

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200780032857.X

[51] Int. Cl.

C07K 16/10 (2006.01)

C07K 16/16 (2006.01)

C07K 16/32 (2006.01)

[43] 公开日 2009年9月9日

[11] 公开号 CN 101528775A

[22] 申请日 2007.7.5

[21] 申请号 200780032857.X

[30] 优先权

[32] 2006.7.5 [33] AT [31] A1145/2006

[86] 国际申请 PCT/AT2007/000343 2007.7.5

[87] 国际公布 WO2008/003116 英 2008.1.10

[85] 进入国家阶段日期 2009.3.5

[71] 申请人 F-星生物科技研究和发展有限公司
地址 奥地利维也纳

[72] 发明人 G·希姆勒 G·雷德尔 F·鲁克

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 林毅斌 黄可峻

权利要求书5页 说明书74页 附图8页

[54] 发明名称

工程改造免疫球蛋白的方法

[57] 摘要

本发明涉及工程改造免疫球蛋白以及测定所述免疫球蛋白与抗原表位结合的方法、由该方法制备的免疫球蛋白、和免疫球蛋白文库，所述免疫球蛋白包含可变结构域和在该免疫球蛋白的至少两个结构环中的至少一个修饰，其中未修饰的免疫球蛋白明显不与所述表位结合，所述方法包含以下步骤：提供编码包含至少两个结构环的免疫球蛋白的核酸，修饰每个所述结构环的至少一个核苷残基，将所述经修饰的核酸转移到表达系统中，表达所述经修饰的免疫球蛋白，使所表达的修饰免疫球蛋白与表位接触，并且测定所述修饰免疫球蛋白是否与所述表位结合。

1. 一种方法，用于工程改造免疫球蛋白并测定其与抗原表位的结合情况，所述免疫球蛋白包括可变结构域和该免疫球蛋白的至少两个结构环中的至少一个修饰，其中未修饰的免疫球蛋白明显不与所述表位结合；所述方法包含以下步骤：

- 提供编码包含至少两个结构环的免疫球蛋白的核酸，
- 修饰每个所述结构环的至少一个核苷酸残基，
- 将所述经修饰的核酸转移到表达系统中，
- 表达所述经修饰的免疫球蛋白，
- 使所表达的修饰免疫球蛋白与表位接触，并且
- 确定所述经修饰的免疫球蛋白是否与所述表位结合。

2. 权利要求 1 的方法，其中所述免疫球蛋白与至少两个不同的表位特异性结合。

3. 权利要求 1-2 中任一项的方法，其特征在于所述免疫球蛋白来源于人、骆驼或鼠。

4. 权利要求 1-3 中任一项的方法，其特征在于所述可变结构域选自 VH、V_κ、V_λ、VHH 和它们的组合。

5. 权利要求 1-4 中任一项的方法，其特征在于：所述 VH、V_κ、V_λ 或 VHH 的环在氨基酸 7-21、氨基酸 25-39、氨基酸 41-81、氨基酸 83-85、氨基酸 89-103 或氨基酸 106-117 中包含至少一个修饰，其中所述结构域的氨基酸位置编号是根据 IMGT 的。

6. 权利要求 1-4 中任一项的方法，其特征在于：所述来源于人的 VH 或 V_κ 或 V_λ 的环在氨基酸 8-20、氨基酸 44-50、氨基酸 67-76 和氨基酸 89-101，最优选氨基酸位置 12-17、氨基酸位置 45-50、氨基酸位置 69-75 和氨基酸位置 93-98 内包含至少一个修饰，其中所述结构域的氨基酸位置编号是根据 IMGT 的。

7. 权利要求 1-4 中任一项的方法，其特征在于：所述来源于鼠的

VH 环在氨基酸 6-20、氨基酸 44-52、氨基酸 67-76 和氨基酸 92-101 内包含至少一个修饰，其中所述结构域的氨基酸位置编号是根据 IMGT 的。

8. 权利要求 1-4 中任一项的方法，其特征在于：所述来源于骆驼的 VHH 环在氨基酸 7-18、氨基酸 43-55、氨基酸 68-75 和氨基酸 91-101 内包含至少一个修饰，其中所述结构域的氨基酸位置编号是根据 IMGT 的。

9. 权利要求 1-8 中任一项的方法，其特征在于：所述经修饰的免疫球蛋白进一步与一个或多个经修饰的免疫球蛋白或与未修饰免疫球蛋白或其组成部分组合，得到组合免疫球蛋白。

10. 权利要求 9 的方法，其特征在于用重组技术组合所述免疫球蛋白。

11. 权利要求 1-10 中任一项的方法，其特征在于：对核酸的至少一个核苷酸的修饰导致由所述核酸编码的免疫球蛋白的一个或多个氨基酸的取代、缺失和/或插入。

12. 权利要求 1-11 中任一项的方法，其特征在于每个所述结构环的至少一个氨基酸被定点随机突变所修饰。

13. 权利要求 12 的方法，其特征在于经随机修饰的核酸分子包含至少一个具有序列 5'-NNS-3'、5'-NNN-3'或 5'-NNK-3'的核苷酸重复单位。

14. 免疫球蛋白用于制备免疫球蛋白文库的用途，所述免疫球蛋白通过权利要求 1-13 中任一项的方法获得。

15. 包含至少 10 个免疫球蛋白的文库，所述免疫球蛋白通过权利要求 1-13 中任一项的方法获得且在至少两个结构环上具有修饰。

16. 权利要求 15 的文库，其特征在于：含有在至少 2 个结构环上具有至少 4 个氨基酸位置突变的免疫球蛋白。

17. 权利要求 15-16 中任何一项的文库，其特征在于：包含至少 10 个免疫球蛋白且所述免疫球蛋白含有在结构环中具有修饰的可变

结构域。

18. 权利要求 15-17 中任何一项的文库，其特征在于：包含具有至少 2 个修饰的可变结构域的免疫球蛋白，所述修饰的可变结构域中的每一个的结构环上具有至少一个修饰。

19. 权利要求 15-18 中任何一项的文库，其特征在于：包含在至少两个结构环上具有至少 4 个修饰的免疫球蛋白。

20. 权利要求 15-19 中任何一项的文库，其特征在于：包含至少 10 个含有经修饰的单链免疫球蛋白的免疫球蛋白，在每个单链免疫球蛋白的结构环上具有至少两个修饰。

21. 权利要求 15-20 中任何一项的文库，其特征在于：包含至少 10 个含有经修饰的 VHH 结构域的免疫球蛋白，在每个 VHH 结构域的结构环上具有至少两个修饰。

22. 权利要求 15-21 中任何一项的文库，其特征在于：包含至少 10 个含有经修饰的人抗体的免疫球蛋白，在每个人抗体的结构环区上具有至少两个修饰。

23. 权利要求 15-22 中任何一项的文库，其特征在于：包含至少 10 个含有经修饰的 Fab 片段的免疫球蛋白，在每个 Fab 片段的结构环区上具有至少两个修饰。

24. 权利要求 15-23 中任何一项的文库，其特征在于：包含选自 VH、V_K、V_λ 和 VHH 的免疫球蛋白。

25. 根据权利要求 1-13 中任何一项的方法可得到的经修饰的免疫球蛋白，其具有通过结构环上至少两个修饰而对抗原特异的结合位点。

26. 权利要求 25 的经修饰的免疫球蛋白，其中所述抗原是人血清白蛋白。

27. 权利要求 26 的经修饰的免疫球蛋白，其中所述抗原是 Fc 受体。

28. 特异性结合和/或检测分子的方法，包含以下步骤：

(a) 使权利要求 15-24 中任何一项的修饰免疫球蛋白的文库或通过权利要求 1-13 中任何一项的方法可得到的修饰免疫球蛋白的文库与含有所述分子的试样接触, 并且任选地

(b) 检测特异性免疫球蛋白/分子复合体的潜在形成情况。

29. 特异性分离与分子结合的经修饰的免疫球蛋白的方法, 包含以下步骤:

(a) 使权利要求 15-24 中任何一项的修饰免疫球蛋白的文库或通过权利要求 1-13 中任何一项的方法可得到的修饰免疫球蛋白的文库与含有所述分子的样品接触,

(b) 分离形成的特异的修饰免疫球蛋白/分子复合物, 并且

(c). 任选从所述复合物分离修饰的免疫球蛋白。

30. 结合配偶体试剂盒, 其含有

(a) 权利要求 15-24 中任何一项的修饰免疫球蛋白的文库或通过权利要求 1-13 中任何一项的方法可得到的修饰免疫球蛋白的文库, 和

(b) 含有抗原表位的结合分子。

31. 权利要求 30 的试剂盒的结合分子的用途, 其用于从权利要求 15-24 中任何一项的文库中选择修饰的免疫球蛋白。

32. 变体可变结构域多肽, 其包含至少两个经修饰的结构环, 其中每个修饰结构环的至少一个氨基酸选自: 色氨酸、酪氨酸、苯丙氨酸、组氨酸、异亮氨酸、丝氨酸、甲硫氨酸、丙氨酸和天冬酰胺。

33. 权利要求 32 的变体可变结构域多肽, 其特征在于: 至少一个修饰结构环包含至少两个所述氨基酸。

34. 权利要求 32-33 中任何一项的变体可变结构域多肽, 其特征在于: 它是来源于人或人源化的变体可变结构域多肽, 并且在位置 12-17、45-50、69-75 和 93-98 中的任何一个包含至少一个酪氨酸, 和/或在位置 12-17、45-50、69、71-75、93-94 和 96-98 中的任何一个包含至少一个色氨酸, 和/或在位置 12-17、46、47、49、50、69-74 和 93-98 中的任何一个包含至少一个组氨酸, 和/或在位置 12-17、45-47、49、

50、70-73、75、94-96 和 98 中的任何一个包含至少一个天冬酰胺，和/或在位置 12-17、46-50、69-71、73-75、93、95、96-98 中的任何一个包含至少一个甲硫氨酸，和/或在位置 13、71、75、94、95 和 98 中的任何一个包含至少一个丝氨酸，和/或在位置 12、14-17、45-50、69、70、72-75、93 和 96-98 中的任何一个包含至少一个异亮氨酸，和/或在位置 15、46、48、70-73、75、93、95 和 98 中的任何一个包含至少一个苯丙氨酸。

工程改造免疫球蛋白的方法

发明领域

本发明涉及工程改造和制备包含经修饰的免疫球蛋白可变区多肽的分子的方法。

一般领域是工程改造蛋白质以赋予其特异性结合的特性。更具体地讲，这里相关的工程蛋白质是免疫球蛋白(抗体)，甚至更具体地讲，是免疫球蛋白的单个可变结构域(variable domain)、或单个可变结构域对或组合、或与其它免疫球蛋白结构域的组合。免疫球蛋白的特异性结合特性是重要的特征，因为它们控制着与其它分子如抗原的相互作用，使得免疫球蛋白可用于诊断和治疗的应用。

背景

抗体的结构

不同物种和不同类别的抗体的基本结构是相似的，在这里用一个完整 IgG1 免疫球蛋白的作为例子来解释。

两个同样的重(H)链和两个同样的轻(L)链结合形成 Y 形抗体分子。重链各有四个结构域。氨基末端可变结构域(VH)在 Y 的顶部。后接有三个恒定区结构域(constant domain) CH1、CH2 和在 Y 主干的底部的羧基端 CH3。一段短的区段：转换区(switch)连接重链可变区和恒定区。铰链将 CH2 和 CH3 (Fc 片段)连接到抗体(Fab 片段)的剩余部分。一个 Fc 和两个同样的 Fab 片段可通过蛋白酶切割完整抗体分子的铰链而产生。轻链由被转换区分隔的可变(VL)和恒定(CL)两个结构域组成。

铰链区的二硫键连接着两条重链。轻链通过另外的二硫键与重链连接。

重链和轻链的可变区(VL)和(VH)位于 Y 的“顶部”，它们定位于此

与抗原反应。该分子的这个顶部是氨基酸序列 N - 末端所在的一端。

Y 的主干表现在某种程度上有效地介导效应功能例如补体活化和与 Fc 受体相互作用、或 ADCC 和 ADCP。其 CH2 和 CH3 结构域凸出以方便与效应蛋白相互作用。氨基酸序列 C - 末端位于顶部的对边，可称为 Y 的“底部”。

免疫球蛋白可变结构域(V 结构域)

抗体分子中每个结构域有一个相似的结构：两个 β 片层在紧缩的反向平行 β 桶中彼此紧压。该守恒的结构被定义为免疫球蛋白折叠。参考 Bork 等(1994)J. Mol. Biol. 242:309-320; Halaby 等(1999) Protein Engineering 12: 563-571; Immunobiology . 5th ed. Janeway, Charles A.; Travers, Paul; Walport, Mark; Shlomchik, Mark. New York and London: Garland Publishing; 2001。

可变结构域折叠有 9 个 β 股，排列在 4 股和 5 股的两个片层中。上述 5 股的片层与恒定结构域的 3 股片层结构同源，但是含有额外的股 C 和股 C'。剩余的股(A、B、C、D、E、F、G)与其在恒定结构域免疫球蛋白折叠中的负体有同样的拓扑结构和相似的结构。如在恒定结构域中一样，二硫键连接在相对的片层中的股 B 和股 F。免疫球蛋白轻链和重链的可变结构域都含有三个高变环，即互补决定区(CDR)。V 结构域的三个 CDR (CDR1、CDR2、CDR3)群集在 β 桶的一端。CDR 是连接免疫球蛋白折叠的 β 链 B-C、C'-C'和 F-G 的环。CDR 上的残基随各种免疫球蛋白分子的不同而不同，赋予每种抗体对抗原的特异性。

在抗体分子顶部的 VL 和 VH 结构域紧压在一起以至 6 个 CDR(每个结构域 3 个)配合构成特异性结合抗原用的表面(或空腔)。因而抗体的天然抗原结合位点由连接轻链可变结构域的股 B-C、股 C'-C' 和股 F-G 和重链可变结构域的股 B-C、股 C'-C' 和股 F-G 的环组成。

用于蛋白质工程的支架

利用蛋白质 3D 结构作为设计的辅助，用蛋白质的核心结构作为

支架，使位于许多蛋白质表面的氨基酸残基随机化。以下参考文献中描述或综述了该策略的实例，以引用的方式并入本文：Nygren PA, Uhlen M., *Curr Opin Struct Biol.* (1997) 7:463-9; Binz HK, Amstutz P, Kohl A, Stumpp MT, Briand C, Forrer P, Grutter MG, Pluckthun A. *Nat Biotechnol.* (2004) 22:575-82; Vogt M, Skerra A. *Chembiochem.* (2004) 5:191-9; US 6,562,617; Hufton 等 *FEBS Letters* (2000) 475: 225; Binz 等 *Nat Biotechnol.* (2005) 23:1257-68; Hosse 等 *Protein Sci.* (2006)15:14-27。

该技术的基本原理是基于对许多蛋白质具有稳定核心的观察，该核心通过如 β 片层或 α 螺旋的二级结构元件的特定排列而形成，而二级结构元件通过环、转角或无规卷曲结构相互连接。通常，后面这三种结构元件对于蛋白质整体结构的重要性较小，在这些结构元素中的氨基酸残基可被交换而常常不会破坏总的蛋白质折叠。这一设计原理的天然例子是如可在抗体、T-细胞受体和其它免疫球蛋白超家族分子中发现的免疫球蛋白样结构域的 CDR。

免疫球蛋白可变结构域 CDR 环的操作

很多现有技术文献显示免疫球蛋白样支架被用于对现有的抗原或配体结合位点进行操作以引入新的结合特性的目的。更确切地讲，主要是对 CDR 区进行抗原结合工程改造，换言之对于免疫球蛋白折叠而言，修饰天然抗原结合位点以改变其结合亲和力或特异性。现有大量文献描述这些被操纵的免疫球蛋白的不同形式，通常以完全抗体、融合产物和/或片段如单链 Fv 片段(scFv)、双抗体、微型抗体、单结构域或 Fab 片段等等的形式表达，或者在噬菌体颗粒或其它病毒和细胞表面展示，或者在各种原核或真核表达系统中可溶性表达。这些技术例如在 Holliger & Hudson., *Nat. Biotechnol.* (2005) 23:1126-36 和 Hoogenboom, *Nat. Biotechnol.* (2005)23:1105-16 中综述。

免疫球蛋白可变结构域的骨架或非 CDR 区

免疫球蛋白可变结构域的 CDR 环决定抗原特异性。分子的剩余部分叫做骨架(FR)。然而这些骨架区由 β 股和环结构组成。

天然免疫球蛋白可变结构域中的非 CDR 环的环没有抗原结合或表位结合特异性，但是有助于免疫球蛋白结构域的正确总体折叠，此也有助于 CDR 的正确定位和结构域之间的相互作用。为了本发明的意图，将这些环称为结构环。

通常出于很多不同的原因对抗体可变结构域进行操作，例如各种抗体形式(format)的构建、CDR 移植(也就是将特定抗体的特异性移植到另一个骨架；例如 Jones 等 *Nature* (1986) 321: 522-525; Kashmiri 等 *Methods* (2005) 36:25-34)、改变可变结构域的表面以使其更具可溶性和稳定性(例如 Ewert 等 *Methods* (2004) 34:184-99; Conrath 等 *J Mol Biol.* (2005) 350:112- 125)、使其变为单体(例如 Dottorini 等 *Biochemistry* (2004) 43: 622-628))或研究可变结构域之间的相互作用(例如 Masuda 等 *FEBS J.* (2006) 273:2184-94)。很多那些操作已经包括了改变分子骨架区；已经完成了一些在可变结构域的结构环内部的氨基酸突变。

从 CDR 移植结果看，远端骨架区对 CDR 环定位的影响是明显的，CDR 移植结果显示骨架氨基酸的频繁突变对于在 CDR 从一个骨架移植到另一个骨架之后恢复抗原结合是必需的(例如 Foote & Winter (1992) *J. Mol. Biol.* 224: 487- 499; Kettleborough 等 *Protein Eng.* (1991) 4:773-783; Wu 等 *J. Mol. Biol.* (1999) 294:151-162)。

Simon & Rajwesky (*Protein Sci.* (1992) 5:229-234)研究了将四个残基插入来自抗-NP 抗体 B1-8 的重链可变区的 FR3 环的影响。插入突变体实现在生物合成中没有重要缺陷的分泌型抗体，表明抗体可变结构域不仅可在互补决定区(CDR)而且在骨架区(FR)环适应长度的变化。在这种情况下，由 CDR 环形成的初始抗原结合位点不受邻近结构环上的修饰的影响。

CDR 环移植到结构环区

EP0640130B1 描述了具有超过一个生物学结合位点(单个 V 结构域或多个 Fv)的嵌合免疫球蛋白超家族蛋白类似物(嵌合蛋白

(chi-protein)。这些蛋白上的结合位点由来源于与免疫球蛋白分子超家族的分子相关的分子的高变区组成，这些分子包括免疫球蛋白、细胞表面抗原(如 T 细胞抗原)和细胞受体(如 Fc 受体)。所述高变区叫做“CDR 样区”并确定配体结合位点。另外，该嵌合蛋白具有至少一个以上的配体结合位点区段，还有 CDR 样区，拼接到 β 桶结构域的 FR 样区中。

因此该嵌合蛋白的每个配体结合位点由来自免疫球蛋白超家族的分子的 CDR 样区组成。例如，配体结合位点由来自其配体是抗原的免疫球蛋白分子的 CDR 组成。

这样，EP0640130B1 教导怎样将来自免疫球蛋白超家族分子的具有给定特异性的 CDR 样区剪接成可变结构域的结构环。推测用该技术可制备双特异性功能抗体。对该技术来说存在以下需要：对于可变结构域而言，使 CDR 样环的相对定位(CDR 环对称性)复制成合理近似于结构环的相对定位。EP0640130B1 声称该 CDR 样环对称性对于结构环的近似是存在的。然而，CDR 环的相对定位与结构环的相对定位在细节和分辨率(resolution)上的足够相似性是令人怀疑的，因此，至今为止还没有用该技术开发双特异性分子的实际可能性的描述。

EP0640130B1 举例说明了 R19.9(特异于对偶氮苯胍酸盐的鼠类单克隆抗体)和 26-10(特异于乌本苷的鼠类单克隆抗体)分别被用作提供主要 CDR 环(primary CDR loop)的骨架，并且将鼠类抗溶菌酶抗体 D1.3 的 CDR 环移植到结构环区。然而，没有描述移植后的功能特异性。

另一个例子描述了特异于乌本苷的单链抗体 26-10 能够在将来自溶菌酶特异抗体的两个 CDR 移植到乌本苷特异的单链 Fv 抗体片段的结构环之后保持其乌本苷特异性。然而，没有描述根据该方法制备的抗体片段还具有溶菌酶结合特异性。

多肽移植到结构环区

WO00244215A2 描述了包含特异性靶结合位点和 Fc 效应肽的结

合分子。Fc 效应肽是 100 多个氨基酸的与效应分子相互作用的多肽。效应肽可以例如插入抗体环区中，前提是结合抗原的能力不受到不利影响。例证了效应肽被插入到免疫球蛋白片段的 CH1 结构域的非 CDR 环。对于可变结构域同样的插入没有进行描述。根据该公开说明书移植到非 CDR 环的每个多肽由于其所在的不同的结构环境而具有很高的失活机会。另外，如果多肽移植到可变结构域的结构环中，那么在亲本免疫球蛋白中保持特异的 CDR 环构象可能是困难的。因此，没有描述效应肽可既不损失抗原结合也不损失效应分子结合而被移植到可变结构域。

PCT/EP2006/050059 描述了工程改造免疫球蛋白以获得新的抗原结合位点的方法，该免疫球蛋白在结构环区包含修饰。该方法广泛应用于免疫球蛋白并可用于制备一系列靶向多种抗原的免疫球蛋白。

