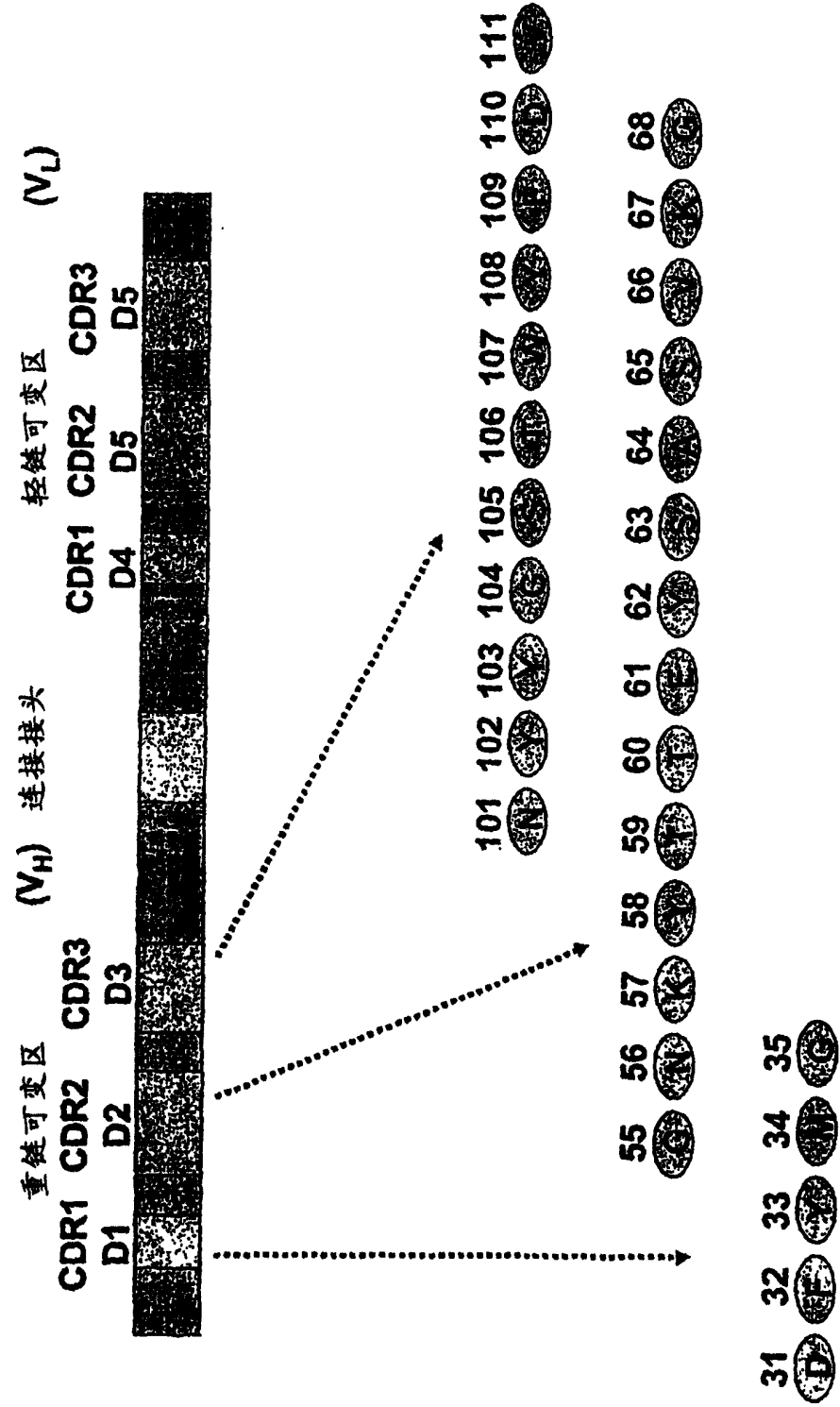
<210> 13
 <211> 254
 <212> PRT
 <213> 人工序列

<400> 13
 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1 5 10 15
 Ala Gln Pro Ala Met Ala Glu Val Gln Leu Gln Gln Ser Gly Ala Glu
 20 25 30
 Leu Val Lys Pro Gly Ala Ser Val Lys Leu Ser Cys Thr Ala Ser Gly
 35 40 45
 Phe Asn Ile Lys Asp Thr Phe Met His Trp Val Lys Gln Arg Pro Glu
 50 55 60
 Gln Gly Leu Glu Trp Ile Gly Arg Ile Asp Pro Ala Asn Gly Asn Thr
 65 70 75 80
 Glu Tyr Asp Pro Lys Phe Gln Gly Lys Ala Thr Ile Thr Ala Asp Thr
 85 90 95
 Ser Ser Asn Thr Val Asn Leu Gln Leu Ser Ser Leu Thr Ser Glu Asp
 100 105 110
 Thr Ala Val Tyr Tyr Cys Ala Ser Gly Gly Glu Leu Gly Phe Pro Tyr
 115 120 125
 Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ala Ala Lys Thr Thr Pro
 130 135 140
 Pro Ser Val Tyr Pro Leu Ala Pro Gly Ser Ala Ala Gln Thr Asn Ser
 145 150 155 160
 Met Val Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu Pro Val
 165 170 175
 Thr Val Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe
 180 185 190
 Pro Ala Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Ser Ser Val Thr
 195 200 205
 Val Pro Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala
 210 215 220
 His Pro Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp
 225 230 235 240
 Cys Thr Ser Gly Gly Gly Gly Ser His His His His His His
 245 250