US2005/266000A1 描述了包含变体重链可变骨架结构域(VFR)的多肽。VFR 是可与抗原接触的抗原结合袋或沟的一部分。VFR 是 CDR 环区(在本文中也被称为 CDR 区)的一部分，位于在 CDR 环旁边的可变结构域，通过 CDR 环区支持抗原结合。没有为了工程改造抗原结合位点而对骨架环(而非 VFR)进行诱变。

本发明的目标是提供具有改良的抗原结合特性的免疫球蛋白和可变免疫球蛋白结构域、以及工程改造和制备具有改良的可变结构域的免疫球蛋白的方法。

发明描述

本发明主题满足此目标的需要。

本发明涉及工程改造免疫球蛋白并测定所述免疫球蛋白与抗原表位结合情况的方法，该免疫球蛋白包含可变结构域和在所述免疫球蛋白的至少两个结构环内的至少一个修饰，其中未修饰的免疫球蛋白明显不与所述表位结合，所述方法包括以下步骤：

- 提供编码包含至少两个结构环的免疫球蛋白的核酸，
- 修饰每个所述结构环的至少一个核苷酸残基，

- 将所述经修饰的核酸转移到表达系统中，
- 表达所述经修饰的免疫球蛋白，
- 使所表达的修饰免疫球蛋白与表位接触，和
- 测定所述的修饰免疫球蛋白是否与所述表位结合。

根据本发明工程改造的结构环优选位于一个或两个结构环区的内部，在某些情况下将修饰置于在甚至超过2个的结构环区内。

为了克服在 WO00244215A2 和 EP0640130B1 中所描述的将结合位点工程改造到在抗体可变结构域的结构环区中的现有技术水平的缺点，本发明提供方法以对多肽(即包含修饰的结构环区的可变结构域肽)进行其天然环境(context) (即抗体可变结构域)中的特异性结合的选择。为了增加可根据抗原结合情况进行选择的变体免疫球蛋白结构域的潜在结构的数量，根据本发明对至少两个结构环或环区进行修饰。因此引入和选择对总体结构域结构干扰最小的许多不同的修饰是可能的。修饰至少两个环或环区的另一个优势是：扩大的表面积有可能与特异性结合配偶体相互作用。根据特异性功能或结合情况，对如此修饰的免疫球蛋白可变结构域进行选择。本发明能够对修饰的免疫球蛋白可变结构域和修饰的免疫球蛋白进行工程改造，这两者使它们的经修饰的结构环或环区以高亲和力结合和/或高特异性与结合配偶体结合。该方法允许选择具有最小限度的修饰且没有破坏免疫球蛋白可变结构域总体结构的特异性结合分子。

为了克服难以保持亲本免疫球蛋白中的特定 CDR 环构象的困难，根据本发明对可变结构域的结构环进行了修饰。因此这是第一次可能为免疫球蛋白增加抗原结合特性而不显著损失亲本抗体的抗原结合乃至效应分子结合。令人异想不到的是，有可能根据本发明工程改造免疫球蛋白而没有明显干扰亲本免疫球蛋白特定 CDR 环构象的结合特性，例如结合亲和力、亲合力和特异性。

特定的 CDR 环构象为抗原结合提供 CDR 环或 CDR 环区。根据本发明可修饰任何经由这种 CDR 环构象与抗原结合的免疫球蛋白，

以获得在免疫球蛋白结构环上的另外结合位点，优选同时该特定的 CDR 环构象和亲本分子的初始抗原结合特性可保持或不变。优选 CDR 环构象的结合特性可维持到亲本分子的结合亲和力没有显著减少的程度，例如，解离常数 K_d 没有显著增加，意味着 K_d 的差异是优选小于 10^4 、更优选小于 10^3 、甚至优选小于 10^2 或小于 10 的倍数。

依照本发明，本发明的一个关键特征是免疫球蛋白或免疫球蛋白可变结构域的工程改造发生在通常不涉及抗原结合的区域，换句话说，发生在抗体可变结构域的 CDR 之外的区域。据观察免疫球蛋白结构域的特异折叠允许在与 CDR 结构类似但序列及结构的位置不同的区域引入随机突变。根据本发明鉴别用于修饰的区域是像 CDR 一样的连接免疫球蛋白折叠的 β 链的环区。如本发明所述，可对这些结构环或环区进行突变而不影响经过 CDR 环介导的免疫球蛋白可变结构域的结合。通过对所述的结构环或环区进行突变，产生新的分子结合表面或结合口袋，其大小和形状与由抗体天然抗原结合位点的 CDR 环形成的结合表面或结合口袋相似。因为通过插入另外的氨基酸也可使结构环延长，新生结合位点的结构可被自动调节到它应结合的靶标。例如，特别适合小分子结合的深的结合口袋可优先通过长环(即具有插入到它们之中的额外氨基酸的结构环)来形成；而适合与具有大且扁平的分子表面的靶标结合的相当平坦的结合表面，优先在对结构环上的残基进行不插入额外残基的突变时形成。更具体地说，本文描述的是，通过在连接人或人源化重链可变结构域的 β 股 A-B, C-D 和 E-F 的环中引入随机突变，可以选出特异性结合人血清蛋白或 Fc γ 受体 III 的突变结构域，而这两者通常不被人或人源化的免疫球蛋白可变结构域识别和结合。引入的突变包括通过用随机选择的残基置换被选择的野生型序列的氨基酸残基的突变，还包括在上述环中的额外氨基酸残基的插入。因此，优选根据本发明提供按上述方法可获得和生产的经修饰的免疫球蛋白，而且其对抗原、特别是血清白蛋白、细胞受体和补体因子、更特别是人血清白蛋白和 Fc 受体具有特异性结合位点。

具体而言，本发明涉及工程改造免疫球蛋白可变结构域和含有这种与抗原表位特异性结合的结构域的免疫球蛋白的方法。

具体地说，本发明的方法包含以下步骤：

- 提供编码与至少一个第一表位特异性结合的免疫球蛋白的核酸，而所述免疫球蛋白包含至少两个结构环或环区，
- 修饰由所述核酸编码的每个所述结构环或环区的至少一个核苷酸残基，
- 将所述经修饰的核酸转移到表达系统中，
- 表达所述经修饰的免疫球蛋白，
- 使所表达的修饰免疫球蛋白可变结构域与所述至少一个第二表位接触，并且
- 测定所述修饰免疫球蛋白可变结构域是否与第二表位特异性结合。

优选将此方法应用于免疫球蛋白可变结构域肽。更优选本发明该实施方案的方法涉及经由其 CDR 区或特定的 CDR 环构象与所述第一表位结合的免疫球蛋白。

本发明的方法进一步涉及：在所述免疫球蛋白可变结构域的至少两个结构环或环区中的每个的至少一个修饰、以及测定所述结构环或环区与至少一种抗原特异性结合的情况，其中含有未修饰结构环或环区的免疫球蛋白可变结构域不与这样的抗原特异性结合。

本文所用术语“免疫球蛋白”包括免疫球蛋白或免疫球蛋白的组成部分、片段或衍生物。因此它包括按照本发明修饰的“免疫球蛋白可变结构域肽”(本文所用术语免疫球蛋白和抗体是可以互换的)；和免疫球蛋白可变结构域、或它的含有结构环的部分如微型结构域(minidomain)、或这种结构域的结构环。免疫球蛋白可被用作分离的多肽或与其它多肽组合的分子。在某些情况下，优选将确定的修饰结构环或结构环区、或其组成部分用作以结合或组合为目的的分离分子。本文所定义的“免疫球蛋白可变结构域”含有这种在修饰和工程改

造后具有特异性结合特性的免疫球蛋白可变结构域肽或多肽。所述肽与免疫球蛋白可变结构域序列同源，优选长度为至少 5 个氨基酸，更优选至少 10 个或甚至至少 50 或 100 个氨基酸，至少部分构成可变结构域的结构环区或非 CDR 环区。优选所述肽排除了那些被认为是非功能氨基酸的插入、杂合或嵌合 CDR 区或 CDR 样区和/或 CDR 区正则结构(canonical structure)。结合特性涉及与特定表位结合、亲和力和亲合力。

根据本发明的免疫球蛋白衍生物是一种或多种本发明免疫球蛋白的任何组合和或融合蛋白，其中本发明免疫球蛋白的任何结构域或微型结构域可融合在一种或多种其它蛋白(如其它免疫球蛋白、配体、支架蛋白、酶、毒素等等)的任何位置。还可通过重组技术或与其它物质通过诸如共价偶联、静电作用、二硫键等的各种化学技术结合，来获得本发明免疫球蛋白的衍生物。

结合到免疫球蛋白的其它物质可以是脂类、糖类、核酸、有机和无机分子或它们的任何组合(例如 PEG、前药或药物)。衍生物还可为具有同样氨基酸序列但完全或部分由非天然的氨基酸或化学修饰的氨基酸制成的免疫球蛋白。

按本发明工程改造的分子被用作独立蛋白以及融合蛋白或衍生物，最典型以这样一种方式融合以致成为较大抗体结构或完整抗体分子的部分、或它的组成部分如 Fab 片段、Fc 片段、Fv 片段和其它。用工程改造的蛋白质生产双特异性、三特异性甚至同时携带多种特异性的分子将是可能的，同时根据这种分子的计划用途的需要同时控制和预选结合价数也是可能的。

本发明另一方面涉及除可变结构域 CDR 环构象之外还具有至少一个环区的免疫球蛋白，其特征为上述的至少一个环区包含形成至少一个修饰环区的至少一个氨基酸修饰，其中上述的至少一个修饰环区与抗原的至少一个表位特异性结合。

优选使至少一个经修饰的抗体结构域与至少一种其它结合分子

进行分子结合，所述经修饰的抗体结构域通过非可变序列或结构环与特异性配偶体结合，所述其它结合分子可以是抗体、抗体片段、可溶性受体、配体或另外的经修饰的抗体结构域。

由本发明免疫球蛋白特异性识别的、起着作为结合对的一部分作用的分子优选选自蛋白质分子、核酸和糖类。

本发明免疫球蛋白的修饰环或环区可与任何种类的结合分子或结构特异性结合，特别是与抗原、蛋白质分子、蛋白质、肽、多肽、核酸、多聚糖、糖类、脂类、小的有机分子、无机分子或它们的组合或融合物特异性结合。当然，所述经修饰的免疫球蛋白可包含至少两个环或环区，由此每个环或环区可与不同的分子或表位特异性结合。

按照本发明，各种细胞表面抗原的抗原结合区或抗原结合位点可被引入到给定抗体结构的结构环。

本发明免疫球蛋白的结合优选一方面通过 CDR 区来进行(前提是该免疫球蛋白含有这样的 CDR 区)，而在另一方面通过由修饰至少两个结构环而形成的额外结合位点来进行。那些结构环置于具有一个或多个修饰的结构环的一个或者至少两个可变结构域上。因此，通过至少两个可变结构域的至少一个修饰或者至少一个可变结构域的至少两个修饰，本发明免疫球蛋白在可变结构域的结构环中含有至少两个修饰。优选经修饰的免疫球蛋白通过改变蛋白质的一级结构或者通过改变三级结构来展示结合位点，以获得构象特异性结合位点。

以此类推，来自任何类别的免疫球蛋白和来自任何物种免疫球蛋白的免疫球蛋白可变结构域都可按照这类型工程改造进行处理。此外，不仅可以对本发明实施例中作为目标的特定环进行操作，还同样可对免疫球蛋白可变结构域中任何连接 β 股的环进行操作。

根据本发明可以将来自任何生物体和任何类别免疫球蛋白的工程免疫球蛋白或免疫球蛋白可变结构域用作例如单结构域或者较大分子的一部分。例如，它们可以是完整免疫球蛋白的组成部分，因此该免疫球蛋白具有其“正常的”由 6 个 CDR 中的至少一个形成的抗原结

合区和新的工程改造抗原结合位点。照这样，可产生多特异性，例如双特异性或三特异性免疫球蛋白。所述工程免疫球蛋白结构域也可能是任何融合蛋白的组成部分。

本发明的免疫球蛋白或免疫球蛋白可变结构域可为完整的抗体分子或抗体的部分，所述抗体例如有 IgG、IgA、IgM、IgD、IgE 等等。本发明免疫球蛋白或免疫球蛋白可变结构域也可包含以下或由以下组成：功能抗体片段如 Fab、Fab2、scFv、Fv 或其组成部分、或免疫球蛋白的其它衍生物或组合如小抗体(minibody)、可变区重链和轻链的结构域(如 Fd、VL，包括 V λ 和 V κ 、VH、VHH)以及由被至少两个结构环连接的两条免疫球蛋白结构域 β 股组成的微型结构域(例如，见 Laffly 等(2005) Hum Antibodies. 2005; 14 : 33-55)，作为分离的结构域或在天然缔合分子的环境(context)中。

根据本发明修饰的免疫球蛋白可能进一步与一个或多个修饰免疫球蛋白或未修饰免疫球蛋白或其组成部分结合，得到组合免疫球蛋白。优选通过重组技术获得组合，也可以通过经吸附、静电相互作用等的缔合，或者通过使用或不使用接头的化学结合来获得。优选的接头序列是天然接头序列或者是功能合适的人工序列。

可以理解的是，术语“免疫球蛋白”、“免疫球蛋白可变结构域肽”和“免疫球蛋白可变结构域”还包括衍生物。衍生物是一种或多种免疫球蛋白的任何组成部分或组合和/或融合蛋白，其中按照本发明获得的免疫球蛋白的任何结构域或微型结构域可在任何位置与一种或多种其它的肽或蛋白质组合或融合(包括但不限于其它的免疫球蛋白、免疫球蛋白结构域、Fc 部分、配体、支架蛋白质、酶、毒素、血清蛋白等等)。利用各种化学技术如共价偶联、静电相互作用、二硫键等通过将本发明未修饰或修饰的免疫球蛋白或免疫球蛋白可变结构域与其它物质相结合，也可获得本发明免疫球蛋白的衍生物。

与免疫球蛋白或免疫球蛋白可变结构域结合的其它物质可以是脂类、糖类、核酸、有机和无机分子或其任何组合(例如 PEG、前药

或药物)。衍生物也是有同样氨基酸序列的免疫球蛋白或免疫球蛋白可变结构域，但它完全或部分由非天然的、人工或化学修饰的氨基酸组成。

按照本发明工程改造的分子将被用作独立蛋白质以及融合蛋白质或衍生物，最通常在修饰前或后以成为较大抗体结构或完全抗体分子或其组成部分的一部分的方式融合。因此本发明的免疫球蛋白还可包含 Fab 片段、Fc 片段、Fv 片段、单链抗体、特别是单链 Fv 片段、双或多特异性 scFv、双抗体、多抗体、免疫球蛋白结构域的多价体或多聚体和其它，或由它们组成。用工程改造的蛋白质制备单特异性、双特异性、三特异性的分子和甚至可能携带更多特异性的分子将是可能的。通过本发明，根据这些分子的计划用途的需要同时控制和预选结合价是可能的。

按照本发明，可将抗原的一个或多个结合位点或者一种或多种抗原的抗原结合位点引入给定抗体可变域结构的结构环或环区。所述抗原可以是天然存在的分子或化学合成的分子或重组分子，在溶液或混悬液中，例如位于如固相的粒子表面或内部、细胞表面或内部或者病毒表面。令人意想不到的是，甚至当抗原仍粘附或结合在将阻碍其结合的分子和结构上时，按照本发明仍能够实现免疫球蛋白与抗原的结合。通过使用在其选择的或天然的环境和结构中的靶抗原，来选出经修饰和设计的免疫球蛋白，鉴定和获得那些最适用于诊断或治疗用途的修饰免疫球蛋白是可能的。

本文所用术语“抗原”将包含选自以下的分子：过敏原、肿瘤相关抗原、包括白蛋白的自身抗原、T 细胞受体、FcRn、细胞表面受体、酶、Fc 受体、补体系统蛋白、血清分子、细菌抗原、真菌抗原、原生动植物抗原和病毒抗原、还有引起传染性海绵状脑炎(TSE)的分子如感染性或非感染性的朊病毒和涉及阿尔茨海默症的标记或分子。

在一个优选的实施方案中，所述免疫球蛋白以其修饰的结构环特异地结合到至少两个这样的表位：为相同抗原或不同抗原的、彼此相

同或不同的表位。

例如，本发明方法涉及：工程改造与至少一个第一表位特异性结合的免疫球蛋白或免疫球蛋白可变结构域，其包含在所述免疫球蛋白的至少两个结构环或环区的每个中的至少一个修饰；测定所述环或环区与至少一个第二表位特异性结合的情况，所述表位选自上述抗原，其中未修饰的结构环或环区(非 CDR 区)不与所述的至少一个第二表位特异地结合。

本文所用术语抗原将具体包括作为整个靶分子或作为这种分子的片段(特别是靶标的亚结构，通常称作“表位”)的、能够被免疫球蛋白或抗体结构的结合位点识别的所有抗原和靶分子。

被本发明免疫球蛋白靶向的优选的抗原是那些已经被证明是或能够产生免疫性的、被免疫应答因子结合的、或免疫学或治疗学相关的、特别是那些临床功效已被检验的抗原或分子。

本发明术语“抗原”将意指已知与或能够与免疫球蛋白 CDR 环区相互作用同的分子或结构。与天然抗体有关的现有技术结构环区不与抗原相互作用但提供整体结构和/或有助于效应分子的结合。

按照一个优选的实施方案，与本发明免疫球蛋白结合的抗原是细胞表面抗原。术语“细胞表面抗原”包括能够被细胞表面的抗体结构识别的所有抗原或分子，以及这些分子的片段。优选的“细胞表面抗原”是那些已经被证明是或能够是免疫学或治疗学相关的、特别是那些临床前或临床功效已经被检验的抗原。那些介导细胞杀伤活性的细胞表面分子出于本发明的意图是尤其相关的。本发明免疫球蛋白与至少两种那些细胞表面分子结合后，免疫系统马上为细胞溶解或细胞死亡作准备，从而可提供攻击人类细胞的有效手段。

优选抗原选自细胞表面抗原，包括受体、特别是选自 erbB 受体酪氨酸激酶(如 EGFR、HER2、HER3 和 HER4，但不仅限于这些)的受体、TNF 受体超家族的分子如 Apo-1 受体、TNFR1、TNFR2、神经生长因子受体 NGFR、CD40、T 细胞表面分子、T 细胞受体、T 细胞抗

原 OX40、TACI 受体、BCMA、Apo-3、DR4、DR5、DR6、诱杀型受体如 DcR1、DcR2、CAR1、HVEM、GITR、ZTNFR-5、NTR-1、TNFL1 但不限于这些分子、B 细胞表面抗原如 CD10、CD19、CD20、CD21、CD22、实体瘤或血液癌细胞的抗原或标记、淋巴瘤或白血病细胞、包括血小板的其它血液细胞，但不限于这些分子。

根据本发明进一步优选的实施方案，与修饰结构环或环区结合的抗原或分子选自肿瘤相关抗原，特别是 EpCAM、肿瘤相关糖蛋白-72(TAG-72)、肿瘤相关抗原 CA 125、前列腺特异性膜抗原(PSMA)、高分子量黑色素瘤相关抗原(HMW-MAA)、表达 Lewis Y 相关糖的肿瘤相关抗原、癌胚抗原(CEA)、CEACAM5、HMFG PEM、粘蛋白 MUC1、MUC18 和细胞角蛋白肿瘤相关抗原、细菌抗原、病毒抗原、过敏原、变态反应相关分子 IgE、cKIT 和 Fc- ϵ -受体 I、IRp60、IL-5 受体、CCR3、红血球受体(CR1)、人血清白蛋白、小鼠血清白蛋白、大鼠血清白蛋白、新生 Fc- γ -受体 FcRn、Fc- γ -受体 Fc- γ -RI、Fc- γ -RII、Fc- γ -RIII、Fc- α -受体、Fc- ϵ -受体、荧光素、溶菌酶、toll 样受体 9、促红细胞生成素、CD2、CD3、CD3E、CD4、CD11、CD11a、CD14、CD16、CD18、CD19、CD20、CD22、CD23、CD25、CD28、CD29、CD30、CD32、CD33 (p67 蛋白)、CD38、CD40、CD40L、CD52、CD54、CD56、CD64、CD80、CD147、GD3、IL-1、IL-1R、IL-2、IL-2R、IL-4、IL-5、IL-6、IL-6R、IL-8、IL-12、IL-15、IL-18、IL-23、 α 干扰素、 β 干扰素、 γ 干扰素；TNF- α 、TNF β 2、TNF α 、TNF $\alpha\beta$ 、TNF-R1、TNF-RII、FasL、CD27L、CD30L、4-1BBL、TRAIL、RANKL、TWEAK、APRIL、BAFF、LIGHT、VEG1、OX40L、TRAIL 受体-1、A1 腺苷受体、淋巴毒素 β 受体、TACI、BAFF-R、EPO；LFA-3、ICAM-1、ICAM-3、整联蛋白 β 1、整联蛋白 β 2、整联蛋白 α 4/ β 7、整联蛋白 α 2、整联蛋白 α 3、整联蛋白 α 4、整联蛋白 α 5、整联蛋白 α 6、整联蛋白 α v、 α V β 3 整联蛋白、FGFR-3、角质形成细胞生长因子、VLA-1、VLA-4、L-选择蛋白、抗-Id、E-选择蛋白、HLA、HLA-DR、CTLA-4、T 细胞