示范的Look-Through突变 (LTM) 文库形式

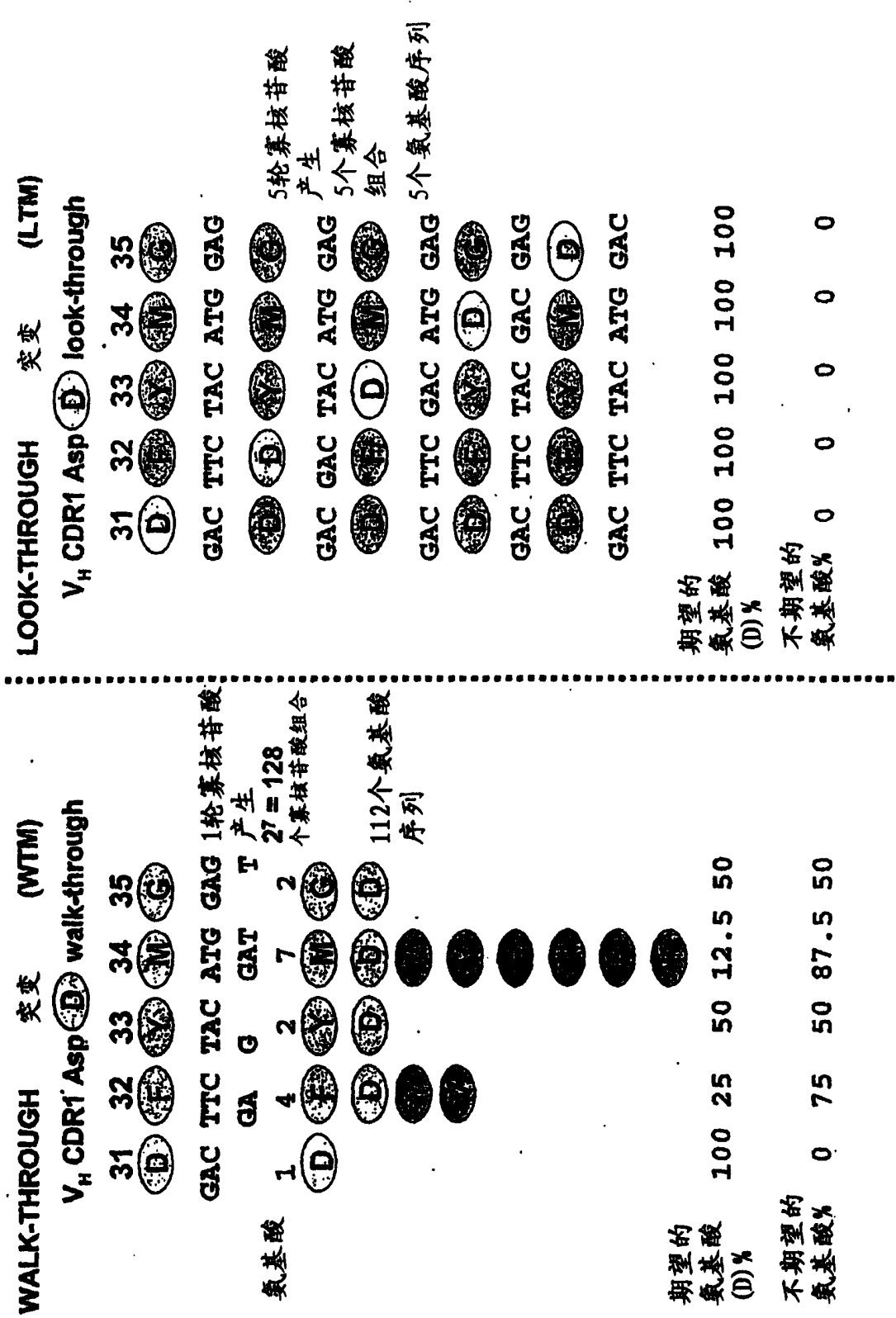


在合成基因 (sFv) 上鉴定 3 个指定区 (D1-D3) 用于 Look-Through 突变 (LTM)

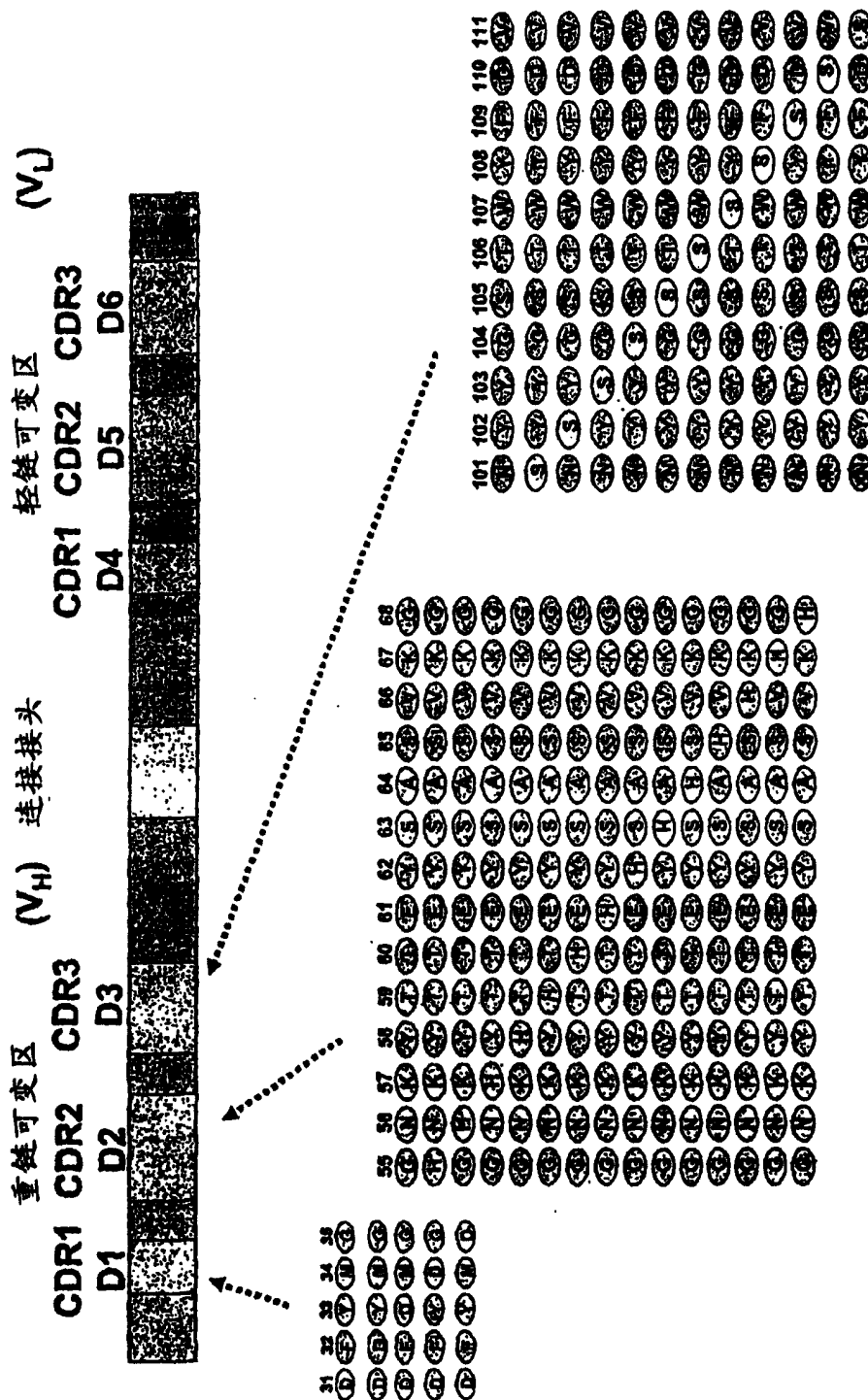


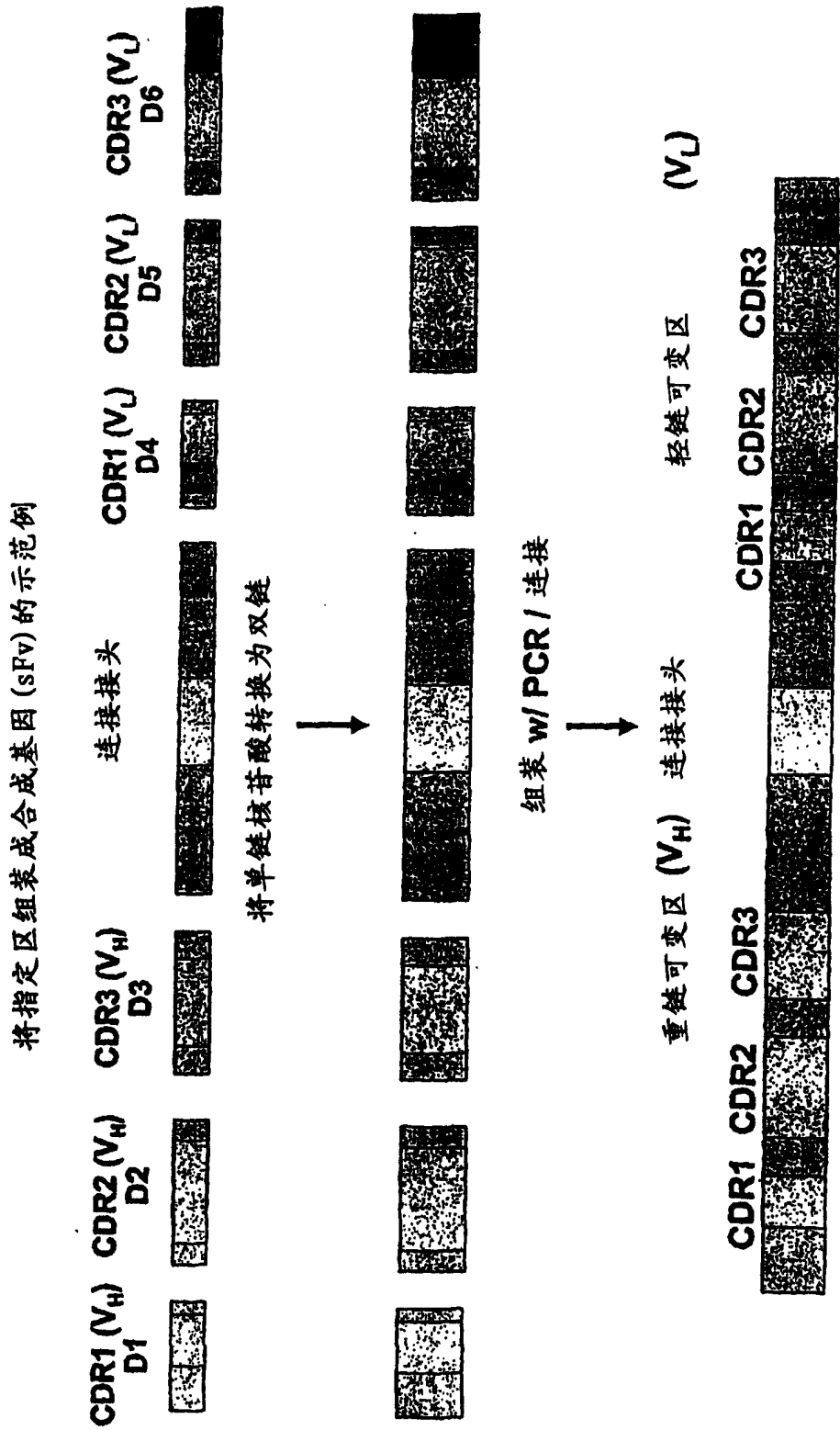
第一个指定区 (D1) 的顺序LTM
(V_HCDR1的第31-35位氨基酸)