受体、B7-1、B7-2、VNR 整联蛋白、TGFβ1、TGFβ2、嗜伊红粒细胞趋化蛋白 1、BLyS(B-淋巴细胞刺激因子)、补体 C5、IgE、IgA、IgD、IgM、IgG、因子 VII、CBL、NCA 90、EGFR (ErbB-1)、Her2/neu (ErbB-2)、Her3 (ErbB-3)、Her4 (ErbB4)、组织因子、VEGF、VEGFR、内皮素受体、VLA-4、糖类如血型抗原和相关糖类、糖基化 Galili、促胃液素、促胃液素受体、肿瘤相关糖类、半抗原 NP-帽 或 NIP-帽、T 细胞受体 α/β、E-选择蛋白、P-糖蛋白、MRP3、MRP5、谷胱甘肽-S-转移酶 pi(多重耐药蛋白)、α-颗粒膜蛋白(GMP)140、地高辛、胎盘碱性磷酸酶(PLAP)和睾丸 PLAP 样碱性磷酸酶、运铁蛋白受体、肝素酶 I、人心肌球蛋白、糖蛋白 II b/IIIa(GP II b/IIIa)、人巨细胞病毒(HCMV)gH 包膜糖蛋白、HIV IIIB gp120、HCMV、呼吸道合胞体病毒 RSV F、RSV Fgp、VNR 整联蛋白、Hep B gp120、CMV、gp II bIIIa、HIV IIIB gp120 V3 环、呼吸道合胞体病毒(RSV) Fgp、单纯疱疹病毒(HSV)gD 糖蛋白、HSV gB 糖蛋白、HCMV gB 包膜糖蛋白、产气荚膜梭菌肠毒素及它们的片段。

优选抗原选自病原体抗原、肿瘤相关抗原、酶、底物、自身抗原、有机分子或过敏原。更优选抗原选自病毒抗原、细菌抗原、来自真核生物或噬菌体病原体的抗原。优选的病毒抗原包括 HAV-、HBV-、HCV-、HIV I -、HIV II -、细小病毒- (Parvovirus-)、流感病毒-(Influenza-)、HSV-、肝炎病毒(Hepatitis Virus)、黄病毒(Flaviviruses)、西尼罗病毒(West Nile Virus)、埃波拉病毒(Ebola Virus)、痘病毒(Pox-Virus)、天花病毒(Smallpox Virus)、麻疹病毒(Measles Virus)、疱疹病毒(Herpes Virus)、腺病毒(Adenovirus)、乳头瘤病毒(Papilloma Virus)、多瘤病毒(Polyoma Virus)、细小病毒(Parvovirus)、鼻病毒(Rhinovirus)、柯萨奇病毒(Coxsackie virus)、脊髓灰质炎病毒(Polio Virus)、埃柯病毒(Echovirus)、日本脑炎黄病毒(Japanese Encephalitis virus)、登革热病毒(Dengue Virus)、蜱传脑炎病毒(Tick Borne Encephalitis Virus)、黄热病病毒(Yellow Fever Virus)、冠状病毒(Coronavirus)、呼吸道合胞病毒(respiratory syncytial virus)、副流感病

毒(parainfluenza virus)、拉克罗斯病毒(La Crosse Virus)、拉沙病毒(Lassa Virus)、狂犬病毒(Rabies Viruse)、轮状病毒(Rotavirus)的抗原;优选的细菌抗原包括假单胞菌 - (Pseudomonas-)、分枝杆菌 - (Mycobacterium-)、葡萄球菌 - (Staphylococcus-)、沙门氏菌 - (Salmonella-)、脑膜炎球菌 - (Meningococcal-)、疏螺旋体 - (Borellia-)、李斯特菌 - (Listeria)、奈瑟氏球菌 - (Neisseria-)、梭菌 - (Clostridium-)、埃希氏菌 - (Escherichia-)、军团菌 - (Legionella-)、芽胞杆菌 - (Bacillus-)、乳酸杆菌 - (Lactobacillus)、链球菌 - (Streptococcus-)、肠球菌 - (Enterococcus-)、棒状杆菌 - (Corynebacterium-)、诺卡氏菌 - (Nocardia-)、红球菌 - (Rhodococcus-)、莫拉氏菌 - (Moraxella-)、布鲁氏菌 (Brucella)、弯曲杆菌 - (Campylobacter-)、心杆菌 - (Cardiobacterium-)、弗朗西斯氏菌 - (Francisella-)、螺杆菌 - (Helicobacter-)、嗜血杆菌 - (Haemophilus-)、克雷伯氏菌 - (Klebsiella-)、志贺氏菌 - (Shigella-)、耶尔森氏菌 - (Yersinia-)、弧菌 - (Vibrio-)、衣原体 - (Chlamydia-)、钩端螺旋体 - (Leptospira-)、立克次氏体 - (Rickettsia-)、分枝杆菌 - (Mycobacterium-)、密螺旋体 - (Treponema-)、巴尔通氏体 - (Bartonella-) 的抗原。优选的致病真核生物的真核抗原包括来自贾第虫(Giardia)、弓浆虫(Toxoplasma)、环孢子虫(Cyclospora)、隐孢子虫(Cryptosporidium)、毛形线虫(Trichinella)、酵母、假丝酵母(Candida)、曲霉(Aspergillus)、隐球菌(Cryptococcus)、芽生酵母(Blastomyces)、组织胞浆菌(Histoplasma)、球孢子菌(Coccidioides)。

本发明的修饰免疫球蛋白可优选与上面公开的分子之一结合。这些分子还包含过敏原和半抗原。

抗原亚结构通常称为“表位”(例如 B-细胞表位、T-细胞表位),只要它们是免疫学相关的,即也可被天然抗体或单克隆抗体识别。本发明术语“表位”意指对于结合结构域或本发明的免疫球蛋白而言可完全构成特异性结合配偶体或可为特异性结合配偶体一部分的分子结构。

从化学上,表位也可由糖类、肽、脂肪酸、无机物或其衍生物和其任何组合组成。如果表位是肽或多肽,那么通常所述肽中包含至少

3个氨基酸、优选8-50个氨基酸、更优选在大约10-20个之间的氨基酸。肽的长度没有临界上限，其可包含几乎多肽序列的全长。表位可以是线性表位或构象表位。线性表位由多肽链一级序列的单个区段组成。线性表位可以是连续的或重叠的。构象表位由通过折叠多肽形成三级结构而集合的氨基酸组成，在线性序列中氨基酸不必相互邻近。

特别地，表位是诊断相关分子的最小组成部分，即样品中是否存在表位与疾病、或健康状况、或制造过程状态、或环境和食物状况定性或定量相关。表位也可是治疗相关分子的最小组成部分，即可被改变病程的特异性结合结构域靶向的分子。

除了按照本发明工程改造的免疫球蛋白抗原结合位点以外，还可进一步将结合能力引入到可变结构域的结构环区或结构环区之外，例如与其它抗原、小分子、药物或酶、酶的催化部位或酶底物、或与酶底物的过渡态类似物结合的结合能力。

优选结构环上的新抗原结合位点对于未修饰的免疫球蛋白可变结构域而言是外源的。因此本质上不被可变结构域结合的外源靶标如效应分子、血清蛋白或Fc-受体或细胞表面受体，优选作为被本发明免疫球蛋白可变结构域结合的抗原或结合分子。

在上下文中术语“外源的”意指抗原不被免疫球蛋白的特异CDR结合区或其它天然或固有的结合区识别。外源的结合配偶体(但不是免疫球蛋白的天然结合配偶体)因此可被新形成的结构环抗原结合位点结合。这意味着天然结合配偶体(如Fc-受体或免疫系统效应物)不认为被对未修饰的免疫球蛋白而言是外源的抗原结合位点结合。

本文所用术语“特异地结合”或“特异性结合”是指在异种分子群中决定目标相关配体的结合反应。因此在指定的条件下(例如免疫测定条件)，特异的抗体可变结构与其特定的“靶标”结合并且没有与存在于样品中的其它分子大量结合。特异性结合意指结合跟据所选择靶标的特性和高、中、低结合亲和力或亲合力而具有选择性。如果结合常数或结合动力学相差至少10倍则通常可实现选择性的结合。

术语“表达系统”是指在可操作的连接中包含所需的编码序列和调控序列的核酸分子，以致用这些序列转化或转染的宿主能够产生所编码的蛋白质。为了实现有效的转化，表达系统可被包含在载体上，而相关 DNA 随后还可整合到宿主染色体中。或者，表达系统可被用作体外转录/转译。

以下列方式理解本发明的“结构环”或“非 CDR 环”：免疫球蛋白由具有所谓的免疫球蛋白折叠的结构域构成。实质上，反向平行 β 片层被环所连接形成紧缩的反向平行 β 桶。在可变结构域中，结构域的一些环基本上贡献抗体的特异性，即与抗原的结合。这些环叫做 CDR 环。

CDR 环位于邻近 CDR 环的在某些情况下也可称作可变骨架区(称为“VFR”)的 CDR 环区内部。已知 VFR 也有助于形成通常主要由 CDR 环确定的抗体的抗原结合口袋。因此，那些 VFR 被认为是 CDR 环区的组成部分，不适合用于本发明的目的。与这些位于 CDR 环区内部或最接近 CDR 环的 VFR 相反，可变结构域的其他 VFR 特别适合于本发明使用。那是位于 CDR 环区对面的或者在免疫球蛋白可变结构域 C 末端的 VFR 结构环。

在 CDR 环区外的抗体可变结构域的所有环贡献了相当多的分子结构。这些环在本文中定义为结构环或非-CDR 环。

免疫球蛋白氨基酸序列的所有编号是根据 IMGT 编号模式进行编号的 (IMGT, the international ImMunoGeneTics information system@imgt.cines.fr; <http://imgt.cines.fr>; Lefranc 等, 1999, Nucleic Acids Res. 27: 209-212; Ruiz 等, 2000 Nucleic Acids Res. 28: 219-221; Lefranc 等, 2001, Nucleic Acids Res. 29: 207-209; Lefranc 等, 2003, Nucleic Acids Res. 31: 307-310; Lefranc 等, 2005, Dev Comp Immunol 29:185-203)。

根据本发明优选的实施方案，免疫球蛋白可变结构域来源于人、骆驼、兔、鸡、鼠、狗、马、羊或鼠科动物。

因为修饰的免疫球蛋白可用于各种目的，特别是在药物组合物方面，免疫球蛋白优选来源于人、骆驼或鼠科动物。当然，修饰的免疫球蛋白也可以是人源化或嵌合免疫球蛋白。在本发明最优选的实施方案中，修饰的可变结构域来源于人或者是任何物种可变结构域的人源化可变结构域。

人源化免疫球蛋白可变结构域具有与天然的人免疫球蛋白可变结构域序列的至少约 50%氨基酸序列同一性，优选至少约 55%氨基酸序列同一性、更优选至少约 60%氨基酸序列同一性、更优选至少约 65%氨基酸序列同一性、更优选至少约 70%氨基酸序列同一性、更优选至少约 75%氨基酸序列同一性、更优选至少约 80%氨基酸序列同一性、更优选至少约 85%氨基酸序列同一性、更优选至少约 90%氨基酸序列同一性、更优选至少约 95%氨基酸序列同一性。

此外，当将所有表面可及的氨基酸与天然的人免疫球蛋白可变结构域序列的表面可及的氨基酸相比较时，人源化的免疫球蛋白可变结构域有至少约 50%氨基酸序列同一性、优选至少约 55%氨基酸序列同一性、更优选至少约 60%氨基酸序列同一性、更优选至少约 65%氨基酸序列同一性、更优选至少约 70%氨基酸序列同一性、更优选至少约 75%氨基酸序列同一性、更优选至少约 80%氨基酸序列同一性、更优选至少约 85%氨基酸序列同一性、更优选至少约 90%氨基酸序列同一性、更优选至少约 95%氨基酸序列同一性。

优选的同源性或序列同一性特别是指骨架区序列的那些同源性或序列同一性。

本发明优选的免疫球蛋白包含与未修饰结构域有至少 50%同源性的结构域。

术语“同源性”是指多肽在它们的一级、二级、三级结构中相应的位置上有相同的或保守的残基。该术语也延伸到两个或多个编码同源多肽的核苷酸序列。

“同源免疫球蛋白结构域”意指本发明的免疫球蛋白，其具有与如

本文所公开的全长天然序列免疫球蛋白结构域序列或全长免疫球蛋白结构域序列的任何其它片段的至少约 50%氨基酸序列同一性。优选同源免疫球蛋白结构域与如本文所公开的天然免疫球蛋白结构域序列或全长免疫球蛋白结构域序列的任何其它具体确定的片段具有至少约 50%氨基酸序列同一性，优选至少约 55%氨基酸序列同一性，更优选至少约 60%氨基酸序列同一性，更优选至少约 65%氨基酸序列同一性，更优选至少约 70%氨基酸序列同一性，更优选至少约 75%氨基酸序列同一性，更优选至少约 80%氨基酸序列同一性，更优选至少约 85%氨基酸序列同一性，更优选至少约 90%氨基酸序列同一性，更优选至少约 95%氨基酸序列同一性。

在本文中鉴定的关于免疫球蛋白结构域序列的“氨基酸序列同一性百分比(%)”定义为：在进行序列比对、必要时为实现最大序列同一性百分比而引入缺口并且不考虑作为序列同一性组成部分的任何保守取代之后，与特定免疫球蛋白可变结构域序列氨基酸残基具有同一性的、候选序列中的氨基酸残基的百分比。为了确定氨基酸序列百分比同一性而进行的比对可以各种本领域熟练方式用公开可得到的计算机软件如 BLAST、BLAST-2、ALIGN 或 Megalign (DNASTAR) 软件来完成。本领域技术人员可以确定用于测量比对的合适参数，包括对将被比较的序列的全长进行最大比对所需的任何运算法则。

氨基酸序列同一性%数值可通过使用 WU-BLAST-2 计算机程序按照如下所述的来获得(Altschul 等, Methods in Enzymology 266:460-480 (1996))。大多数 WU-BLAST-2 检索参数设置为默认值。那些没有设置为默认值的参数(即可调参数)用下列值进行设定：重叠跨度(overlap span)=1，重叠分数(overlap fraction)=0.125，字域(word threshold) (T)=11，得分矩阵=BLOSUM62。当使用 WU-BLAST-2 时，氨基酸序列同一性%数值通过将以下(a)除以(b)来确定：(a)根据 WU-BLAST-2 测定，具有来源于天然免疫球蛋白可变结构域序列的目标免疫球蛋白可变结构域氨基酸序列与目标比较氨基酸序列(即与可为未

修饰的免疫球蛋白可变结构域的目标免疫球蛋白可变结构域进行比较的序列)之间配对相同的氨基酸残基数; (b) 目标免疫球蛋白可变结构域的非随机化部分的氨基酸残基总数。例如, 在“包含与氨基酸序列 Y 有至少 80%氨基酸序列同一性的氨基酸序列 X 的多肽”的陈述中, 氨基酸序列 A 是目标比较氨基酸序列, 氨基酸序列 B 是目标免疫球蛋白可变结构域的氨基酸序列。

在优选的实施方案中, 本发明的多肽是双特异性抗体、或双特异性单链抗体、或双特异性 Fab 或双特异性 sdAb。进一步优选的是, 所述多肽包含双特异结构域或其部分。具体实例是指 Fab 或 dAb 分子, 其通过位于 CDR 区对面的新结合位点而具有另外的功能性。为 Fab 分子的本发明优选的免疫球蛋白可例如在 CHI 和/或 CL 结构域的 C 末端环侧上含有一个或两个额外的结合位点。为 dAb 分子的本发明优选的免疫球蛋白可例如在 Vh, Vhh 或 Vl 结构域的 C 末端环侧上含有额外的结合位点。通过这些额外的结合位点, 可将额外的功能性如延长的半寿期(例如通过与 FcRn 或血清蛋白如白蛋白或 IgG 的结合)、或效应功能(例如通过与 T 细胞受体、C1q 或 CD64 的结合)加到分子中去。根据这些实例, 将以合适的大小制备专用的 Fab 或 dAb 文库, 使得能够选出针对特定结合配偶体的特异性结合物(binder)。

本发明优选的可变结构域选自 VH、VL, 包括 V κ 和 V λ 、VHH 和其组合。当将修饰引入 VH、V κ 、V λ 或 VHH 的环或环区, 且经修饰的环或环区包含在氨基酸 7-21、氨基酸 25-39、氨基酸 41-81、氨基酸 83-85、氨基酸 89-103 或氨基酸 106-117 内的至少一个修饰时, 结果发现那些修饰是特别有利的。

根据本发明所修饰的来源于人或人源化的免疫球蛋白的结构环或环区或免疫球蛋白可变结构域优选选自包含氨基酸 8-20、氨基酸 44-50、氨基酸 67-76 和氨基酸 89-101, 最优选氨基酸位置 12-17、氨基酸位置 45-50、氨基酸位置 69-75 和氨基酸位置 93-98 的结构环。

在另一个优选的实施方案中, 包含氨基酸 93-98 的结构环或环区

中的修饰与包含氨基酸 8-20 的结构环或环区中的修饰相结合。

以上确定的单独的免疫球蛋白的氨基酸区域指定为适合本发明修饰目的环或环区。优选这些修饰的组合(例如至少两个指定环或环区中的修饰组合)工程改造到本发明的免疫球蛋白中。

优选在包含氨基酸 93-98 的结构环或环区上的修饰与在一个或多个其它结构环上的修饰结合。

在一个优选的实施方案中,在包含氨基酸 93-98 的结构环或环区的修饰与在包含氨基酸 69-75 的结构环区的修饰相结合。

最优选包含氨基酸 93-98、氨基酸 69-75 和氨基酸 8-20 的各个结构环含有至少一个氨基酸修饰。

在另一个优选的实施方案中,包含氨基酸 93-98、氨基酸 69-75、氨基酸 44-50 和氨基酸 8-20 的各个结构环含有至少一个氨基酸修饰。

根据本发明优选的实施方案,鼠科动物来源的免疫球蛋白的结构环或环区或免疫球蛋白可变结构域(例如 VH)包含氨基酸 6-20、氨基酸 44-52、氨基酸 67-76 和氨基酸 92-101。

根据本发明另一个优选的实施方案,骆驼来源的免疫球蛋白的结构环或环区或免疫球蛋白可变结构域(例如 VHH)包含氨基酸 7-18、氨基酸 43-55、氨基酸 68-75 和氨基酸 91-101。

骆驼来源的可变结构域或骆驼来源的人源化变体具有以下优势:它们能够很容易地与其它可变结构域结合,例如与其它骆驼来源的修饰的或天然的 VHH 结合。骆驼来源的 VHH 组合的可能性是多价免疫球蛋白的基础。因此,根据本发明,骆驼来源的经特定修饰的可变结构域是多价组合,优选具有至少 3、更优选具有至少 4 或 5 价或 VHH。

优选将结构环上的新的抗原结合位点引入到免疫球蛋白,该免疫球蛋白由至少有一个核苷酸取代、缺失和/或插入的经选择的核酸所编码。

根据本发明另一个优选的实施方案,在各个至少两个结构环或环区的至少一个核苷酸的修饰导致在由所述核酸编码的免疫球蛋白或免

疫球蛋白可变结构域取代、缺失和/或插入。

免疫球蛋白或抗体可变结构域的至少两个环或环区的修饰可导致两个或更多个氨基酸的取代、缺失和/或插入，优选点突变、整个环的氨基酸改变，更优选至少 2、3、4、5、6、7、8、9、10、11、12、13、14、直到 30 个氨基酸的改变。然而，插入免疫球蛋白或免疫球蛋白可变结构域的结构环或环区的氨基酸最大数目在特定情况下可不超过 30 个，优选 25 个，更优选 20 个氨基酸。

因此，经修饰的序列包含：不包括在结构环的保守区内的氨基酸、为天然存在但对于修饰位点来说是外源的新引入的氨基酸、或天然存在的氨基酸的取代。当外源的氨基酸选自特定类别的氨基酸如带有特定极性或疏水性的氨基酸时，可根据本发明获得在随机化位置上富含特定类别氨基酸的文库。这样的文库也被叫做“集中”库(“focused library”)。

优选通过随机、半随机或特别是定点随机诱变的方法对至少两个环或环区进行突变或修饰。

引入修饰的优选方法是定点随机突变。以此方法用随机生成的插入使环的两个或更多个特定氨基酸残基交换或引入到所述结构环中。或者优选的是使用组合的方法。

在另一个优选的实施方案中，通过随机的、半随机或特别是定点随机诱变方法来突变或修饰免疫球蛋白或免疫球蛋白可变结构域的至少三个结构环或环区。

这些方法可用于在本发明的免疫球蛋白或免疫球蛋白可变结构域的所需位置上进行氨基酸修饰。在这些情况下位置的选择是随机的，或用某种规则对氨基酸进行改变。例如可将某些残基突变成任何氨基酸，而可将其它残基突变成限定组别(restricted set)的氨基酸。这可以逐步的方式通过突变和选择的循环交替或同时来实现。

经随机修饰的核酸分子可包含本文所确定的、编码所有已知的天然存在的氨基酸或其子集(subset)的重复单位。那些含有修饰序列的、

其中特定子集的氨基酸用作修饰目的的文库叫做“集中”库。这种文库的成员在修饰位置上具有这种子集氨基酸的可能性增加，至少比平常高出两倍，优选至少高3倍乃至至少4倍。这种文库还具有有限或较低数量的文库成员，以致实际的文库成员数量达到了理论的文库成员数量。在某些情况下，聚焦文库的文库成员数量不少于理论数量的 10^3 倍，优选不少于 10^2 倍，最优选不少于10倍。

本发明的文库可被设计成包含特定免疫球蛋白形式的文库成员的或由这些文库成员组成的文库。为本发明的目的，那些由特定免疫球蛋白分子形式组成的文库叫做专用文库(dedicated library)。专用文库优选含有占大部分的特定形式，至少50%、优选至少60%、更优选至少70%、更优选至少80%、更优选至少90%、，或基本上由特定抗体形式组成的文库。优选特定抗体形式，以致根据本发明优选的文库选自：VH文库、VHH文库、V_K文库、V_λ文库、Fab文库、CHI/CL文库和CH3文库。特别优选特征为组成分子(composite molecule)内含有超过一个抗体结构域的文库，如IgG文库或Fc文库。其它优选的文库是那些含有T细胞受体，形成T细胞受体的文库。进一步优选的文库是表位文库，其中融合蛋白包含具有表位变体的分子，还能够选择出有相似结合功能但不同功能性的竞争分子。例示性的文库是TNF α 文库，其中由单个基因包(genetic package)展示了TNF α 融合蛋白的三聚体。