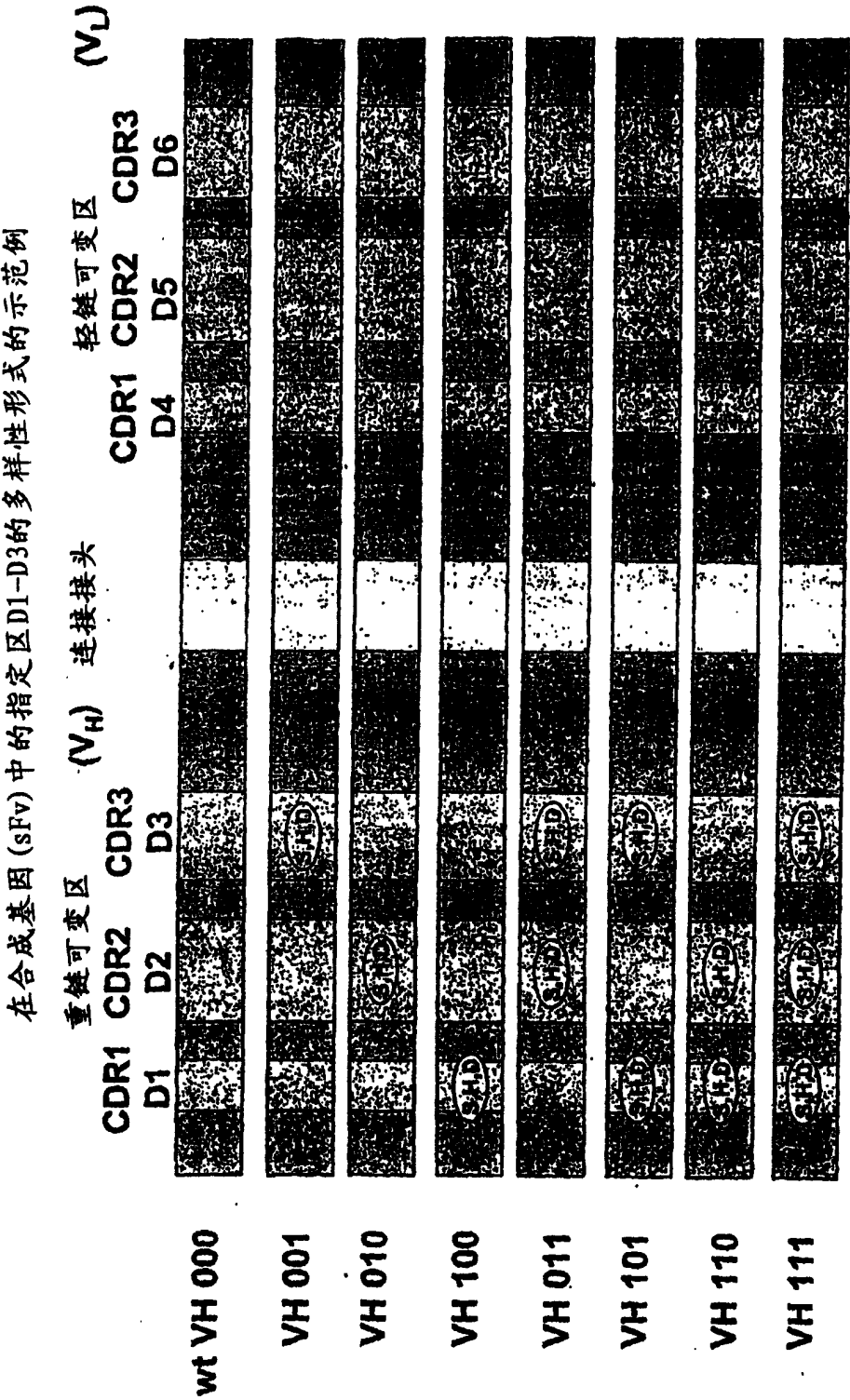
31	32	33	34	35
D	F	N	N	G
D	D	N	M	G
D	D	D	M	C
D	F	D	D	G
D	F	Y	D	D
D	F	Y	M	D



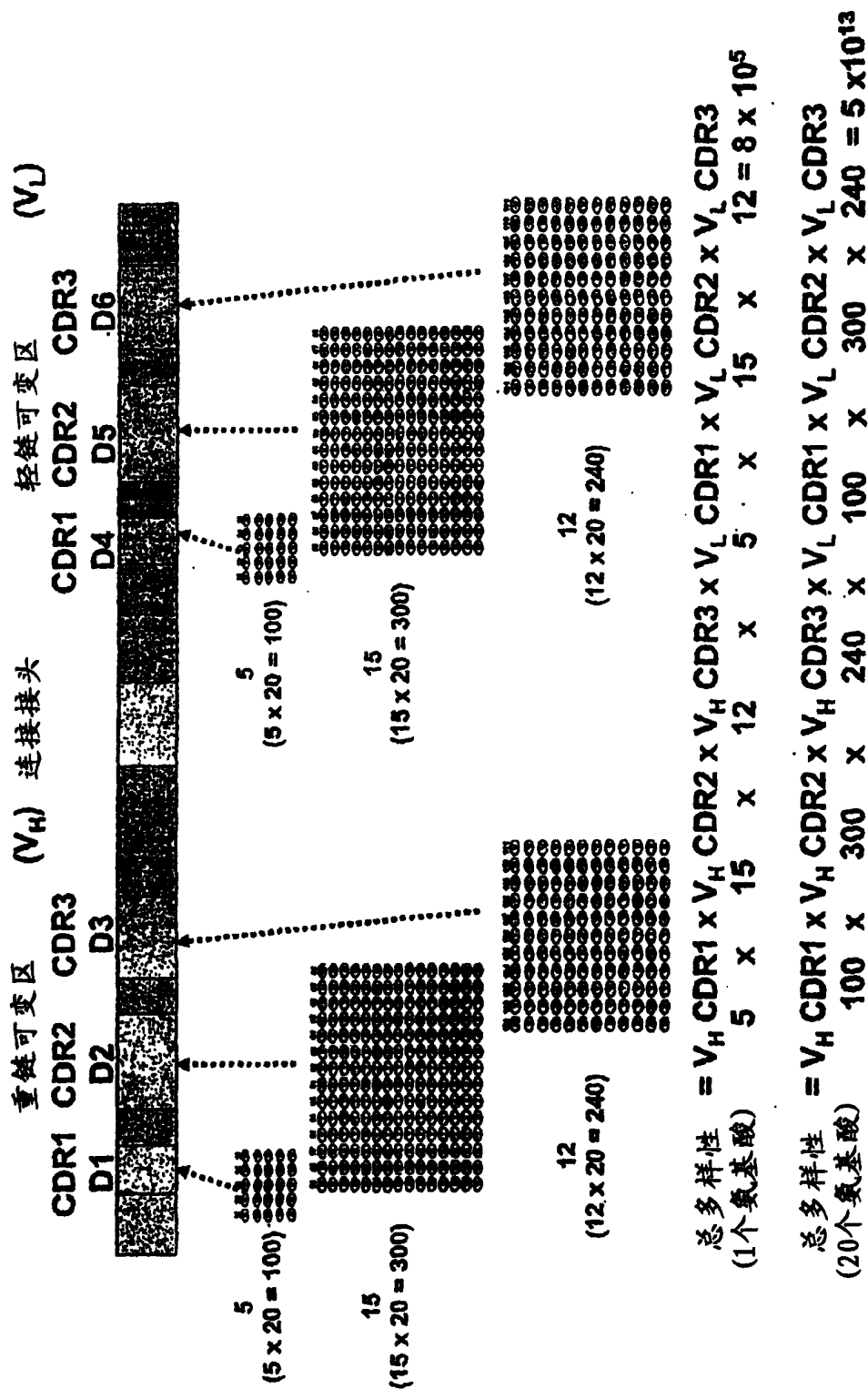
将指定区D1-D3在LTM后整合成为合成基团 (sFv)