然而，插入免疫球蛋白环或环区的氨基酸的最大数目最好不超过30，优选25个，更优选最多20个氨基酸。优选通过本领域已知的和在本专利申请中公开的方法，用所有可能的氨基酸或选出优选的用于随机目的的氨基酸，使氨基酸的取代和插入随机或半随机地发生。

修饰位点可以位于特异的单个结构环或结构环区。环区通常由至少两个、优选至少3个或至少4个彼此邻近的环组成，可通过形成抗原结合位点或抗原结合口袋促进抗原的结合。优选一个或多个修饰位点位于10个氨基酸区域内，更优选在20、30、40、50、60、70、80、

90 直至 100 个氨基酸内，特别在结构环区内形成抗原可空间上接近环区的表面或口袋。

优选将至少一个环或环区进行突变或修饰以产生文库，优选通过随机、半随机或特别是通过定点随机诱变的方法，特别是缺失、交换或将随机生成的插入引入结构环。另外的优选是使用组合的方法。可使用任何已知的诱变方法，其中之一是盒式诱变。这些方法可用来在本发明免疫球蛋白预期的位置上进行氨基酸修饰。在一些情况下，位置是随机选择的，例如，用任何可能的氨基酸或选择优选的氨基酸使环序列随机化，或用简化的规则产生氨基酸变化。例如，优选所有残基可被突变成特定的氨基酸如丙氨酸，称为氨基酸或丙氨酸扫描(scanning)。这样的方法可与使用选择方法的更尖端的工程改造手段相结合，以筛选更高水平序列多样性。

本发明优选的方法是指编码免疫球蛋白、免疫球蛋白结构域或其部分的随机修饰的核酸分子，其包含在具有序列 5' -NNS-3'、5'-NNN-3'、5'-NNB-3' 或 5'-NNK-3' 的结构环编码区内的至少一个核苷酸重复单位。在一些实施方案中，修饰的核酸包含选自 TMT、WMT、BMT、RMC、RMG、MRT、SRC、KMT、RST、YMT、MKC、RSA、RRC、NNK、NNN、NNS 或其任何其组合的核苷密码子(编码依照 IUPAC)。

随机修饰的核酸分子可包含上述确定的编码所有已知天然存在的氨基酸或其子集的重复单位。

核酸分子的修饰可通过将合成的寡核苷酸引入到较大区段的核酸中或通过完整核酸分子的从头合成来完成。如果将要编码氨基酸的子集，那么可用能够减少无义序列组合数量的三核苷酸结构单元来完成核酸的合成(例如 Yanez 等 *Nucleic Acids Res.* (2004) 32:e158; Virnekas 等 *Nucleic Acids Res.* (1994) 22:5600-5607)。

优选被修饰的位置是暴露于表面的氨基酸。可从已知的抗体可变结构域的蛋白质结构和通过这种没有实验上确定的结构可用的氨基

酸序列的相似或同源性，来判断结构环氨基酸的表面暴露。

在本发明优选的实施方案中，引入到至少两个结构环的修饰包含至少 1、2、3、4、5、或 6 个外源氨基酸或在非修饰免疫球蛋白或免疫球蛋白结构域的结构环各自的位点上的非天然存在的氨基酸。

氨基酸的修饰优先可为偏爱(bias)的，以便引入到已知经常参与蛋白质之间相互作用的结构环或环区氨基酸(例如，Lea & Stewart (1995) FASEB J. 9:87- 93; Fellhouse 等 (2006) J. Mol. Biol. 357:100-114; Adib-Conquy 等 (1998) International Immunology 10:341-346; Lo Conte 等 (1999) J. Mol. Biol. 285:2177-2198; Zemlin 等 (2003) J. Mol. Biol. 334:733-749)。

根据发明的一个实施方案，依据发明方法得到的免疫球蛋白用于制备具有展示或编码本发明免疫球蛋白的成员的文库，特别是蛋白质、融合蛋白、细胞的文库、尤其是微生物细胞，像细菌或酵母细胞、噬菌体、病毒、核酸或核糖体文库。

下列说明不仅涉及多肽或蛋白质变体的文库，当然还涉及例如上述的用于表达本发明免疫球蛋白的备选文库。

在一个优选的实施方案中，包含本发明免疫球蛋白或免疫球蛋白可变结构域的多肽变体文库被用作选择的库(pool)，其中所述修饰含有或引入至少一个、更优选至少两个氨基酸/个经修饰的结构环，所述氨基酸来自色氨酸、酪氨酸、苯丙氨酸、组氨酸、异亮氨酸、丝氨酸、甲硫氨酸、丙氨酸和天冬酰胺。

根据本发明结果发现，可以给变体可变结构域多肽提供对于天然多肽而言是外源的特定突变。色氨酸、酪氨酸、苯丙氨酸、组氨酸、异亮氨酸、丝氨酸、甲硫氨酸、丙氨酸和天冬酰胺中的任何一种氨基酸都不存在于人天然免疫球蛋白的结构环中，因此被认为是“外源的”。通过至少一个结构环的修饰并且为了构成结合位点，本发明变体多肽在结构环中可含有至少两个所述的外源氨基酸。

如果修饰的免疫球蛋白或免疫球蛋白可变结构域是来源于人的

或人源化的免疫球蛋白可变结构域,那么优选的修饰是在 12-17、45-50、69-75 和 93-98 的任何一个位置上至少一个酪氨酸的掺入,和/或在 12-17、45-50、69、71-75、93-94 和 96-98 的任何一个位置上至少一个色氨酸的掺入,和/或在 12-17、46、47、49、50、69-74 和 93-98 的任何一个位置上至少一个组氨酸的掺入,和/或在 12-17、45-47、49、50、70-73、75、94-96 和 98 的任何一个位置上至少一个天冬酰胺的掺入,和/或在 12-17、46-50、69-71、73-75、93、95、96 和 98 的任何一个位置上至少一个甲硫氨酸的掺入,和/或 13、71、75、94、95 和 98 的任何一个位置上至少一个丝氨酸的掺入,和/或在 12、14-17、45-50、69、70、72-75、93 和 96-98 的任何一个位置上至少一个异亮氨酸的掺入,和/或在 15、46、48、70-73、75、93、95 和 98 的任何一个位置上至少一个苯丙氨酸的掺入。

根据本发明另一个优选的实施方案,在人或人源化单结构域抗体的 15-17、29-34、85.4-85.3、92-94、97-98 和/或 108-110 位置上的至少两个氨基酸残基被修饰。

可将编码经修饰的免疫球蛋白或免疫球蛋白结构域(并且贯穿整篇说明总是包括:包含修饰的免疫球蛋白可变结构域的免疫球蛋白和免疫球蛋白片段)的核酸分子克隆到宿主细胞,使其表达并测定其结合特异性。用熟知的程序来完成这些实践,在 *Molecular Cloning-A Laboratory Manual*, 3rd Ed. (Maniatis, Cold Spring Harbor Laboratory Press, New York, 2001)和 *Current Protocols in Molecular Biology* (John Wiley & Sons)中描述了可在本发明中找到使用的多种方法。可将编码本发明修饰的免疫球蛋白或免疫球蛋白可变结构域的核酸掺入到表达载体中,以便表达所述的免疫球蛋白。表达载体通常包含与调控或调节序列可操作连接(即被放在与调控或调节序列有功能关系的位置)的免疫球蛋白、选择标记、任何融合配偶体和/或另外元件。可在诱导或引起修饰的免疫球蛋白表达的适当条件下,通过培养用核酸、优选含有编码修饰免疫球蛋白的核酸的表达载体转化的宿

主细胞来产生本发明修饰的免疫球蛋白，。将外源核酸分子引入宿主细胞的方法在本技术领域是为人熟知的，并随所使用的宿主而变化。当然，也可使用适合于修饰免疫球蛋白表达的非细胞或无细胞表达系统。

在本发明优选的实施方案中，将表达后的修饰的免疫球蛋白纯化或分离。可用为本技术领域所熟知的多种方式对修饰的免疫球蛋白进行分离或纯化。标准的纯化方法包括色谱技术、电泳、免疫、沉淀、透析、过滤、浓缩和层析聚焦技术。通常用特殊的融合配偶体可使纯化成为可能。例如，如果使用了 GST 融合，可用谷胱甘肽树脂纯化抗体，如果使用了 His 标签，可用 Ni^{2+} -亲和色谱法纯化抗体，或者如果使用了 flag 标签，可用固定化的抗-flag 抗体纯化抗体。对于合适的纯化技术的一般指南，参见例如 Scopes, "Protein Purification: Principles and Practice", 1994, 第 3 版, Springer-Science and Business Media Inc., NY or Roe, "Protein Purification Techniques : A Practical Approach", 2001, Oxford University Press。当然，在宿主表面特别是在细菌、昆虫细胞或酵母细胞表面或在噬菌体或病毒表面表达本发明修饰的免疫球蛋白也是可能的。

可用多种方法筛选本发明的修饰的免疫球蛋白，包括但不限于那些使用体外测定、体内和基于细胞测定的方法和选择技术。在筛选程序中可利用自动操作和高通量筛选技术。筛选可利用融合配偶体或标记的用途，例如酶、免疫标记、同位素标记或小分子标记如荧光或比色染料或发光分子。

在一个优选的实施方案中，在体外测定中筛选免疫球蛋白的功能和/或生物物理特性。在一个优选的实施方案中，对抗体功能性进行筛选，例如它催化反应的能力或其结合特异性、交叉反应性和/或与其靶标的亲和力。

在另一个优选的实施方案中，可在体内选择有利的修饰免疫球蛋白，例如通过将其引入细胞或有机体来进行。可从体液如血液或淋巴

液中或从特定器官中分离特异性结合的变体，视所需要的修饰结构域的特性而定。

可使用多种检测方法进行测定，包括但不限于产色的、荧光的、发光的或同位素标记物。

如在本领域所熟知的，有多种选择技术可用于鉴定和分离具有某些结合特性和亲和力的蛋白质，包括例如如下面所描述的展示技术如噬菌体展示、核糖体展示、细胞表面展示等等。制备和筛选抗体变体的方法在本领域是为人熟知的。在 *Antibody Engineering*, 由 Duebel & Kontermann 编辑, Springer-Verlag, Heidelberg, 2001; 和 Hayhurst & Georgiou, 2001, *Curr Opin Chem Biol* 5:683-689; Maynard & Georgiou, 2000, *Annu Rev Biomed Eng* 2:339-76 中描述了关于抗体分子生物学、表达、纯化和筛选的一般方法。

如领域中所知的，一些筛选方法选择有利的文库成员。这些方法在本文中称为“选择方法”，这些方法在本发明中在筛选修饰的免疫球蛋白方面得到应用。当用选择方法筛选变体免疫球蛋白可变结构域文库时，只有那些有利的、满足某些选择标准的文库成员被增殖、分离和/或观测到。正如将被意识到的，因为只有最合适的变体才被观测到，所以这样的方法比起通过逐个测定文库成员适合性的方法能够筛选更大的文库。通过任何方法、技术、或共价或非共价连接免疫球蛋白表型与其基因型(即连接抗体的功能与其编码核酸)的融合配偶体，都可使选择得以实现。例如通过使文库成员与噬菌体外壳蛋白(最常使用的是丝状噬菌体基因 III 蛋白，也可使用其它外壳蛋白如蛋白 VIII、蛋白 VII、蛋白 VI 和蛋白 IX)融合，使得噬菌体展示能够作为选择方法使用。以这种方式，对满足某些标准(例如与免疫球蛋白靶标的结合亲和力)的修饰免疫球蛋白的选择或分离也选择或分离编码它的核酸。编码修饰免疫球蛋白的单个基因或多个基因一经分离后马上可进行扩增。可重复进行这一被称作淘选的分选和扩增过程，使文库中有利的抗体可变结构域变体得以富集。对附着的核酸进行核酸测序最终允许

基因鉴定。

可在本发明发现的、用于筛选免疫球蛋白或免疫球蛋白可变结构域文库的多种选择方法为本领域已知的。这些方法包括但不限于噬菌体展示(Phage display of peptides and antibodies: a laboratory manual, Kay 等, 1996, Academic Press, San Diego, Calif., 1996; Lowman 等, 1991, *Biochemistry* 30:10832-10838; Smith, 1985, *Science* 228:1315-1317)和其衍生方法例如选择性噬菌体感染(Malmborg 等, 1997, *J Mol Biol* 273:544-551)、选择性感染的噬菌体(Krebber 等, 1997, *J Mol Biol* 268:619-630)和延迟感染性淘选(panning) (Benhar 等, 2000, *J Mol Biol* 301:893-904), 细胞表面展示(Wittrup, 2001, *Curr Opin Biotechnol*, 12:395-399)例如细菌展示(Georgiou 等, 1997, *Nat Biotechnol* 15:29-34; Georgiou 等, 1993, *Trends Biotechnol* 11:6-10; Lee 等, 2000, *Nat Biotechnol* 18:645-648; Jun 等, 1998, *Nat Biotechnol* 16:576-80)、酵母展示(Boder & Wittrup, 2000, *Methods Enzymol* 328:430-44; Boder & Wittrup, 1997, *Nat Biotechnol* 15:553-557)和哺乳动物细胞展示(Whitehorn 等, 1995, *Bio/technology* 13:1215-1219), 以及体外展示技术(Amstutz 等, 2001, *Curr Opin Biotechnol* 12:400-405)如多核糖体展示(Mattheakis 等, 1994, *Proc Natl Acad Sci USA* 91:9022-9026)、核糖体展示(Hanes 等, 1997, *Proc Natl Acad Sci USA* 94:4937-4942)、mRNA 展示(Roberts & Szostak, 1997, *Proc Natl Acad Sci USA* 94:12297-12302; Nemoto 等, 1997, *FEBS Lett* 414:405-408)和核糖体灭活展示系统(Zhou 等, 2002, *J Am Chem Soc* 124, 538-543)。

在本发明发现使用的其它选择方法包括不依靠展示的方法, 例如体内方法包括但不限于: 周质表达和细胞计量筛选(Chen 等, 2001, *Nat Biotechnol* 19:537-542)、抗体片段互补测定(Johnsson & Varshavsky, 1994, *Proc Natl Acad Sci USA* 91:10340-10344; Pelletier 等, 1998, *Proc Natl Acad Sci USA* 95:12141-12146)和酵母双杂交筛选(Fields & Song, 1989, *Nature* 340:245-246)及以选择模式被使用的酵母双杂交筛选

(Visintin et al., 1999, Proc Natl Acad Sci USA 96:11723-11728)。在一个备选的实施案中，与表达载体上特定序列结合的融合配偶体使选择得以进行，因此共价或非共价地将融合配偶体和相关的免疫球蛋白文库成员与其编码核酸连接。例如，WO9308278 描述了在本发明中可发现使用的这样的融合配偶体和技术。在一个备选的实施案中，如果抗体的表达赋予细胞某些生长、繁殖或存活优势，那么可在体内进行选择。

一些选择方法称为“定向进化”方法。那些方法包括在选择过程中对有利序列的进行配对(mating)或选育，有时掺入新的突变。正如为本领域技术人员所了解的那样，定向进化方法能促进在大多数多肽中最有利序列的鉴定，能增加被筛选序列的多样性。为本领域所了解的多种定向进化方法可在本发明中发现用于生成和筛选抗体可变结构域变体，包括但不限于DNA改组(PCT WO00/42561; PCT WO 01/70947)、外显子改组(Kolkman & Stemmer, 2001, Nat Biotechnol 19:423-428)、家族改组(Cramer 等, 1998, Nature 391:288-291)、选择性组合随机化(WO03012100, WO04018674A1)、瞬时模板的随机嵌合方法(Coco 等, 2001, Nat Biotechnol 19:354-359)、由在体外重组的交错延伸程序(StEP)进行的分子进化(Zhao 等, 1998, Nat Biotechnol 16:258- 261; Shao 等, 1998, Nucleic Acids Res 26:681-683)、核酸外切酶介导的基因装配(美国专利第 6,352,842 号; 美国专利第 6,361,974 号)、基因位点饱和诱变(美国专利第 6,358,709 号)、基因重新装配(美国专利第 6,358,709 号)、SCRATCHY(Lutz 等, 2001, Proc Natl Acad Sci USA 98:11248- 11253)、DNA 片段化方法(Kikuchi 等, Gene 236:159- 167)、单链 DNA 改组(Kikuchi 等, 2000, Gene 243:133-137)、定向进化抗体工程技术(应用分子进化)(美国专利第 5,824,514 号; 美国专利第 5,817,483 号; 美国专利第 5,814,476 号; 美国专利第 5,763,192 号; 美国专利第 5,723,323 号)。

在一个优选的实施案中，用一种或多种基于细胞测定或体内测

定筛选免疫球蛋白或抗体可变结构域变体。对于这些测定，通常将纯化的或未纯化的修饰免疫球蛋白或免疫球蛋白可变结构域从外部加入，以使细胞暴露于单独的免疫球蛋白或修饰的免疫球蛋白可变结构域、或属于文库的修饰免疫球蛋白可变结构域的库。这些测定通常但不总是基于免疫球蛋白或免疫球蛋白可变结构域的所需功能，即根据本发明修饰的抗体或抗体可变结构域与其靶标结合的能力、以及介导某些生化事件(例如效应功能、配体/受体结合抑制、凋亡等等)的能力。这样的测定经常涉及监测细胞对抗体可变结构域的反应，例如细胞存活、细胞死亡、细胞形态变化、或转录激活例如天然基因或报告基因的细胞表达。例如，这样的测定可测量可变免疫球蛋白结构域变体诱导 ADCC、ADCP 或 CDC 的能力。对于一些测定，额外的细胞或组分(即除了靶细胞之外的细胞和组分)可能需要加入，例如血清补体、或效应细胞例如外周血单核细胞(PBMC)、NK 细胞、巨噬细胞等等。这样的额外细胞可来自任何生物体，优选人、小鼠、大鼠、兔和猴。免疫球蛋白可引起某些表达靶标的细胞系的凋亡，或者它们可通过已被加入测定的免疫细胞介导对靶细胞的攻击。在本技术领域已知的监测细胞死亡或生存力的方法包括使用染料、免疫化学的、细胞化学的和放射性的试剂。例如，半胱氨酸蛋白水解酶染色试验能够测量凋亡，放射性底物或荧光染料的吸收或释放能够监测细胞生长或活化。

可供选择的是，可通过测量一种或多种天然胞内组分例如乳酸脱氢酶的释放，监测死亡或受损的靶细胞。转录激活也可作为基于细胞的测定中测定功能的方法。在这种情况下，可通过测定可能被上调的天然基因来对反应进行监测，例如可测量某些白细胞介素的释放，或者可另外借助报告系统读出。基于细胞的测定也可包括测量作为对修饰免疫球蛋白可变结构域的存在作出反应的细胞形态变化。用于这种测定的细胞类型可以是原核的或真核的，可使用本技术领域所熟知的多种细胞系。

可供选择的是，可使用经变体免疫球蛋白可变结构域的编码核酸

转化或转染的细胞，来完成基于细胞的筛选。在这种情况下，不将本发明抗体可变结构域变体从外部加入所述细胞中(例如 Auf der Maur, 2004, *Methods*, 34:215-224)。在另一个备选的方法中，基于细胞的筛选利用细胞表面展示。可使用能在细胞表面展示修饰的免疫球蛋白可变结构域的融合配偶体(Wittrup, 2001, *Curr Opin Biotechnol*, 12:395-399)。

在一个优选的实施方案中，可用一种或多种免疫学的或基于细胞的检测，在实验上改变和确定修饰的免疫球蛋白的免疫原性(例如 Koren 等, 2002, *Current Pharmaceutical Biotechnology* 3:349-360; Chirino 等, 2004, *Drug Discovery Today* 9:82-90; Tangri 等, 2005, *J. Immunol.* 174:3187-3196; Hermeling 等, 2004, *Pharm. Res.* 21:897-903)。在一个优选的实施方案中，可利用离体 T-细胞活化测定在实验上对免疫原性进行定量。在该方法中，用目标肽或完整抗体或免疫球蛋白对来自匹配供体的抗原呈递细胞和幼稚 T-细胞进行一次或多次攻击。可用很多方法检测 T-细胞活化，例如通过监测细胞因子释放或测量氘-胸腺嘧啶脱氧核苷吸收。在优选的实施方案中，LUMINEX 技术用于测量细胞因子释放(例如 de Jager 等, *Clin. Diagn. Lab. Immunol.*, 2003, 10:133-139)或者用 Elispot 测定法监测 γ 干扰素生产(Schmittel 等, 2000, *J. Immunol. Meth.*, 24: 17-24)。

本发明修饰的免疫球蛋白或免疫球蛋白可变结构域的生物学特性或功能特性可在细胞、组织和完整生物体实验中表现出来。正如在本领域已知的，为了测量药物治疗疾病或疾病模型的疗效或者测量药物的药物动力学、毒性和其它特性，经常用包括但不限于小鼠、大鼠、兔、狗、猫、猪和猴的动物进行药物试验。这些动物可称为疾病模型。治疗剂经常在包括但不限于裸鼠、SCID 小鼠、异种移植小鼠和转基因小鼠(包括基因敲入和敲除突变体)的小类中进行试验。这样的实验方法可为确定多肽变体用作治疗剂的潜力提供有意义的的数据。任何生物体，优选哺乳动物可用于试验。由于与人类具有遗传相似性，猴可

能是合适的治疗模型，因此可被用来试验本发明修饰的免疫球蛋白的疗效、毒性、药物动力学或其它特性。在人体中试验的候选药物通常需要通过作为治疗剂的审批，因此当然这些实验是被考虑在内的。因此可在人体内试验本发明修饰的免疫球蛋白或免疫球蛋白可变结构域以确定它们的治效、毒性、免疫原性、药物动力学、药效学和/或其它临床特性。

本发明的免疫球蛋白可用于在免疫球蛋白技术领域已知的任何目的，也允许依赖于由本发明引入的特异性组合的应用。

在一个实施方案中，本发明的抗体变体作为诊断、工业化合物或研究试剂、优选治疗剂用于治疗或预防、制备或分析的用途。抗体变体可在单克隆、寡克隆或多克隆抗体组成中得到应用。在优选的实施方案中，本发明修饰的免疫球蛋白或免疫球蛋白可变结构域用于杀死荷有靶抗原的靶细胞，例如癌细胞。在一个备选的实施方案中，本发明的修饰的免疫球蛋白被用来阻断、拮抗或对抗靶抗原，例如通过拮抗细胞因子或细胞因子受体。在另一个优选的实施方案中，本发明的修饰免疫球蛋白被用来阻断、拮抗或对抗靶抗原并杀死荷有靶抗原的靶细胞。

在备选的优选实施方案中，本发明的修饰的免疫球蛋白被用来阻断、拮抗或对抗生长因子或生长因子受体并杀死荷有或需要靶抗原的靶细胞。在备选的优选实施方案中，本发明的修饰的免疫球蛋白被用来阻断、拮抗或对抗酶或酶的底物。在另一个备选的优选实施方案中，本发明的修饰的免疫球蛋白可变结构域被用来中和传染性病原体如病毒、小病毒、朊病毒、细菌或真菌。

本发明的修饰的免疫球蛋白或免疫球蛋白可变结构域可用于各种治疗目的。在一个优选的实施方案中，将包含修饰的免疫球蛋白或免疫球蛋白可变结构域的抗体给予患者以治疗特定的病症。本发明目的“患者”包括人类和其它动物，优选哺乳动物，最优选人。本文“特定的病症”意指可通过给予包含本发明修饰的免疫球蛋白或免疫球蛋

白可变结构域药物组合物而得以改善的病症。

在一个实施方案中，本发明的修饰的免疫球蛋白或免疫球蛋白可变结构域是给予患者的唯一的治疗活性剂。或者，本发明的修饰的免疫球蛋白或免疫球蛋白可变结构域与一种或多种其它的治疗剂联合给药，这些治疗剂包括但不限于细胞毒剂、化疗药物、细胞因子、生长抑制剂、抗激素剂、激酶抑制剂、抗血管形成剂、保心药或其它治疗剂。修饰的免疫球蛋白或免疫球蛋白可变结构域可伴随一种或多种其它治疗方案进行给药。例如，可配制本发明的抗体变体制剂并连同化学疗法、放射疗法、或化学疗法和放射疗法一起给予患者。在一个实施方案中，本发明的修饰的免疫球蛋白或免疫球蛋白可变结构域可与一种或多种可包含或者不包含本发明的抗体变体的抗体联合给药。依照本发明的另一个实施方案，使用本发明的修饰的免疫球蛋白或免疫球蛋白可变结构域和一种或多种其它的抗癌疗法来离体治疗癌细胞。考虑的是这样的离体治疗在骨髓移植特别是自体骨髓移植中是有用的。当然考虑本发明的抗体可与其它治疗技术如外科学结合使用。

多种其它的治疗剂可用于与本发明修饰的免疫球蛋白一起给药。在一个实施方案中，修饰的免疫球蛋白与在某种程度上阻断或干扰血管发展的抗血管形成剂一起给药。所述抗血管形成因子可例如为结合到涉及促进血管生成的生长因子或生长因子受体的小分子或蛋白质，例如抗体、Fc 融合蛋白、或细胞因子。本文优选的抗血管生成因子是结合到血管内皮生长因子(VEGF)的抗体。在一个备选的实施方案中，修饰的免疫球蛋白与诱导和加强获得性免疫应答的治疗剂例如靶向 CTLA-4 的抗体一起给药。在一个备选的实施方案中，所述修饰的免疫球蛋白与在一定程度上抑制酪氨酸激酶活性的酪氨酸激酶抑制剂一起给药。在一个备选的实施方案中，本发明的修饰的免疫球蛋白与细胞因子一起给药。本文所使用的“细胞因子”意指由一个细胞群释放的、对另一种细胞起胞间介质包括趋化因子作用的蛋白质的通称。

考虑药物组合物，其中本发明的修饰的免疫球蛋白和一种或多种

治疗活性剂进行配制。通过将具有所需纯度的所述修饰的免疫球蛋白或免疫球蛋白可变结构域与任选的可药用载体、赋形剂或稳定剂 (Remington's Pharmaceutical Sciences, 1980, 16th edition, Osol, A. Ed.) 混合, 以冻干制剂或水溶液的形式来制备本发明多肽变体制剂, 用于贮存。用于体内给药的制剂优选是无菌的。这通过经无菌过滤膜的过滤作用或其它方法容易地完成。所述修饰的免疫球蛋白和本文公开的其它治疗活性剂也可配制成免疫脂质体和/或包埋在微囊内。

包含本发明修饰的免疫球蛋白或免疫球蛋白可变结构域的药物组合物的给药, 优选以无菌水溶液的形式, 可以多种方式实施, 包括但不限于口服给药、皮下给药、静脉内给药、鼻内给药、intraotically、透皮给药、局部给药(例如, 凝胶剂、药膏、洗剂、面霜等等)、腹膜内给药、肌内给药、肺内给药、阴道给药、胃肠外给药、直肠给药、或眼内给药。

本发明另一个方面涉及制备由免疫球蛋白或免疫球蛋白可变结构域组成的分子或其药物制剂、以及测定所述分子与抗原表位的结合情况的方法, 所述分子在所述免疫球蛋白或免疫球蛋白可变结构域的两个结构环或环区的每个上包含至少一个修饰, 其中未修饰的分子明显不与所述的表位结合, 所述方法包含以下步骤:

- 提供编码包含至少两个结构环或环区的免疫球蛋白可变结构域的核酸,
- 在每个所述的结构环或环区修饰至少一个核苷酸残基,
- 将所述经修饰的核酸转移到表达系统中,
- 表达所述经修饰的免疫球蛋白,
- 使所表达的修饰的免疫球蛋白与表位接触,
- 确定所述经修饰的免疫球蛋白是否与所述表位结合, 和
- 提供与所述表位结合的修饰的免疫球蛋白并任选将其制备成药物制剂。

具体而言, 本发明涉及一种方法, 所述方法用于制备与至少一个

第一分子特异性结合的多特异性分子或其药物制剂，所述特异性分子在免疫球蛋白或免疫球蛋白可变结构域的至少两个结构环或环区中的每个上包含至少一个修饰；和测定所述至少两个环或环区与至少一个选自抗原的第二分子的特异性结合情况。含有未修饰的结构环或环区的免疫球蛋白或免疫球蛋白结构域不与所述至少一个第二分子特异性结合。

所述方法明确包含以下步骤：

-提供编码与至少一个第一分子特异性结合的、包含至少一个结构环或环区的免疫球蛋白的核酸，

-修饰由所述核酸编码的至少一个所述环或环区的至少一个核苷酸残基，

-将所述经修饰的核酸转移到表达系统中，

-表达所述经修饰的免疫球蛋白，

-将所表达的修饰的免疫球蛋白与所述的至少一个第二分子接触，
和

-确定所述经修饰的免疫球蛋白是否与所述第二分子特异性结合
并且

-提供与至少一个第二分子特异性结合的经修饰的免疫球蛋白并
任选将其制备成药物制剂。

优选将多于一个结合位点或至少两种特异性工程改造到特异性结合的配对的成员中(Kufer等(2004) Trends in Biotechnology 第22卷第238-244页)。

已经做出很多尝试以产生多特异性，例如双特异性、单克隆抗体或抗体片段。制备由两组不同多肽链(重链和轻链)构成的双特异抗体的一个问题是必须在一个细胞中表达4条不同的链(2条重链和2条轻链)，这致使产生许多不同的分子组合，它们必须与混合物中的所需的双特异性分子分离。由于它们具有相似性，所以这些分子的分离既困难又昂贵。已经使用了很多技术使这种不需要的配对的出现最小化

(Carter (2001) *Journal of Immunological Methods*, 第 248 卷, 第 7-15 页)。

对该问题的一个解决办法是制备具有两种特异性的一条多肽链像例如互相连接两条的 scFv 或制备所谓的双抗体。这样的分子已经表明难以折叠成天然分子, 且众所周知难以制备(LeGall 等. (2004) *Protein Engineering, Design & Selection* 第 17 卷 第 357- 366 页)

根据本发明优选的实施方案, 所述表达系统包含载体。本领域已知的任何表达载体可适当地用于此目的。

修饰的免疫球蛋白优选在宿主中表达, 优选在细菌、酵母、植物细胞、昆虫细胞、动物细胞或哺乳动物细胞或动植物器官或完整的动物或植物中表达。

各种各样合适的宿主细胞可用于表达本发明修饰的多肽, 这些宿主细胞包括但不限于哺乳动物细胞(动物细胞)、植物细胞、细菌(例如枯草芽孢杆菌(*Bacillus subtilis*)、大肠杆菌(*Escherichia coli*))、昆虫细胞和酵母(例如巴斯德毕赤酵母(*Pichia pastoris*), 酿酒酵母(*Saccharomyces cerevisiae*))。例如, 在由美国典型培养物保藏中心提供的 ATCC 细胞系目录中描述了很多可被用于本发明的细胞系。此外, 植物和动物也可用于本发明免疫球蛋白的表达宿主。根据所使用的宿主可选择表达和转染载体或盒。

当然还可使用非细胞或无细胞蛋白质表达系统。可生成足够量蛋白质的体外转录/转译蛋白质表达平台, 提供无细胞蛋白质表达系统的很多优势, 排除通常与基于细胞的表达系统相关的繁琐的上下游步骤(例如宿主细胞转化、培养或溶解)。

双特异抗体当前设计的另一个问题是: 虽然亲本抗体与它们各自的结合配偶体二价结合(例如 IgG), 但是产生的双特异抗体对于各自的结合配偶体中的一每种都是单价的。

本发明优选的多特异分子解决了这些问题:

使双特异分子作为一条多肽链表达是可能的(具有两种结合特异性的修饰的免疫球蛋白可变结构域, 见实施例部分), 这比表达两条抗

体多肽链更容易完成(Cabilly 等 Proc. Natl. Acad. Sci. USA 81:3273-3277 (1984))。

也可将其制成抗体样分子(即由两条多肽链构成)。由于事实上第二特异性位于可变结构域的非 CDR 部分, 所以不需要两条不同的重链或者不同的轻链。因此, 不存在两条链错配的可能性。

本发明抗体可由共同形成与特异性结合配偶体结合的可变 CDR 环区(即特异 CDR 环构象)的重链和轻链组成, 而第二特异性可通过包含在免疫球蛋白分子内的经修饰的结构环或环区(例如重链或轻链可变结构域的结构环)、同时保持特异的 CDR 环构象来形成。结合位点还可通过在两个可变结构域(例如可以是结构上相邻的重链可变结构域和轻链可变结构域)上的至少一个或多于一个非 CDR 环形成。

修饰的抗体可以是完全抗体或抗体片段(例如 Fab、scFv、Fv、小抗体、dAb)或其衍生物, 所述抗体片段包含至少一个免疫球蛋白可变结构域和在结构环或环区上的修饰。

所述修饰的抗体可与结合配偶体单价或多价结合乃至以不同价对应不同结合配偶体, 视设计而定。例如, 可以这种方式工程改造 Fab 片段或者等价的 scFv: 分别对 VH 和 VL 结构域的结构环进行工程改造, 使其结合与由 CDR 形成的结合位点所结合的相同的表位, 从而分别产生三价的 Fab 片段或 scFv。在另一个实施方案中, 含有同样的工程改造 VH 和 VL 结构域的完整免疫球蛋白六价地与其靶表位结合。如果例如由 CDR 形成的天然结合位点识别与工程改造的 Vh 和 Vl 结构域所识别的不同的靶表位, 那么产生的 Fab 片段或 scFv 将会与第一靶标单价结合, 与第二靶标二价结合, 所述第二靶标被 VH 和 VL 结构域的修饰结构环分别独立地结合。可以多种不同的方式应用该标准设计原理对于本领域技术人员而言是显而易见的。

因为有很多不同结构环可用于选择和设计重链和轻链结构域的非 CDR 区域的特异性结合位点, 所以设计具有甚至多于两种特异性的抗体衍生物是可能的。例如, 可分别工程改造通过其 CDR 识别第

一靶标的 VH 和 VL 结构域，以使它们通过修饰的结构环所介导的相互作用特异地与不同的(第二和第三)靶标结合。从而可产生与其靶标中的每一种单价结合的三特异 Fab 片段或 scFv。如果以完整大小 IgG 的形式工程改造该 Fab 的修饰的可变结构域，那么将产生三特异的且每种三特异性两价地结合的工程改造 IgG。

在一条多肽链内部的特异性结合结构域可用或不用肽接头进行连接。

一些抗体类别本性上可被视为多特异的，特别是双特异的：它们用可变结构域与抗原(例如通常是外源结构或癌症相关的结构)结合，而用 Fc 部分与 Fc 效应分子(例如在各种免疫细胞上的 Fc 受体或补体蛋白)结合从而激活如 ADCC、ADCP 或 CDC 的效应。

Fc 效应分子可被免疫球蛋白分子的 Fc 部分(对于 IgG1 而言它由结构域 CH2 和 CH3 组成)结合，已经描述了很多方法优化效应功能，通过糖基化工程改造技术(US 6,602,684)、或通过直接在 Fc 上(US 2005/0054832)或间接在 Fc 外部进行工程改造(US 2005/02444403)的蛋白质工程。这样的技术已经改变了 Fc 区域与 Fc 受体的结合和/或与补体蛋白如 Cq1 的结合。通常人们尝试去提高与这样的 Fc 效应分子的结合亲合力，因为这与效应功能的提高相关联。

以当前的发明，设计在天然 Fc 结合区域外与 Fc 效应分子结合的抗体是可能的。可从经修饰的环结构文库中选择抗体可变结构域的经修饰的环(而非参与与“天然”Fc 效应分子结合的环)或将其设计成与一个或多个 Fc 效应分子结合。具有这种另外的 Fc 效应分子结合位点的抗体将会对某些 Fc 效应分子或展示 Fc 效应分子的效应细胞具有更强的亲和性，因而比糖基化工程抗体或者其他改良的 Fc 区域具更强的效应。

与完整抗体相比，抗体片段具有某些优势。片段通常具有好的生物分布特性，且更易制备。然而，大多数抗体片段设计缺乏效应功能且体内半寿期短(Holliger P,等 Nat Biotechnol . (2005) 23:1126-36)。

CH1、CK 和 C λ 结构域都不能介导效应功能，这就是为什么 Fab 分子通常不表现 ADCC、ADCP 或 CDC 的原因。

WO 02/44215 描述了由抗体的抗原结合位点和结合 Fc 效应分子的肽组成的结合分子。以这种方式可构建出展示效应功能的抗体片段。所述肽被掺入到所述结合分子的位置上既不损害抗原结合又不破坏所述肽与 Fc 效应分子结合能力。

然而根据本发明，经修饰的免疫球蛋白或免疫球蛋白可变结构域可用作实现与 Fc 效应分子的结合，为此根据与 Fc 效应分子的结合情况，已经对其进行从免疫球蛋白或免疫球蛋白可变结构域的固定支架内的两个、三个或四个随机化的结构环序列文库中选择。因此，选出如果从 Ig 结构域支架中分离则不与 Fc 效应分子结合的特异环序列是可能的。因此本发明所产生的多肽可优选由多于 100 个氨基酸组成并可包含一个或多个免疫球蛋白可变结构域。

为了选择本发明的这种可变结构域的潜在效应功能，可根据与 Fc 受体和/或补体因子如 C1q 结合情况，来对包含突变可变结构域的抗体或抗体片段文库进行选择。可在天然表达相应受体的细胞表面或者通过相应受体胞外部分的表达和纯化，提供用于选择的 Fc γ 受体。经 IFN-g 刺激的 U937 细胞(CRL-1503, 美国典型培养物保藏中心)可被用作靶细胞以分离经噬菌体展示的与高亲和力 IgG 受体、Fc γ RI 特异性结合的修饰免疫球蛋白可变结构域(Berntzen 等, 2006, Protein Eng Des Sel. 19(3): 121-8)。与 Fc 受体的结合情况可用 U937 细胞作为靶细胞通过 FACS 进行检验，细胞可用选择的修饰免疫球蛋白可变结构域特异地染色。此外，可将人 Fc γ 受体的胞外结构域进行克隆、作为可溶性蛋白或融合蛋白表达并用于分析潜在结合配偶体的特异性结合情况(例如在 Berntzen 等, 2005, J Immunol Methods. 298 (1-2): 93-104)。可基本上类似地实施对修饰的免疫球蛋白可变结构域与补体因子 C1q 特异性结合的鉴定和鉴别(例如在 Lauvrak 等 1997 Biol Chem. 378 (12): 1509-19)。

为了增加包含这种可变结构域或由其组成的分子的体内半衰期，可根据与 FcRn 或血清白蛋白的结合情况，对本发明的突变可变结构域文库进行选择。为了与将被设计成体内半寿期增加的分子的那些分子结合，那些负责通过与血清蛋白或补体蛋白特异性结合来延长分子的半寿期的修饰的结构环可用作单独的结构环或用在免疫球蛋白或其部分的前后序列中。

可在天然表达相应受体的细胞表面或者通过表达和纯化相应受体的胞外部分，来提供用于选择的 FcRn 受体或其它细胞受体。为本发明的目的，针对 FcRn 的第一轮筛选可选出突变的可变结构域(或包含这种可变结构域分子)，可进一步对其进行体外试验和更进一步在 FACS 实验中通过与表达 FcRn 受体的细胞结合情况进行鉴定。筛选和选择也可考虑与 FcRn 结合的 pH 依赖性(如在 PCT WO02/060919; PCT WO97/34631 中所述)。可通过与各种重组 FcRn、同种型和同种异型结合的亲和力分级用例如表面等离子共振技术进一步进行表征(例如在 Dall' Acqua 等 *Journal of Immunology*, 2002, 169: 5171-5180)。

本发明的修饰的免疫球蛋白可包含重链或轻链或其组成部分和至少一个可变结构域。

本发明的免疫球蛋白优选包含至少一个免疫球蛋白恒定结构域和/或至少一个可变结构域或其组成部分。

可变结构域通常被看作免疫球蛋白可变部分的免疫球蛋白折叠单位，也称为可变区的结构域(例如 VH、Vk、Vl、Vd)。

本发明另一优选的免疫球蛋白由具有至少两个结构环或环区的重链或轻链可变结构域或其部分组成，其特征是：所述的至少两个修饰结构环或环区包含至少两个氨基酸修饰而形成至少两个修饰的结构环或环区，其中所述至少两个修饰结构环或环区与至少一个抗原表位特异地结合。在这样一个优选的本发明免疫球蛋白中，所述的至少两个氨基酸修饰可位于一个或两个结构环或环区中或者位于一个或两个结构环上，以构成抗原结合位点。

根据本发明优选的实施方案，修饰的多肽与分子的特异性结合是由选自免疫学测定、优选酶联免疫吸附测定法(ELISA)、表面等离子共振分析、饱和转移差异核磁共振波谱法、转移 NOE (trNOE)核磁共振波谱法、竞争测定法、组织结合测定、活细胞结合分析和细胞提取测定(cellular extract assay)的结合分析所决定的。

可用多种本领域已知的方法进行结合测定，包括但不限于以 FRET (荧光共振能量转移) 和 BRET (生物发光共振能量转移) 为基础的测定法、放大发光临近同质分析(Amplified Luminescent Proximity Homogeneous Assay)、闪烁亲近测定法、ELISA (酶联免疫吸附测定)、SPR (表面等离子共振)、等温滴定量热法、差示扫描量热法、凝胶电泳和包括凝胶过滤的色谱法。

本发明的修饰的多肽优选与标记缀合，所述标记选自有机分子、酶标记、放射性标记、颜色标记、荧光标记、产色标记、发光标记、半抗原、地高辛配基、生物素、金属络合物、金属、胶体金以及它们的混合物。

与上述标记缀合的修饰多肽可用在例如诊断方法中。

修饰的免疫球蛋白可与允许在例如结合测定(例如 ELISA)和结合研究中对所述缀合物进行简单检测的其它分子缀合。

本发明另一方面涉及具有至少两个环或环区的、包含轻链或重链或其组合的可变结构域的多肽，其特征是：所述至少两个结构环或环区中的每个包含至少一个氨基酸修饰而形成至少两个修饰的结构环或环区，其中所述的至少两个修饰结构环或环区与至少一个抗原表位特异地结合。

优选使至少一个修饰的抗体可变结构域与至少一个其它的结合分子进行分子结合(combine molecularly) (=通过非可变序列或结构环与特异配偶体结合)，所述结合分子可以是抗体、抗体片段、可溶性受体、配体或另外的修饰抗体结构域。

与本发明的至少一个修饰的抗体可变结构域结合的其它结合分

子选自蛋白质分子、核酸和糖类。

根据发明的修饰免疫蛋白的结构环或环区可与任何种类的结合分子特异性结合，特别是蛋白质分子、蛋白、肽、多肽、核酸、多糖、糖类、脂类、小和大的有机分子、无机分子。当然，本发明的修饰的免疫球蛋白可包含至少两个环或环区，而每一个环或环区可特异地与不同的分子或表位结合。

本发明的另一个方面涉及本发明的或通过本发明方法可得到的免疫球蛋白或免疫球蛋白可变结构域用于制备自动免疫疫苗的应用。因此，所述免疫球蛋白被用作抗原药物制备疫苗或者用于诱捕(fishing)或捕获抗原结构而在疫苗配制中使用。