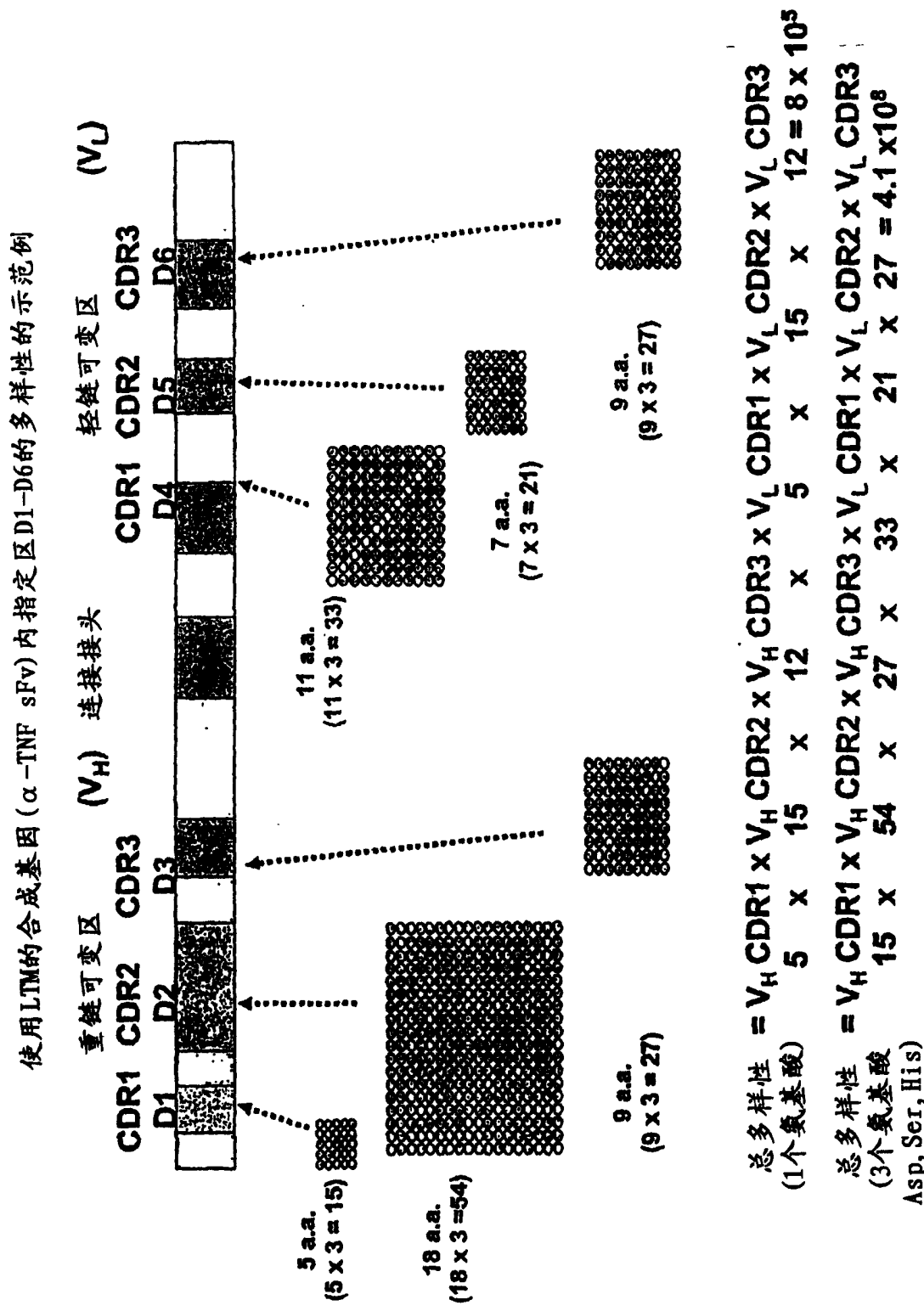


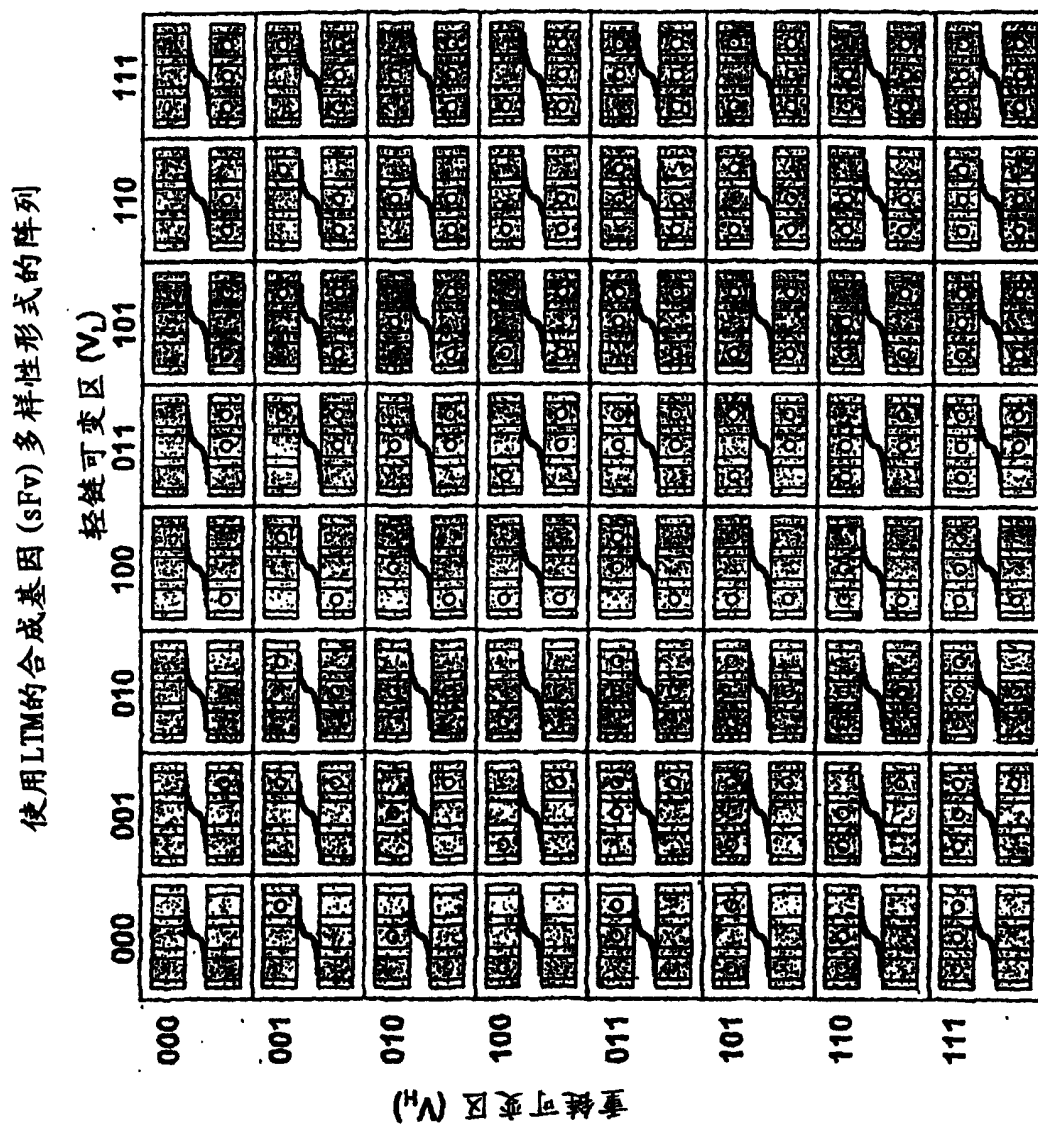


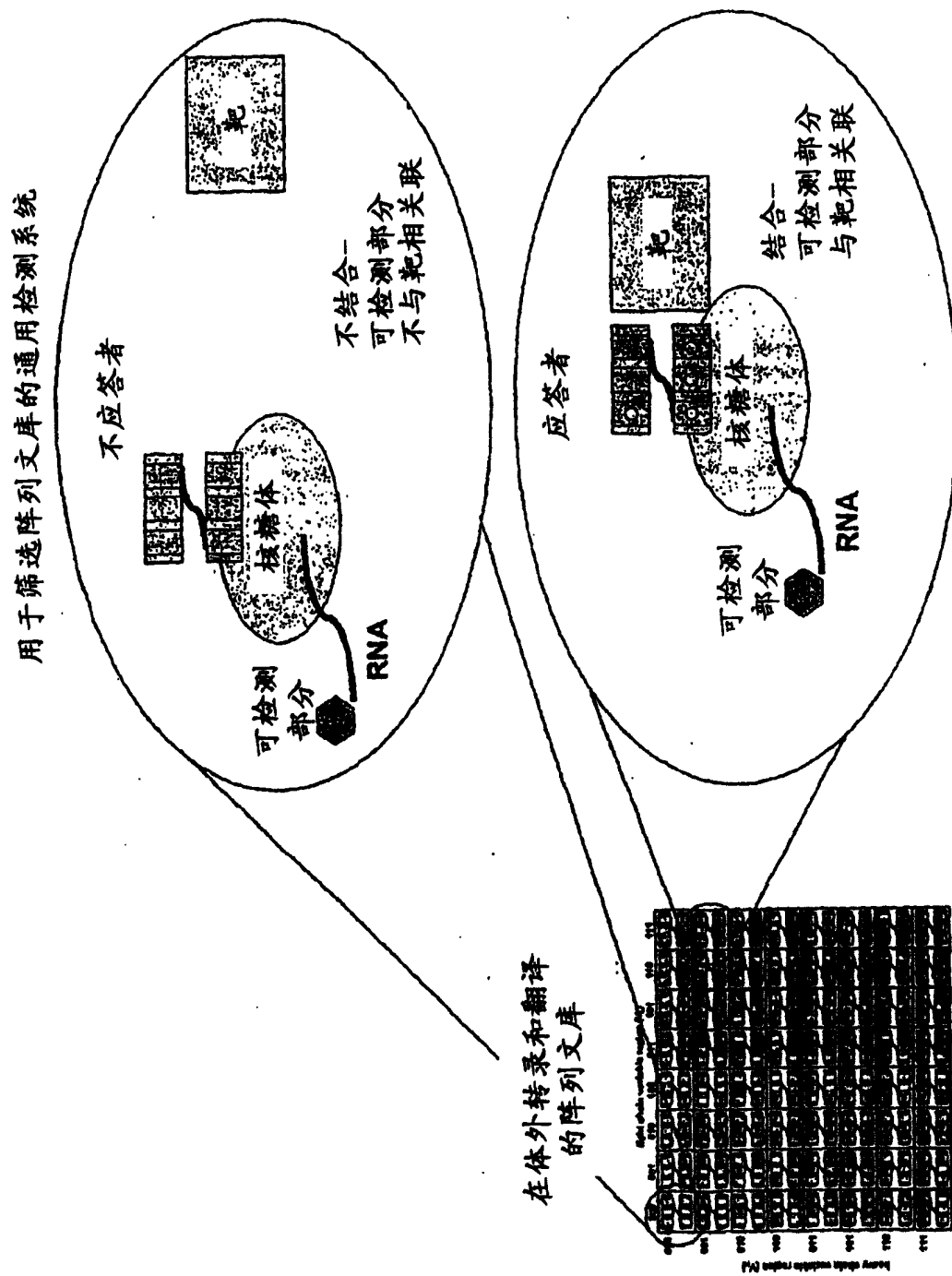


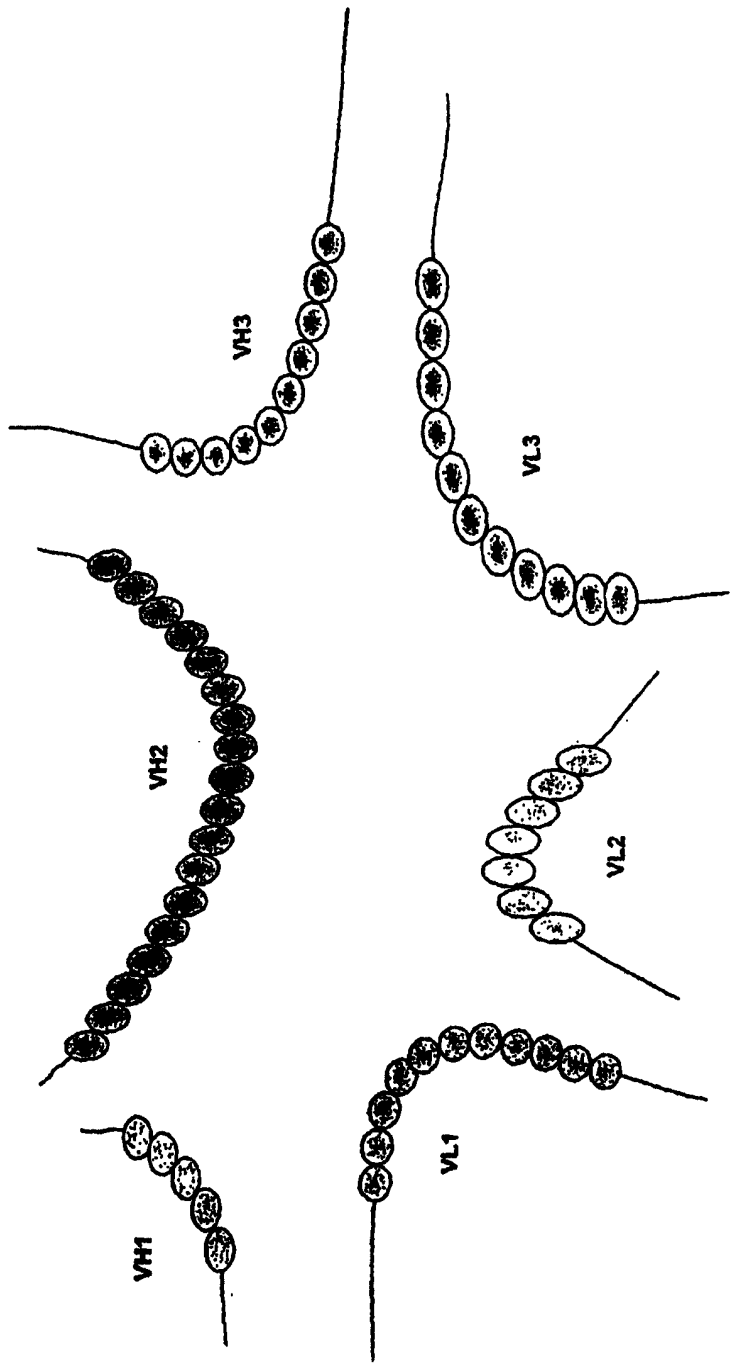
使用LTM的合成基因 (sFv) 内指定区D1-D6的多样性的示范例











cA2			
VH1: 5 A.A.	5 个寡核苷酸	5 个基因类似物	VL1: 11 A.A. 11 个寡核苷酸
VH2: 18 A.A.	18 个寡核苷酸	18 个基因类似物	VL2: 7 A.A. 7 个寡核苷酸
VH3: 9 A.A.	9 个寡核苷酸	9 个基因类似物	VL3: 9 A.A. 9 个寡核苷酸

商品化的 α -TNF scFv 的比对

1	A	V	D	E	N	Q	G	A	E	V	K	F	G	S	V	K	V	S	C	K	A	S	G	Y	I	F	I	D	Y	N	V	D	W	V	K	Q	A	CDP571 VH		
1	E	V	K	L	E	E	S	G	G	L	V	Q	P	G	S	N	K	L	S	C	V	A	S	G	F	I	F	S	N	R	W	N	N	V	R	Q	S	CA2 VH AA		