本发明的另一个方面涉及本发明的或者通过本发明可得到的免疫球蛋白用于制备包含修饰免疫球蛋白可变结构域的多肽文库的用途。

本发明的另一个方面更涉及特异性结合和/或检测靶分子的方法，包含下列步骤：

(a)将包含根据本发明的修饰免疫球蛋白或免疫球蛋白可变结构域分子或包含通过本发明的方法可获得的修饰免疫球蛋白可变结构域分子与含有或怀疑含有所述靶分子的试样接触，并任选

(b)检测特异免疫球蛋白/分子或免疫球蛋白可变结构域/分子复合体的潜在形成情况。

本发明优选的方法用于特异地结合和/或检测分子，包含下列步骤：

(a)将本发明的修饰免疫球蛋白(modified immunoglobulins or a modified immunoglobulin)文库与含有所述分子的试样接触，并任选

(b)检测特异免疫球蛋白/分子复合体的潜在形成情况。

试样可以使人或动物样品，如血样或其它体液和细胞悬浮液，这些样品可能含有与用于捕获和/或检测的免疫球蛋白特异性结合的靶分子。

本发明另一方面涉及特异地分离靶分子的方法，包含步骤：

- (a)将包含根据本发明修饰的免疫球蛋白或免疫球蛋白可变结构域分子或从通过本发明的方法可获得的包含修饰免疫球蛋白可变结构域分子与含有所述靶分子的样品接触，
- (b)分离所形成的特异免疫球蛋白可变结构域/靶分子复合物，并且
- (c)从所述复合物随意分离靶分子。

根据本发明优选的方法用于特异地分离与分子结合的修饰免疫球蛋白，包含步骤：

- (a)将本发明修饰的免疫球蛋白文库与含有所述分子的样品接触，
- (b)分离所形成的特异的修饰免疫球蛋白/分子复合物，并且
- (c)任选从所述复合物分离修饰的免疫球蛋白。

这些样品通常被看作是用于分离制备那些分子的来源，例如复杂的天然来源像动物、人或植物来源或微生物来源或细胞悬液和培养物。

根据本发明的免疫球蛋白或免疫球蛋白可变结构域可被用于从样品中特异地分离靶分子。如果使用多特异免疫球蛋白或免疫球蛋白可变结构域，那么可从样品中分离出超过一种靶分子。在这些方法中使用免疫球蛋白或免疫球蛋白可变结构域特别具有优势，因为它允许例如生成具有均一表面(homogeneous surface)的基质(matrix)，所述均一表面具有确定数量的能够与待分离的靶分子结合的结合配偶体(即修饰的免疫球蛋白可变结构域)固定于其上。相反，如果使用单特异性结合配偶体，则不会产生均匀基质，因为单结合配偶体不以同样的效率结合到基质。

本发明另一方面涉及使化合物靶向靶标的方法，包含步骤：

- (a)接触包含本发明的修饰的免疫球蛋白可变结构域分子或包含通过本发明的方法可获得的修饰免疫球蛋白分子，所述分子能够特异地与所述化合物结合，
- (b)向靶标提供包含免疫球蛋白可变结构域/化合物复合物的分子。

本发明修饰的免疫球蛋白或免疫球蛋白可变结构域可用于通过使靶标结合至修饰的结构环区，来向靶标提供通过特异 CDR 环构象结合至 CDR 的至少一种化合物。可用这种免疫球蛋白在疾病治疗过程中将治疗剂靶向优选的作用位点。

本发明的另一方面涉及包含、表达或编码本发明或通过本发明方法可获得的免疫球蛋白或免疫球蛋白可变结构域的分子文库。

本发明优选的免疫球蛋白或免疫球蛋白可变结构域文库包含至少 10 个免疫球蛋白或免疫球蛋白可变结构域，优选 100，更优选 1000，更优选 10000，更优选 100000，最优选超过 1000000 个变体免疫球蛋白或可变结构域，其在至少两个结构环或环区上具有修饰。

通常本发明的文库包含至少 10 个融合蛋白或结合因子(binding agent)，优选至少 100，更优选至少 1000，更优选至少 10^4 ，更优选至少 10^5 ，更优选至少 10^6 ，更优选至少 10^7 ，更优选至少 10^8 ，更优选至少 10^9 ，更优选至少 10^{10} ，更优选至少 10^{11} ，一直到 10^{12} ，在核糖体展示的情况下甚至更高的数量也是可行的。

结果发现最优选的文库成员在至少两个结构环或环区具有至少 4 个、甚至 5 个或 6 个氨基酸位置的突变。因此，本发明特别优选的文库由在至少两个结构环上具有至少 2、3 或 4 个氨基酸位置的突变的成员组成。

本发明的文库也可包含选自 VH、V_K、V_λ 和 VHH 的免疫球蛋白可变结构域的其中之一或其混合物或者由它们组成，适合于为商业原因确定结合配偶体的目的。

构建所述文库的优选方法可在上面以及实施例中找到。本发明的文库可用于鉴定与特定分子结合的免疫球蛋白或免疫球蛋白可变结构域。

本发明特别涉及多肽的蛋白质文库用于设计免疫球蛋白衍生物的用途，其包含本发明的或者通过本发明的方法获得的免疫球蛋白或免疫球蛋白可变结构域。可通过使用至少 10 个、优选 100、更优选

1000、更优选 10000、更优选 100000、最优选超过 1000000 的免疫球蛋白或变体可变结构域(每个具有至少两个修饰结构环)的相应免疫球蛋白或可变结构域的蛋白质文库,将抗原结合位点引入到免疫球蛋白或可变结构域来改变现有的免疫球蛋白。然后根据与特异性抗原的结合情况,对文库进行筛选。在根据所需特性进行分子鉴定后,通过基因工程技术将选出的免疫球蛋白或可变结构域克隆到原免疫球蛋白以取代野生型区。或者,可仅使编码修饰环或编码突变氨基酸的 DNA 进行交换,以获得具有额外特异抗原结合位点的免疫球蛋白。或者,可用其天然环境(context)的可变结构域,例如以 Fab、scFv 或完整免疫球蛋白分子的形式完成可变结构域结构环上的修饰。除了当制备单结构域免疫球蛋白、单免疫球蛋白结构域链或单链免疫球蛋白如 scFv 或单体(unibody) (单价免疫球蛋白片段)时之外,通常本发明的免疫球蛋白作为二聚体、优选异二聚体提供。

对突变的、抗原特异结构环的位点的选择取决于原免疫球蛋白的结构和额外结合位点的目的。如果例如原免疫球蛋白(即亲本免疫球蛋白)是 Fab 或 scFv,那么对轻链和/或重链可变结构域中的至少两个结构环进行修饰是可能的,而且优选对 CH1 和/或 CL 结构域中的至少两个结构环进行修饰以生成本发明的免疫球蛋白。因此,本发明的 Fab 分子可通过 CH1 和或 CL 结构域中的修饰的环区而含有新的结合位点。在此情况下通常最重要的是保持特异的 CDR 环构像和亲本免疫球蛋白、scFv 或 Fab 的天然结合特性。

为了生成文库,可制备在一个或多个可变结构域的两个或多个结构环中具有突变的突变体初始分子文库。对完全突变(complete mutate)的初始分子进行选择可具有一些优势,因为根据与抗原的结合情况对修饰的结构环进行选择将提供立体上有利的修饰。例如,如果完全分子是 scFv,那么将会有利的是,根据与抗原的结合情况对突变的初始 scFv 的文库进行筛选,随后根据与被 CDR 环识别(初始特异性)的抗原的结合情况对特异性结合物进行筛选。在一个备选的选择程序

中，初始(第一)抗原可在根据与抗原结合的情况对修饰的结构环进行筛选期间结合到 CDR 环上。如果与抗原的结合受到与第一抗原结合的影响，那么这种同时筛选可允许挽救在连续的选择过程中丢失的克隆。

本发明的优选实施方案是包含以下可变结构域的、或由该可变结构域组成的变体免疫球蛋白文库：在至少两个结构环的每一个上具有至少一个变体氨基酸位置的可变结构域。该文库可包含重链和轻链的免疫球蛋白结构域或其混合物和分子组合。

另一个优选的实施方案是包含以下 VHH 结构域或这种骆驼结构域的人源化形式的、或者由其组成的文库：在至少两个结构环或环区的每一个上具有至少一个变体氨基酸位置的 VHH 结构域或其人源化形式。

本发明另一个优选的实施方案是包含单链抗体的、或由其组成的文库例如 scFv 文库，其在任何单链抗体或 scFv 的可变结构域的至少两个结构环或环区的每一个上具有至少一个变体氨基酸位置。

本发明另一个的优选实施方案是双抗体文库，其在双抗体任何可变结构域的至少两个结构环或环区的每一个中包含至少一个变体氨基酸位置或由其组成。

本发明另一个优选的实施方案是小抗体文库，其在小抗体的任何可变结构域的至少两个结构环或环区的每一个中包含至少一个变体氨基酸位置或由其组成。

本发明另一个优选的实施方案是 Fab 文库，其在 Fab 的任何可变结构域的至少两个结构环或环区的每一个里包含至少一个变体氨基酸位置或由其组成。

本发明另一个优选的实施方案是抗体或 IgG 文库，优选人抗体文库，其在抗体或 IgG 结构域的任何可变结构域的至少两个结构环或环区的每一个里包含至少一个变体氨基酸位置或由其组成。

包含不同突变的免疫球蛋白或免疫球蛋白可变结构域或突变的

可变抗体结构域的融合分子的文库的大小需要(例如变体的数量)取决于工作任务。一般来说,重新产生抗体结合位点的文库需比用于进一步修饰(例如提高亲和力或改变对抗原的细微特异性)由修饰的结构环或环区形成的现存的工程抗原结合位点的文库大。

本发明还涉及多肽文库或核酸文库,其包含多种含有免疫球蛋白或免疫球蛋白可变结构域或至少两个包含于微型结构域内的结构环或环区的多肽或者其编码核酸。所述文库含有具有不同修饰的成员,其中大多数由在至少两个结构环或环区的修饰所确定。核酸文库优选包括至少10个不同的成员(具有至少两个潜在的氨基酸修饰)和更优选包括至少100,更优选1000或10000个不同的成员(例如通过随机化策略或组合技术进行设计)。还优选更具多样性的个体成员数目,如至少1000000或至少10000000,更优选至少 10^8 ,更优选至少 10^9 ,更优选至少 10^{10} ,更优选至少 10^{11} ,直到 10^{12} ,在核糖体展示的情况下更高的数目也是可行的。

本发明进一步的方面是:为了生成多特异免疫球蛋白,从至少两个本发明文库中选择两个不同的免疫球蛋白或免疫球蛋白可变结构域的组合。这些选择的特异免疫球蛋白可变结构域可彼此组合或与其它分子组合,与结构单元相似,设计结构域的最佳排列得到所需的特性如特异性和/或价数的组合。

此外,可在各种或所有不同的蛋白质位点将本发明的一个或多个修饰免疫球蛋白或免疫球蛋白可变结构域引入,而不破坏蛋白质结构。通过这种“结构域改组”技术产生可根据所需特性再次进行选择的新文库。

优选的文库含有本发明的免疫球蛋白或免疫球蛋白可变结构域或其衍生物。

本发明优选的实施方案是包含至少一个免疫球蛋白或免疫球蛋白可变结构域和根据本发明修饰的、与抗原结合的至少两个结构环或环区的结合抗原的分子(抗原结合分子),其中所述的结合分子与它

的 CDR 环没有相关的和/或特异的结合活性。它可包含用作抗体活性的其它部分(例如天然的或修饰的效应区(序列));然而,它缺乏“天然的”抗体结合区,即在它们天然位置上活性 CDR 环。本发明这些抗原结合分子具有以上对于该分子所描述的优势,然而没有特异的抗体结合活性;可是在结构环或环区有新引入的特异性结合活性。

对于本发明的抗原结合分子而言,还优选通过随机化技术(即通过随机化技术修饰至少两个结构环的一个或多个氨基酸残基或者通过将随机生成的插入引入这些结构环内中),将新的抗原结合位点引入到结构环中。备选优选的是使用组合方法。

根据另一方面,本发明涉及具有抗原结合位点的修饰的免疫球蛋白,所述抗原结合位点对于未修饰的免疫球蛋白来说是外源的并合并到免疫球蛋白或免疫球蛋白可变结构域的一个、两个、三个或多个结构环中。本文中术语“外源的”是指抗原结合位点不是由免疫球蛋白可变结构域的特异结构环或环区天然形成。

还根据另一方面,本发明涉及具有抗原结合位点的修饰的免疫球蛋白,所述抗原结合位点对于未修饰的免疫球蛋白而言是外源的并合并到可变结构域的一个、两个、三个或多个结构环上,其中所述修饰的免疫球蛋白以至少 10^3 mol^{-1} 、至少 10^4 mol^{-1} 、至少 10^5 mol^{-1} 、至少 10^6 mol^{-1} 、至少 10^7 mol^{-1} 、至少 10^8 mol^{-1} 、或至少 10^9 mol^{-1} 的亲合力与所述抗原结合。

根据本发明通常提供具有中或高亲和力的结合物。优选那些具有中等亲和力的结合物表现解离速率 K_d 介于 10^{-5} - 10^{-7} 之间,那些具有高亲和力的结合物有已证实的 K_d 介于 10^{-8} - 10^{-10} 之间,那些 K_d 低于 10^{-9} 的结合物最优选作为高亲和力的结合物。在某些情况下选择 K_d 甚至更低例如低于 10^{-11} ,通常低到 10^{-12} 的结合物是合适的。

本发明优选的免疫球蛋白或免疫球蛋白可变结构域包含至少两个抗原结合位点,第一位点与第一表位结合,第二位点与第二表位结合。

根据优选的实施方案，本免疫球蛋白或免疫球蛋白可变结构域包含至少三个环或环区，第一环或环区与第一表位结合，第二和第三环或环区与第二表位结合。至少第一或至少第二和第三环或环区或者二者都含有结构环。本发明免疫球蛋白或免疫球蛋白可变结构域包括其在本领域已知是功能性的片段，其含有根据本发明的基本元件：根据本发明的修饰结构环或环区。

本发明的免疫球蛋白优选由至少两个免疫球蛋白结构域或其包括微型结构域的部分组成，每个结构域含有至少一个抗原结合位点。优选的免疫球蛋白结构域对之一是 CL/CH1 对，可对其进行在位于 CL/CH1 对 C 末端的结构环区的工程改造。由此对一个或两个新的结合位点进行改造工程。在对特异的 CL/CH1 结合结构域对进行选择时，它可与可变结构域 VL 和 VH 进行重组，得到在 CDR 区具有一个“天然”结合位点和在其对面(即在 CL/CH1 结构域的 C 末端结构环区)具有一个或两个额外的结合位点的 Fab 分子。

因此，通过修饰含有可变结构域的免疫球蛋白亲本分子可获得本发明的免疫球蛋白。或者，可对免疫球蛋白恒定结构域进行工程改造，以获得在结构环区的结合位点，然后将此结构用作结构单元产生与可变免疫球蛋白结构域和任选与其它恒定结构域的组合，从而产生含有可变结构域及由结构环或结构环区形成的新结合位点的本发明免疫球蛋白。

根据本发明该优选的实施方案，提供一种免疫球蛋白，其包含至少一个免疫球蛋白可变区结构域和至少一个免疫球蛋白恒定区结构域；例如，在连接到 CH1 结构域的至少两个结构环中被修饰的可变结构域。

本发明另一方面涉及结合配偶体的试剂盒，包含

(a)包含具有合并到两个或多个结构环的抗原结合位点的修饰免疫球蛋白可变结构域的多肽，并且

(b)包含所述抗原表位的结合分子。

优选本发明的结合配偶体试剂盒包含

(a)本发明修饰的免疫球蛋白文库，并且

(b)包含抗原表位的结合分子。

根据本发明该试剂盒的这种结合分子可被用来筛选和辨别样品中或来自文库的天然的或修饰的本发明免疫球蛋白。它可进一步用于鉴定包含本发明修饰的免疫球蛋白或免疫球蛋白可变结构域的多肽的结合特异性。通过使用本发明的该试剂盒的结合分子，可测定本发明修饰的多肽的效能。

于此所定义的效能是指本发明修饰的分子与其抗原的结合特性。为了质量控制的目的，可用本领域已知的测定方法根据特异性和/或亲和力和/或亲合力定性或定量测定结合情况。

此外，本发明试剂盒的结合分子可用于从以上详细说明的文库中选择包含本发明修饰的免疫球蛋白或免疫球蛋白可变结构域的多肽，所述文库优选由为至少 10、优选至少 100、更优选至少 1000、更优选至少 10000、尤其至少 100000 的在结构环上具有不同修饰的多肽组成。

下列实施例进一步说明本发明。

实施例 1: VHH 文库的设计

骆驼 VHH 结构域 D2-L24 与鸡蛋白溶菌酶的复合晶体结构(在 Brookhaven Database 发表，其登录号为 1ZVH. Pdb)用来帮助设计突变的 VHH 结构域。SEQ ID No. 1 给出结构文件 1ZVH. Pdb 的 A 链序列。

```
SEQ ID No.1  
PREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFLL  
YSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSLSPGKAAA
```

被用作构建 VHH 文库基础的序列是来自专利 WO04041862A21、克隆 3E 的抗 TNF α VHH 结构域的序列，在 SEQ ID No. 2 中给出。

SEQ ID No.2

```

ccatggcccc ccgagaacca caggtgtaca ccctgcccc atcccgtgac
gagctcnnsn nsnnscaagt cagcctgacc tgcctgggtca aaggcttcta
tcccagcgac atcgccgtgg agtgggagag caatgggcagccggagaaca
actacaagac cacgcctccc gtgctggact ccgacggctc cttcttctc
tacagcaagc ttaccgtgnn snnsnnsagg tgnnsnns ggaacgtctt
ctcatgctcc gtgatgcatg aggctctgca caaccactac acacagaaga
gcctctccct gtctccgggt aaagcggccg ca

```

详细分析 1ZVH. Pdb 结构并对形成连接 β 链的环的残基进行目测之后, 决定使 SEQ ID No. 2 的残基 13、15、(即在 β 链 A 和 B 之间的环) 89、90、92 和 93(即在 β 链 E 和 F 之间的环) 随机化以生成文库。另外, 决定将三个随机化位置插入到 SEQ ID No. 2 的残基 14 和 15 之间, 决定将三个随机化位置插入到 SEQ ID No. 2 的残基 92 和 93 之间。

实施例 2: VHH 文库的构建

通过 PCR 组装以合成基因的方式制备编码 VHH 序列的工程改造基因。图 2 显示了序列及其翻译。用下划线标注在文库构建中待随机化的氨基酸残基。用于克隆的限制性位点包括如下并在图 2 中显示的核苷酸序列中用下划线标注: NcoI、BglII 和 NotI。

合成基因所编码的氨基酸序列在 SEQ ID. No.3 中给出。

SEQ ID No.3

```

MAPREPQVYTLPPSRDELXXXQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSF
FLYSKLTVXXXRWXXGNVFSVMSVHEALHNHYTQKSLSLSPGKAAA

```

最前的两个残基和最后两个残基由用于克隆的限制性位点所致。编码 SEQ ID. No.3 的核苷酸序列在 SEQ ID No. 4 中给出。

```

SEQ ID No.4: cttgccatgg ccccccgaga accacaggtg tac

```

最前的两个密码子和最后两个密码子由用于克隆的限制性位点所致。

用于合成基因 PCR 组装的寡核苷酸通过使用公用的软件工具 DNAWorks 3.1(<http://helixweb.nih.gov/dnaworks/>)进行设计, 按照标准实验方案用表 1 所出示的 SEQ ID No. 5-SEQ ID No. 22 的 18 条寡核苷酸通过 PCR 进行装配(Hoover DM, Lubkowski J., DNAWorks: 用于设计基于 PCR 的基因合成用的寡核苷酸的自动化方法(DNAWorks: an

automated method for designing oligonucleotides for PCR-based gene synthesis), *Nucleic Acids Res.* 2002 May 15 ; 30 (10) : e43).