41	P	G	D	E	N	Q	G	A	E	V	K	F	G	S	V	K	V	S	C	K	A	S	G	Y	I	F	I	D	Y	N	V	D	W	V	K	Q	A	CDP571 VH		
41	P	E	K	G	I	E	V	V	A	E	I	N	S	K	E	I	N	S	A	T	H	Y	A	E	S	V	K	G	R	F	I	S	R	D	D	S	K	S	A	CA2 VH AA
79	A	V	D	E	N	Q	G	A	E	V	K	F	G	S	V	K	V	S	C	K	A	S	G	Y	I	F	I	D	Y	N	V	D	W	V	K	Q	A	CDP571 VH		
81	V	L	Q	E	I	D	L	R	T	E	D	T	G	V	Y	C	S	R	N	Y	Y	G	-	-	-	S	T	Y	D	Y	N	G	Q	G	T	T	L	T	CA2 VH AA	
119	V	S	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CDP571 VH			
149	V	S	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CA2 VH AA			
1	D	E	K	N	Q	G	A	E	V	K	F	G	S	V	K	V	S	C	K	A	S	G	Y	I	F	I	D	Y	N	V	D	W	V	K	Q	A	CDP571 VL			
1	D	E	K	N	Q	G	A	E	V	K	F	G	S	V	K	V	S	C	R	A	S	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CA2 VL AA			
41	F	V	Q	Q	R	T	W	G	S	P	R	L	I	S	W	A	S	T	R	E	S	G	V	P	S	R	F	I	G	S	G	S	G	T	E	F	T	L	T	CDP571 VL
35	F	V	Q	Q	R	T	W	G	S	P	R	L	I	S	W	A	S	T	R	E	S	N	S	G	I	P	S	R	F	S	G	S	G	T	D	F	T	L	S	CA2 VL AA
491	T	E	S	L	G	P	D	E	V	I	T	F	Q	Y	Y	D	Y	P	E	T	F	G	Q	G	T	K	Y	E	I	K	R	-	-	-	-	-	-	CDP571 VL		
25	T	E	S	L	G	P	D	E	V	I	T	F	Q	Y	Y	D	Y	P	E	T	F	G	Q	G	T	K	Y	E	I	K	R	-	-	-	-	-	-	CA2 VL AA		