1. CCATGGCAAGTTCAGCTGCAGGAAAGCGGTGGCGGCCTG (SEQ ID No. 5)
2. AGACGCAGGCTGCCGCCAGGCTGGACCAGGCCGCCACCGC (SEQ ID No. 6)
3. CGGCAGCCTGCGTCTGAGCTGTGCGGCCAGCGCCGTACC (SEQ ID No. 7)
4. AGGTGTAGCCGCTATGGTCGCTAAAGGTACGGCCGCTGGC (SEQ ID No. 8)
5. ACCATAGCGGCTACACCTATAACATTGGCTGGTTTCGTCA (SEQ ID No. 9)
6. TCACGTTCTTTTCCTGGCGCCTGACGAAACCAGCCAATGG (SEQ ID No. 10)
7. CGCCAGGAAAAGAACGTGAATTTGTGGCGCGTATTTACTG (SEQ ID No. 11)
8. ATAGGTATTGCCGCTGCTCCAGTAAATACGCGCCACAAAT (SEQ ID No. 12)
9. GAGCAGCGGCAATACCTATTATGCGGATAGCGTGAAAGGC (SEQ ID No. 13)
10. TGTCGCGGCTAATCGCGAAACGGCCTTTCACGCTATCCGC (SEQ ID No. 14)
11. GCGATTAGCCGCGACATTGCCAAGAACACGGTAGATCTTA (SEQ ID No. 15)
12. GGCTCCAGGTTGTTTCATCGTAAGATCTACCGTGTCTTGGC (SEQ ID No. 16)
13. CGATGAACAACCTGGAGCCCCGAAGACACAGCCGTGTATTA (SEQ ID No. 17)
14. GCCATCCCAGCCGCGCAATAATACAGGCTGTGTCTTCG (SEQ ID No. 18)
15. GCGGCTCGGGATGGCATTCCGACCAGCCGTAGCGTGGAAA (SEQ ID No. 19)
16. CCCTGGCCCCAGTAATTGTAGCTTTCACGCTACGGCTGG (SEQ ID No. 20)
17. CAATTACTGGGGCCAGGGCACCCAGGTGACCGTCAGCTCT (SEQ ID No. 21)
18. GCGGCCGCAGAGCTGACGGTCACCTG (SEQ ID No. 22)

表 1: 用于组装编码工程改造的 VHH 基因的合成基因的寡核苷酸
 简要地, 将等体积寡核苷酸溶液(每个浓度为约 1 mg/ml)混合并用水稀释使每种寡核苷酸的终浓度为约 1 ng/ μ l。用 PCR 溶液将寡核苷酸混合物稀释 5 倍。组份的终浓度是: 每种寡核苷酸 0.2 ng/ μ l, Tris-HCl (pH 8.8) 20 mM, KCl 10 mM, $(\text{NH}_4)_2\text{SO}_4$ 10 mM, MgSO_4 6 mM, Triton X-100 0.1% (v/v), 牛血清白蛋白 0.1 mg/ml, 每种 dNTP 0.2 mM, Pfu 聚合酶 2.5 U。用于基因组装的 PCR 方案以一个 5 分钟 95 $^\circ\text{C}$ 的变性步骤开始, 期间加入聚合酶避免任何可能的引物错配(mispriming) (“热启动” PCR)。该步骤后接包括以下各步的 25 个循环: 变性温度 95 $^\circ\text{C}$ 30 s, 退火温度 55 $^\circ\text{C}$ 30 s, 延伸温度 72 $^\circ\text{C}$ 1.5 min。该方案的最后一步是 72 $^\circ\text{C}$ 10 min 保温循环。为了进行基因扩增, 用基因组装反应所产生的 1 μ l 混合物作为模板, 用最外面的寡核苷酸作引物(SEQ ID No. 5 和 SEQ ID No. 22)。用于 PCR 基因扩增的实验方案与用于基因组装的基本上一样。随后通过在载体 pET27b 上的 NcoI 和 NotI 限制性位点克隆组装

的 合 成 VHH 基 因 (Novagen, <http://www.merckbiosciences.co.uk/product/69863>; <http://www.novagen.com>)并通过 DNA 测序对序列进行验证。

然后用 PCR 构建随机化文库。用于最初 2 次 PCR 反应的模板是如上所述的所克隆的合成 VHH 基因。用于最初 2 次 PCR 反应的引物对如下:

```

3esynmu1 (gactccatgg
caagtgcaac tgcaggaaag cggagggcggc ctggttnnsc cannsnnsnn
snnsggcagc ctgcgtctga gct (SEQ ID No. 23)) 和 3esynmu2
(catgagatct acgggtgttct tggcg (SEQ ID No. 24)); 3esynmu3
(catgagatct tacgatgnns nnsttgnnsn nsnnnsnnsnn sgaagatacg
gcggtgtatt attg (SEQ ID No. 25)) 和 3esyn2 (aatagcggcc
gcagagctca cggtcacc (SEQ ID No. 26)).

```

产生的 PCR 产物用 BglII 消化、连接,连接产物用作 PCR 反应的模板,其引物为 3esyn1 (acgtccatgg caagtgcaac tgcag (SEQ ID No. 27)) 和 3esyn2 (aatagcggcc gcagagctca cggtcacc (SEQ ID No. 26))。将在引物 3esynmu1(SEQ ID No. 23)和 3esynmu3 (SEQ ID No. 25)上的 NNS 密码子在序列中引入随机位置。选择编码所有 20 种天然存在氨基酸的密码子 NNS(IUPAC 编码, S 是指 C 或 G),但要避免 3 个当中 2 个终止密码子。图 3 显示了 PCR 反应和连接程序的示意图。水平箭头显示 PCR 引物的位置和方向,垂直线分别显示 NcoI、BglII 和 NotI 位点的位置(从左到右)。

随后通过 NcoI 限制位点将这种随机化 PCR 产物克隆到噬菌粒克隆载体 pHEN1 中,与 pelB 分泌信号在同一读码框内(Hoogenboom HR, Griffiths AD, Johnson KS, Chiswell DJ, Hudson P, Winter G. 在丝状噬菌体表面上的多亚基蛋白: 用于展示抗体(Fab)重链和轻链的方法 (Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains). *Nucleic Acids Res.* 1991 Aug 11; 19 (15) : 4133-7)。所述基因 3' 端的 NotI 限制位点使 VHH 文库插入到与 pHEN1 载体所包含的丝状噬菌体

fd 的次要外壳蛋白(蛋白 III)编码基因在同一个读码框内。随机 VHH 文库插入的工程改造序列作为 SEQ ID No. 28 中的核苷酸序列给出, 并翻译成 SEQ ID No. 29 的氨基酸序列。SEQ ID No. 29 中的字母 X 表示随机氨基酸残基。

SEQ ID No. 28:

```

1  ccattgcaag tgcaactgca gaaagcgga ggcggtctgg tttnsccann snnsnnsnns
61  ggcagcctgc gtctgagctg cgcggcgctcc ggccgtacct ttagcgacca ttcgggctat
121 acctatacca ttggctggtt cgcgcaggcg ccagggaaag aacgtgaatt tgtggcgcgt
181 atttactgga gcagcggcaa tacctactat gcggatagcg tgaaaggccg ttttgcgatt
241 agccgcgaca tcgccaagaa caccgtagat cttacgatgn nsnnsttgnn snnsnnsnns
301 nnsagaagata cggcgggtgta ttattgcgca gcgcgtgacg gcattccgac ctcccgtagc
361 gtggaagct acaattactg gggccagggc acccaggtga ccgtgagctc tgcggccgc

```

SEQ ID No. 29:

```

PWQVQLQESGGGLVXPXXXXGSLRSLSCAASGRTFSDHSGYTYTIGWFRQAPGKEREVARIYWSSGNTYYADSVKGR
FAISRDIKNTVDLTMXXLXXXXEDTAVYYCAARDGIPTSRVESYNYWGQGTQVTVSSAA

```

然后将连接产物转化到大肠杆菌 TGI 中, 确定得到的克隆数, 通过限制性分析和 DNA 测序对许多所选的克隆进行检验。对于表面展示文库噬菌体制备的下列步骤而言, 遵照标准实验方案。简言之, 通过电穿孔将连接混合物转化到大肠杆菌 TGI 细胞。随后, 用辅助噬菌体 M13-KO7 使噬菌体颗粒从 E. coli TGI 细胞中释放出来。然后用 PEG/NaCl 在两个步骤中将噬菌体颗粒从培养悬浮液沉淀下来, 用水溶解并通过淘选进行选择或者将其储藏在 -80 °C。

实施例 3: 针对人血清白蛋白(HSA)淘选 VHH 噬菌体文库

按照标准方案进行 3 轮淘选。(例如肽和抗体的噬菌体展示: 实验室手册(Phage display of peptides and antibodies: a laboratory manual), Kay 等, 1996, Academic Press, San Diego, Calif.)简言之, 应用下列方法。用 HAS 包被 Maxisorp96 孔板(Nunc)。向每个孔加入 200 μ l 下列溶液: 0.1M 碳酸钠缓冲液(pH 9.6)、以下浓度的 HAS: 第一轮淘选: 2 mg/ml HAS; 第二轮淘选: 1 mg/ml HAS; 第三轮淘选: 1 mg/ml HAS。于 37°C 孵育 1 小时, 随后每个孔加 2%奶粉(M-PBS)200 μ l 于室温下 1

小时进行封闭。然后通过加入 100 μ l 噬菌体悬浮液和 100 μ l 4% 奶粉 (M-PBS) 允许表面展示噬菌体文库与被结合的 HAS 进行反应, 随后室温下振荡 45 分钟、静置 90 分钟来进行孵育。

按如下步骤洗去未结合的噬菌体颗粒。

在第一轮淘选之后: 10 x 300 μ l T-PBS、5x 300 μ l PBS;

在第二轮淘选之后: 15 x 300 μ l T-PBS、10x 300 μ l PBS;

在第三轮淘选之后: 20 x 300 μ l T-PBS、20 x 300 μ l PBS。

结合的噬菌体粒子的洗脱通过每孔加入 200 μ l 0.1M 甘氨酸(pH 2.2)、室温下振荡孵育 30 分钟来完成。随后通过加入 60 μ l 2M Tris 碱以中和噬菌体悬浮液, 然后将 10 ml 对数生长的大肠杆菌 TGI 培养物与 0.5 ml 洗脱的噬菌体混合并于 37°C 保温 30 分钟, 使噬菌体转染到大肠杆菌 TGI 细胞中。最后, 将经转染的细菌涂布于含有 1% 葡萄糖和 100 μ g/ml 氨苄青霉素的 TYE 培养基, 于 30°C 温育过夜。

实施例 4: 根据 HAS 选择所选出的 VHH 突变体克隆的可溶性表达克隆

以中量制备来分离经过 3 轮淘选所选出的噬菌体的噬菌粒 DNA。用 PCR 对编码突变的 VHH 结构域区的 DNA 进行分批扩增并克隆 NcoI-NotI 到载体 pNOTBAD/Myc-His, 它是具有插入的 NotI 限制位点方便克隆的大肠杆菌表达载体 pBAD/Myc-His(Invitrogen)。将所连接的构建体用电穿孔转化到大肠杆菌 LMG194 细胞(Invitrogen), 于 30°C 在含有 1% 葡萄糖和氨苄青霉素的 TYE 培养基上培养过夜。将所选择的克隆接种到含有氨苄青霉素的 200 μ l 2 \times YT 培养基, 于 30°C 培养过夜, 通过加入终浓度为 0.1% 的 L-阿拉伯糖进行诱导。在 16°C 过夜表达之后, 离心收集细胞并用 100 μ l 硼酸钠缓冲液(pH 8.0)于 4°C 过夜处理, 以制备胞质提取物。将 50 μ l 胞质提取物用于 ELISA(见下)。

实施例 5: 根据 HSA 筛选所得的 VHH 突变体的 ELISA

根据以下方案用 ELISA 检验根据人血清白蛋白结合情况所选择的 VHH 突变体的胞质提取物

包被: 微量滴定板(NUNC, Maxisorp), 每孔 100 μ l

含 100 μ g HSA /ml 的 PBS, 4°C 过夜。

洗涤: 3x 200 μ l PBS

封闭: 含 1%酪蛋白封闭剂(Blocker Casein)的 PBS 溶液(Pierce) RT 下 1h

洗涤: 3x 200 μ l PBS

胞质提取物结合: 50 μ l 胞质提取物(实施例 4), 50 μ l PBS 0.05% 吐温 20, 室温过夜

洗涤: 3x 200 μ l PBS

第一抗体: 抗 His4(Qiagen), 1 : 1000 在 PBS 0.05% 吐温 20

RT 下 90 min, 每孔 100 μ l

洗涤: 3x 200 μ l PBS

第二抗体: 山羊抗鼠*HRP(SIGMA), 1:1000 在 PBS 0.05%吐温 20, RT 下 90 min(室温), 每孔 100 μ l

洗涤: 3x 200 μ l PBS

检测: 含 3mg/ml OPD 的柠檬酸钠/磷酸钠缓冲液, pH 4.5, 0.4 μ l 30% H_2O_2

在背景达到太高之前停止: 100 ml 3M H_2SO_4

读取吸光度: 492/620 nm

实施例 6: 其中只对一个环(C'D 环)进行随机化的文库的实例

使用编码经工程改造的 VHH 的合成基因(以上实施例 2 中所描述的基因)作为两个 PCR 反应模板, 其中用到以下引物对: SEQ ID No. 30 (actgctcgag agacatcgcc aagaacac; esynmu4)和 SEQ ID No. 26 (3esyn2) 以及 SEQ ID No. 31 (cacactcgag atcgcaasn nsnnsnncac snnsnncgca tagtagtat tgcc; 3esynmu5)和 SEQ ID No. 27 (3esyn1)。产生的 PCR 产物用 XhoI 消化、连接, 使用连接产物作为模板与引物 3esyn1(SEQ ID No. 27)和 3esyn2 (SEQ ID No. 26)进行 PCR 反应。与在实施例 2 中所描述的相似, 引物 3esynmu4 (SEQ ID No. 30)和 3esynmu5 (SEQ ID No.

31)的 NNS 密码子将随机位置引入到序列中。图 4 显示了 PCR 反应和连接程序的示意图。水平箭头显示 PCR 引物的位置和方向，垂直线分别显示 NcoI、XhoI 和 NotI 位点的位置(从左到右)。

随后将该随机化的 PCR 产物克隆到噬菌粒克隆载体 pHEN1, 与 pelB 分泌信号和载体 pHEN1 所包含的丝状噬菌体 fd 的次要外壳蛋白(蛋白 III)在同一个读码框中(Hoogenboom HR, Griffiths AD, Johnson KS, Chiswell DJ, Hudson P, Winter G. 在丝状噬菌体表面上的多亚基蛋白: 用于展示抗体(Fab)重链和轻链的方法(Multi-subunit proteins on the surface of filamentous phage : methodologies for displaying antibody (Fab) heavy and light chains). Nucleic Acids Res . 1991 Aug 11 ; 19 (15): 4133-7), 如在实施例 2 中所描述。随机化 VHH 文库插入的工程改造序列作为 SEQ ID No. 32 中的核苷酸序列给出, 并且被翻译成 SEQ ID No. 33 中的氨基酸序列。SEQ ID No. 33 中的字母 X 表示随机化的氨基酸残基。

SEQ ID No. 32

```

1 ccatggcaag ttcagctgca gaaagcggg ggcgccctgg tccagcctgg cggcagcctg
61 cgtctgagct gtgcggccag cggccgtacc ttagcgacc atagcggcta cacctatacc
121 attggctggg ttcgtcaggc gccaggaaaa gaactgaaat ttgtggcgcg tatttactgg
181 agcagcggca atacctatta tgcgnnsnns gtgnnsnnsn nsttcgcat ctcgagagac
241 attgccaaga acacggtaga tcttacgatg aacaacctgg agcccgaaga cacagccgtg
301 tattattgcg cggctcggga tggcattccg accagccgta gcgtggaag ctacaattac
361 tggggccagg gcaccagggt gaccgtcagc tctgcgccg c

```

SEQ ID No. 33

```

PWQVQLQESGGGLVQPGGSLRLSCAASGRTFSDHSGYTYTIGWFRQAPGKEREFVARIYWSSGNTYYAXVXXX
FAISRDIKNTVDLTMNLEPEDTAVYYCAARDGIPTSRSVESYNYWGQGTQVTVSSAA

```

然后将连接产物转化到大肠杆菌 TGI, 测定所得到的克隆数量, 通过限制性分析和 DNA 测序对许多克隆进行检验。对于下列表面展示文库噬菌体制备的步骤而言, 如在实施例 2 中所描述的那样遵照标准实验方案。特异性结合克隆的淘选、选择和鉴定步骤基本上如实施例 3、4 和 5 中所描述的来进行。

实施例 7: 其中对三个环(AB、EF 和 C'D 环)进行随机化的文库的实

例

使用实施例 2 中所描述的在 AB 环和 EF 环上具有随机化残基的文库作为模板用于 PCR, 其中使用与实施例 6 中的相同的引物对: SEQ ID No. 30 (esynmu4) 和 SEQ ID No. 26 (3esyn2) 以及 SEQ ID No. 31 (3esynmu5) 和 SEQ ID No. 27 (3esyn1)。用于文库构建、特异性结合克隆的克隆、淘选、选择和鉴定的随后步骤与实施例 2、3、4 和 5 中所描述的基本相同。

实施例 8: 在一个、二个和三个结构环中具有随机化氨基酸位置的可变结构域文库的比较

将文库用于以各种抗原进行的淘选。

鸡蛋溶菌酶作为抗原:

进行 3 轮淘选。通过每孔加入 200 μ l 下列溶液, 用鸡蛋溶菌酶包被 Maxisorp 96 孔板, :

PBS, 和下列浓度的溶解的鸡蛋溶菌酶(HEL):

第一轮淘选: 2 mg/ml HEL

第二轮淘选: 1 mg/ml HEL

第三轮淘选: 1 mg/ml HEL

37°C 孵育 1 小时, 随后用每孔 200 μ l 的 2% 奶粉(M-PBS) 室温封闭 1 小时。

然后通过加入 100 μ l 噬菌体悬浮液和 100 μ l 4% 奶粉(M-PBS), 随后于室温下振荡 45 分钟并静置 90 分钟进行孵育, 允许表面展示噬菌体文库与被结合的鸡蛋白溶菌酶进行反应。

按照以下洗去未结合的噬菌体颗粒:

第一轮淘选: 10 x 300 μ l T-PBS、5x 300 μ l PBS

第二轮淘选: 15 x 300 μ l T-PBS、10x 300 μ l PBS

第三轮淘选: 20 x 300 μ l T-PBS、20 x 300 μ l PBS

通过每孔加入 200 μ l 0.1 M 甘氨酸(pH 2.2) 并于室温振荡保温 30 分钟洗脱被结合的噬菌体颗粒。随后加入 60 μ l 2M Tris 碱中和噬菌体

悬液，再通过将 10 ml 对数生长培养物与 0.5 ml 洗脱的噬菌体混合并于 37°C 孵育 30 分钟，使其感染到大肠杆菌 TGI 细胞中。最后，将被感染的细菌涂布于含有 1%葡萄糖和 100 µg/ml 氨苄青霉素的 TYE 培养基，于 30°C 保温过夜。

人血清白蛋白作抗原：

将实施例 2、6 和 7 的文库用于以上所描述的各轮淘选中。具体而言，将噬菌体文库悬浮在结合缓冲液中(PBS、1%卵清蛋白、0.005%吐温 20)并用直接固定在 maxisorp 板上的人血清白蛋白进行淘选(在 PBS 中 10µg/ml, 4°C 过夜；用酪蛋白封闭剂(Pierce)封闭该板。两小时后，通过反复的洗涤(PBS、0.05% 吐温 20)除去未结合的噬菌体，被结合的噬菌体用 500 mM KCl、10mM HCl(pH2)进行洗脱。

完成每轮 HSA 特异的淘选之后，根据与 TNFα 的结合情况对得到的克隆进行选择或检验。按照 WO2004/041862 的实施例 1 中所描述的根据 TNFα 特异性进行选择 and 试验。

FcRn 作为抗原：

如 WO02060919 实施例 6.2 所描述的进行淘选。简言之，将噬菌体文库重新悬浮于 5 ml 20 mM MES(pH 6.0)/5%脱脂牛奶/0.05% 吐温 20 中，并加入(100µl 的 5×10^{12} PFU/ml/孔)到预先用 1µg 鼠类 FcRn 包被的 Maxisorp 免疫板的 20 个孔内，用 5%脱脂牛奶进行阻断。37 °C 孵育 2 小时之后，用 20 mM MES(pH 6.0)/0.2% 吐温 20/0.3 M NaCl 洗涤孔 10-30 次，在 100 µl PBS (pH 7.4)/孔中、于 37°C 下孵育 30 分钟来洗脱噬菌体。如实施例 3 所描述，将噬菌体用于重新感染对数生长长期的大肠杆菌 TGI。

完成每一轮针对 FcRn 的淘选之后，如以上所描述的，根据与 TNFα 的结合情况对得到的克隆进行选择或检验。

Fc-γ 受体作为抗原：

如在 Berntzen 等(2006) Protein Eng Des Sel 19:121-128 中所描述的一样，针对 Fc-γRI、Fc-γRIIA 和 Fc-γRIIA 重组融合蛋白进行淘选。

完成每一轮针对 Fc-受体的淘选之后, 如以上所描述的, 根据与 TNF α 的结合情况对克隆进行选择或检验。

实施例 9:

该实施例证明了向抗体片段引入新的功能或额外结合特异性的可能性。用作修饰起始点的分子是鼠类单链抗体片段 sFv 26-10(Huston 等(1988) Proc Natl Acad Sci U S A. 85:5879-5883)。用随机氨基酸序列对五个不同的文库进行构建, 以修饰不同的结构环序列。

文库 26-10-1: (SEQ ID NO. 34)

```
EVQLQQSGPELVKPGASVRMSCKSSGYIFTDFYMNWVRQSHGKSLDYIGYISPYSVGTGYNQKF
K GKATLTVDKSSSTAYMELRSLTSEDSAVYYCAGSSGNKWAMDYWGHGASVTVSSGGGGSGGGG
SGGGGSDVVMQTPTPLSLPVSLGDQASISCRSSQSLVHSNGNTYLNWYLQKAGQSPKLLIYKVS
N RFSGXXXXFSGSGSGTDFTLTKISXXXXXXXXGIYFCSQTTHVPPTFGGGTKLEIKR
```

文库 26-10-2: (SEQ ID NO. 35)

```
EVQLQQSGPELVKPGASVRMSCKSSGYIFTDFYMNWVRQSHGKSLDYIGYISPYSVGTGYNQKF
K GKATLTVDKSSSTAYMELRSLTSEDSAVYYCAGSSGNKWAMDYWGHGASVTVSSGGGGSGGGG
SGGGGSDVVMQTPTPLSLPXXXXXQASISCRSSQSLVHSNGNTYLNWYLQKAGQSPKLLIYKVS
N RFSGXXXXFSGSGSGTDFTLTKISXXXXXXXXGIYFCSQTTHVPPTFGGGTKLEIKR
```

文库 26-10-3: (SEQ ID NO. 36)

```
EVQLQQSGPELVKPGASVRMSCKSSGYIFTDFYMNWVRQSHGKSLDYIGYISPYSVGTGYNQKF
K GKATLTVDKSSSTAYMELRSLTSEDSAVYYCAGSSGNKWAMDYWGHGASVTVSSGGGGSGGGG
SGGGGSDVVMQTPTPLSLPXXXXXQASISCRSSQSLVHSNGNTYLNWYLQKAGQSPKLLIYKVS
N RFSGXVPDRFSGSGSGTDFTLTKISXXXXXXXXGIYFCSQTTHVPPTFGGGTKLEIKR
```

文库 26-10-4: (SEQ ID NO. 37)

```
EVQLQQSGPELVKPGASVRMSCKSSGYIFTDFYMNWVRQSHGKSLDYIGYISPYSVGTGYNQKF
K GKATLTVDKSSSTAYMELRSLTSEDSAVYYCAGSSGNKWAMDYWGHGASVTVSSGGGGSGGGG
SGGGGSDVVMQTPTPLSLPVSLGDQASISCRSSQSLVHSNGNTYLNWYLQKAGQSPKLLIYKVS
N RFSGXXXXXXXXFSGSGSGTDFTLTKISXXXXXXXXGIYFCSQTTHVPPTFGGGTKLEIKR
```

文库 26-10-5: (SEQ ID NO. 38)

```
EVQLQQSGPELVKPGASVRMSCKSSGYIFTDFYMNWVRQXXXXXXXXDYIGYISPYSVGTGYNQKF
K GKATLTVDKSSSTAYMELRSLTSEDSAVYYCAGSSGNKWAMDYWGHGASVTVSSGGGGSGGGG
SGGGGSDVVMQTPTPLSLPVSLGDQASISCRSSQSLVHSNGNTYLNWYLQXXXXXXXXKLLIYKVS
N RFSGVPPDRFSGSGSGTDFTLTKISRVEAEDLGIYFCSQTTHVPPTFGGGTKLEIKR
```

通过序列反向翻译制备文库, 其中随机化的氨基酸位置由核苷三联体 NNK 所编码。

文库 26-10-1 基因: (SEQ ID NO. 39)

GAAGTTCAGCTGCAGCAGTCTGGTCCGGAAGTGGTTAAACCGGGTGCTTCTGTTTCGTATGTCTT
GCAAATCTTCTGGTTACATCTTCACCGACTTCTACATGAACTGGGTTTCGTACGTCTCACGGTAA
ATCTCTGGACTACATCGGTTACATCTCTCCGTAAGTCTGGTGTACCGGTTACAACCAGAAATTC
AAAGGTAAAGCTACCCTGACCGTTGACAAATCTTCTTCTACCGCTTACATGGAAGTGCCTTCTC
TGACCTCTGAAGACTCTGCTGTTTACTACTGCGCTGGTCTTCTGGTAACAAATGGGCTATGGA
CTACTGGGGTCACGGTGCTTCTGTTACCGTTTCTTCTGGTGGTGGTGGTCTGGTGGTGGTGGT
TCTGGTGGTGGTGGTCTGACGTTGTTATGACCCAGACCCCGCTGTCTCTGCCGTTTCTCTGG
GTGACCAGGCTTCTATCTCTTGCCGTTCTTCTCAGTCTCTGGTTCACTCTAACGGTAACACCTA
CCTGAACTGGTACCTGCAGAAAGCTGGTACGCTCCGAAACTGCTGATCTACAAAGTTTCTAAC
CGTTTCTCTGGTNNKNNKNNKNNKNTTCTCTGGTCTGGTCTGGTACCGACTTACCCTGAAAA
TCTCTNNKNNKNNKNNKNNKNNKNNKGGTATCTACTTCTGCTCTCAGACCACCCACGTTCCGCC
GACCTTCGGTGGTGGTACCAAAGTGGAAATCAAACGT

文库 26-10-2 基因: (SEQ ID NO. 40)

GAAGTTCAGCTGCAGCAGTCTGGTCCGGAAGTGGTTAAACCGGGTGCTTCTGTTTCGTATGTCTT
GCAAATCTTCTGGTTACATCTTCACCGACTTCTACATGAACTGGGTTTCGTACGTCTCACGGTAA
ATCTCTGGACTACATCGGTTACATCTCTCCGTAAGTCTGGTGTACCGGTTACAACCAGAAATTC
AAAGGTAAAGCTACCCTGACCGTTGACAAATCTTCTTCTACCGCTTACATGGAAGTGCCTTCTC
TGACCTCTGAAGACTCTGCTGTTTACTACTGCGCTGGTCTTCTGGTAACAAATGGGCTATGGA
CTACTGGGGTCACGGTGCTTCTGTTACCGTTTCTTCTGGTGGTGGTGGTCTGGTGGTGGTGGT
TCTGGTGGTGGTGGTCTGACGTTGTTATGACCCAGACCCCGCTGTCTCTGCCGNNKNNKNNKNN
NKNNKCAGGCTTCTATCTCTTGCCGTTCTTCTCAGTCTCTGGTTCACTCTAACGGTAACACCTA
CCTGAACTGGTACCTGCAGAAAGCTGGTACGCTCCGAAACTGCTGATCTACAAAGTTTCTAAC
CGTTTCTCTGGTNNKNNKNNKNNKNTTCTCTGGTCTGGTCTGGTACCGACTTACCCTGAAAA
TCTCTNNKNNKNNKNNKNNKNNKNNKGGTATCTACTTCTGCTCTCAGACCACCCACGTTCCGCC
GACCTTCGGTGGTGGTACCAAAGTGGAAATCAAACGT

文库 26-10-3 基因: (SEQ ID NO. 41)

GAAGTTCAGCTGCAGCAGTCTGGTCCGGAAGTGGTTAAACCGGGTGCTTCTGTTTCGTATGTCTT
GCAAATCTTCTGGTTACATCTTCACCGACTTCTACATGAACTGGGTTTCGTACGTCTCACGGTAA
ATCTCTGGACTACATCGGTTACATCTCTCCGTAAGTCTGGTGTACCGGTTACAACCAGAAATTC
AAAGGTAAAGCTACCCTGACCGTTGACAAATCTTCTTCTACCGCTTACATGGAAGTGCCTTCTC
TGACCTCTGAAGACTCTGCTGTTTACTACTGCGCTGGTCTTCTGGTAACAAATGGGCTATGGA
CTACTGGGGTCACGGTGCTTCTGTTACCGTTTCTTCTGGTGGTGGTGGTCTGGTGGTGGTGGT
TCTGGTGGTGGTGGTCTGACGTTGTTATGACCCAGACCCCGCTGTCTCTGCCGNNKNNKNNKNN
NKNNKCAGGCTTCTATCTCTTGCCGTTCTTCTCAGTCTCTGGTTCACTCTAACGGTAACACCTA
CCTGAACTGGTACCTGCAGAAAGCTGGTACGCTCCGAAACTGCTGATCTACAAAGTTTCTAAC
CGTTTCTCTGGTGTCCGGACCGTTTCTCTGGTCTGGTCTGGTACCGACTTACCCTGAAAA
TCTCTNNKNNKNNKNNKNNKNNKNNKGGTATCTACTTCTGCTCTCAGACCACCCACGTTCCGCC
GACCTTCGGTGGTGGTACCAAAGTGGAAATCAAACGT

文库 26-10-4 基因: (SEQ ID NO. 42)

并表达为 scFv。如在实施例 9 中所描述的那样来进行表达和噬菌体展示载体构建。从而完成噬菌体选择。

野生型 scFv 序列和相应文库如图 5-8 所示。

实施例 11: Fab 文库, 其中恒定结构域的结构环被随机化

为了从免疫球蛋白结构域文库中选择其中位于结构环的残基被随机化的特异性结合分子, 可应用各种形式。可使用单结构域如 VL、VH、CH1、CH2、CH3、CH4 或 CL 结构域, 可使用由 CH2 和 CH3 结构域组成的 Fc 片段(包括或不包括铰链区域或其组成部分), 可使用 Fv 或单链 Fv 片段, 可使用完整抗体或免疫球蛋白结构域其它组合。特别受关注的、用于选择特异性结合分子的一种形式是 Fab 片段, 它是两条链(即抗体的 VL-CL 部分和 VH-CH1 部分)的异源二聚体。Fab 片段已被认识很长时间, 可通过木瓜蛋白酶水解 IgG 产生, 也可在很大范围内的不同的表达系统中重组产生, 如大肠杆菌(*Escherichia coli*)、酿酒酵母(*Saccharomyces cerevisiae*)、巴斯德毕赤酵母(*Pichia pastoris*)、昆虫细胞或哺乳动物细胞。

为本领域已知的是, 表面展示系统如噬菌体展示、酵母展示和其它系统如核糖体展示等可用于从大文库中富集和选择特异性结合分子如的 Fab 片段(例如见 Hoogenboom HR、Griffiths AD、Johnson KS、Chiswell DJ、Hudson P、Winter G. 在丝状噬菌体表面上的多亚基蛋白: 用于展示抗体(Fab)重链和轻链的方法(Multi-subunit proteins on the surface of filamentous phage:methodologies for displaying antibody (Fab) heavy and light chains). *Nucleic Acids Res.* 1991 Aug 11; 19 (15) : 4133-7.; Kang AS, Barbas CF, Janda KD, Benkovic SJ, Lerner RA. 通过在噬菌体表面组装组合的抗体 Fab 文库连接识别功能和复制功能(Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces). *Proc Natl Acad Sci U S A.* 1991 May 15; 88 (10) : 4363-6.; Kang X, Yang BA, Hu Y, Zhao H, Xiong W, Yang Y, Si B, Zhu Q. 通过噬菌体展示制备针对重症急性呼吸道综合症冠状病毒的人综合 Fab 分子(Human

neutralizing Fab molecules against severe acute respiratory syndrome coronavirus generated by phage display). *Clin Vaccine Immunol.* 2006 Aug; 13 (8) : 953-7. ; Weaver-Feldhaus JM, Lou J, Coleman JR, Siegel RW, Marks JD, Feldhaus MJ. 用于制备组合 Fab 文库和表面展示的酵母杂合(Yeast mating for combinatorial Fab library generation and surface display). *FEBS Lett.* 2004 Apr 23; 564 (1-2) : 24-34.)。

例如, 如果将噬菌体展示系统应用于展示 Fab 片段(例如 Fab 片段文库), 那么可将该 Fab 片段中的一条链例如 VH-CHI 链与例如噬菌体 M13 的蛋白 III 表达为融合蛋白从而使该链展示在噬菌体表面上, 而将另一条链 VL-CL 以可溶的形式表达并与表面锚定的 VH-CHI 链形成天然异源二聚体。在一个典型的 Fab 表面展示文库中, 存在不同的 VH 和 VL 序列, 通常可以来源于鼠或人供体, 但通过体外的方法如定点诱变也可产生多样性。不同的结合位点因此而产生, 通过使用合适的展示或其它富集或选择方法, 可从这样的通过由 VH 和 VL 形成的结合位点与它们的结合配偶体结合的文库中分离出特异性结合克隆。

生成的 Fab 片段通过其 VH 和 VL 形成的天然结合位点与一个靶标结合, 通过由它们结构环形成的结合位点与另一个靶目标结合(或与同样的靶标第二次结合)。为了获得这样的工程改造 Fab 片段, 首先制备 Fab 文库, 其中结构环的残基被随机化残基置换。还可进行额外残基的插入。可用这种方法工程改造的结构环可以是 VH 或 VL 的 C-末端环(“底”环), 或者是 CHI 或 CL 结构域的 N-末端(“顶”环)或 C-末端(“底”环)。结构域与在任何这些位置上的改造结构环的不同组合也是可能的。可用于选择特异性结合结构域的一种形式是作为单结构域、作为 Fv 或单链 Fv 片段, 或如下所详述的作为 Fab 片段。

分别编码与人 Her2 结合的抗体 4D5 (Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW Jr, Leahy DJ. 单独的或与赫赛汀 Fab 复合的 HER2 胞外区的结构(Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab). *Nature.*

2003 Feb 13; 421 (6924): 756-60)的 VH-CHI 和 VL-CL 的基因用于该实施例。构建由所述基因的 4D5 VL-CL 编码部分构成的合成基因,使其 5' 端侧翼带有 NcoI 位点,用于插入到与噬菌粒载体 pHEN1 所含的 pelB 信号序列在同一个读码框内(Hoogenboom HR, Griffiths AD, Johnson KS, Chiswell DJ, Hudson P, Winter G. 在丝状噬菌体表面上的多亚基蛋白: 用于展示抗体(Fab)重链和轻链的方法(Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains). Nucleic Acids Res. 1991 Aug 11; 19 (15): 4133-7), 后接终止密码子。在从 VL 到 CL 之间的序列过度位置中, 包含唯一的 BsiWI 限制位点, 其与位于 VL-CL 编码基因终止密码子下游的唯一 AscI 限制位点结合, 随后用于以含有随机化结构环的文库插入物来置换野生型 CL 序列。其前面是取自 Carter 等.(Carter P, Kelley RF, Rodrigues ML, Snedecor B, Covarrubias M, Velligan MD, Wong WL, Rowland AM, Kotts CE, Carver ME 等, 二价人源化抗体片段的高水平大肠杆菌表达和制备(High level Escherichia coli expression and production of a bivalent humanized antibody fragment). Biotechnology (N Y) . 1992 Feb; 10 (2) : 163-7)的含有核糖体结合位点的序列(Shine-Dalgarno 序列), 后接编码热稳定肠毒素 II (st II) 信号序列的基因片段, 该基因片段与编码抗体 4D5 的部分的 VH-CHI 融合在同一个读码框中, 随后是用于框内插入到载体 pHEN1 的 NotI 位点。该二顺反子结构导致一方面表达 VL-CL(其蛋白质序列在 SEQ ID No. 44 给出)和另一方面表达融合到噬菌体 M13 蛋白质 III 的被 pHEN1 编码的 VH-CHI(其蛋白质序列在 SEQ ID No. 45 中给出)。这里所描述的 4D5 Fab 展示载体的完整序列在 SEQ ID No. 46 中作为核苷酸序列给出。

为了构建其结构环中的残基被随机化的 CL 结构域文库, 制备编码人 κ 恒定结构域(CL)的合成基因, 其中某些密码子被简并密码子例如 NNB (IUPAC 编码, 其中 N 代表 C、G、T 和 A; B 代表 T、C 和 G)置换。将额外的残基插入到序列中也是可能的。在该实施例中, 分

别将 3、4、或 5 个残基插入到残基 127 和 128 之间，而对残基 182-185 和 187-189 进行随机化(Kabat 编号方式)。得到的基因的序列作为核苷酸序列在 SEQ ID No. 47、48 和 49 中给出(127 和 128 之间分别有 3、4 或 5 个插入)，作为氨基酸序列在 SEQ ID No. 50、51 和 52 中给出(字母 X 代表 20 个天然编码的氨基酸的任何一个)。这些核苷酸序列包括用于克隆到 SEQ ID No. 46 的 5'端的 BsiWI 位点和 3'端的 AscI 位点。

为了构建噬菌体展示文库，用限制性酶 BsiWI 和 AscI 酶切 Fab 4D5 展示载体(SEQ ID No. 46)，用制备性琼脂糖凝胶电泳制备大的片段。除去对应于编码野生型 CL 基因的小片段。如上所述的将文库混合物插入(SEQ ID Nos. 47-49)，同样用 BsiWI 和 AscI 酶切并连接到纯化的载体片段中。将连接混合物用例如电穿孔转化到合适的大肠杆菌菌株如 TG1，产生大量单菌落(例如 10^8 、 10^9 或更多)。收集经转化的细菌并用辅助噬菌体(例如 M13K07)释放噬菌体颗粒。用标准程序制备噬菌体颗粒并将其用于文库的淘选。

针对给定的靶标来淘选文库，产生不但与 Her2 结合(由于结合位点由抗体 4D5 的 VH 和 VL 形成)而且与对它们针对性选择的靶标结合的 Fab 片段。

在该实施例中，描述了 Fab 展示文库的设计、制备和使用，其中 CL 结构域的结构环是被修饰的。以相似的方式，可制备和使用在其它结构域如 CH1、VH 或 VL 结构域的结构环中被随机化的文库。

序列:

SEQ ID No. 44

```

      5      10      15      20      25      30
1  M K Y L L P T A A A G L L L L A A Q P A M A D I Q M T Q S P
31 S S L S A S V G D R V T I T C R A S Q D V N T A V A W Y Q Q
61 K P G K A P K L L I Y S A S F L Y S G V P S R F S G S R S G
91 T D F T L T I S S L Q P E D F A T Y Y C Q Q H Y T T P P T F
121 G Q G T K V E I K R T V A A P S V F I F P P S D E Q L K S G
151 T A S V V C L L N N F Y P R E A K V Q W K V D N A L Q S G N
181 S Q E S V T E Q D S K D S T Y S L S S T L T L S K A D Y E K
211 H K V Y A C E V T H Q G L S S P V T K S F N R G E C

```

///

SEQ ID No. 45

```

      5      10      15      20      25      30
1  M K K N I A F L L A S M F V F S I A T N A Y A E V Q L V E S
31 G G G L V Q P G G S L R L S C A A S G F N I K D T Y I H W V
61 R Q A P G K G L E W V A R I Y P T N G Y T R Y A D S V K G R
91 F T I S A D T S K N T A Y L Q M N S L R A E D T A V Y Y C S
121 R W G G D G F Y A M D Y W G Q G T L V T V S S A S T K G P S
151 V F P L A P S S K S T S G G T A A L G C L V K D Y F P E P V
181 T V S W N S G A L T S G V H T F P A V L Q S S G L Y S L S S
211 V V T V P S S S L G T Q T Y I C N V N H K P S N T K V D K K
241 V E P K S C A A A E Q K L I S E E D L N G A A @ T V E S C L
271 A K P H T E N S F T N V W K D D K T L D R Y A N Y E G C L W
301 N A T G V V V C T G D E T Q C Y G T W V P I G L A I P E N E
331 G G G S E G G G S E G G G S E G G G T K P P E Y G D T P I P
361 G Y T Y I N P L D G T Y P P G T E Q N P A N P N P S L E E S
391 Q P L N T F M F Q N N R F R N R Q G A L T V Y T G T V T Q G
421 T D P V K T Y Y Q Y T P V S S K A M Y D A Y W N G K F R D C
451 A F H S G F N E D P F V C E Y Q G Q S S D L P Q P P V N A G
481 G G S G G G S G G G S E G G G S E G G G S E G G G S E G G G
511 S G G G S G S G D F D Y E K M A N A N K G A M T E N A D E N
541 A L Q S D A K G K L D S V A T D Y G A A I D G F I G D V S G
571 L A N G N G A T G D F A G S N S Q M A Q V G D G D N S P L M
601 N N F R Q Y L P S L P Q S V E C R P Y V F G A G K P Y E F S
631 I D C D K I N L F R G V F A F L L Y V A T F M Y V F S T F A
661 N I L H K E S

```

SEQ ID No. 46

```

1  gacgaaaggg cctcgtgata cgcctathtt tataggttaa tgtcatgata ataatggttt
61  cttagacgtc aggtggcact tttcggggaa atgtgcgagg aaccctatt tgtttathtt
121 tctaaataca ttcaaatacg tatccgctca tgagacaata accctgataa atgcttcaat
181 aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgcctt attcctttt
241 ttgcggcatt ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaagatg

```

301 ctgaagatca gttgggtgca cgagtgggtt acatogaact ggatctcaac agcggtaaga
 361 tccttgagag ttttcgcccc gaagaacggt ttccaatgat gagcactttt aaagttctgc
 421 tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatac
 481 actattctca gaatgacttg gttgagtact caccagtcaac agaaaagcat ctacggtag
 541 gcatgacagt aagagaatta tgcagtgtcg ccataacocat gagtataaac actgcccaca
 601 acttacttct gacaacgac ggaggaccga aggagctaac cgcttttttg cacacatgg
 661 gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg
 721 acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg
 781 gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag
 841 ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttttgct gataaatctg
 901 gagccgggtga gcgtgggtct cgcggtatca ttgcagcact ggggcccagat ggtaagccct
 961 cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaatagac
 1021 agatcgcgtga gatagggtcc tcaactgatta agcattggta actgtcagac caagtttact
 1081 catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga
 1141 tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt
 1201 cagaccccggt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaactc
 1261 gctgcttgca acaaaaaaaa ccaccgctac cagcgggtgg ttggttgccg gatacaagagc
 1321 taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc
 1381 ttctagtgtga gccgtagtta ggccaccact tcaagaactc tgtagcaccc cctacatacc
 1441 togctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagctg tgtcttaccg
 1501 ggttggaactc aagacgatag ttaccggata agggcagcgg gtcgggctga acggggggtt
 1561 cgtgcacaca gccagccttg gagcgaacga cctacaccga actgagatac ctacagcgtg
 1621 agcattgaga aagcgcaccg cttcccgaag ggagaaagcc ggacaggtat ccgtaagcg
 1681 gcagggctcg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatctt
 1741 atagtcctgt cgggttctgc cactctgac ttgagcctcg atttttgta tctcgtcag
 1801 gggggcggag cctatggaaa aacgccagca acgogccctt tttaccggtc ctggcctttt
 1861 gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta
 1921 ttaccgcctt tgagttagct gatacgcctc gccgcagccg aacgaccgag cgcagcaggt
 1981 cagtgagcga ggaagcggaa gagcgcocaa tacgcaaacc gcctctcccc gcgcgttggc
 2041 cgattcatta atgcagctgg cagcagaggt ttcccagctg gaaagcgggc agtgagcga
 2101 acgcaattaa tgtgagttag ctcaactcatt aggcacccca ggctttacac tttatgctt
 2161 cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg
 2221 accatgatta cgccaagctt gcatgcaaat tctatttcaa ggagacagtc ataatgaaat
 2281 acctattgccc tacggcagcc gctggattgt tattactcgc ggcccagccg gccatggccg
 2341 atatocagat gaccagctcc ccgagctccc tgtccgcctc tgtggcgat aggttaccac
 2401 tcacctgccc tgccagctcag gatgtgaata ctgctgtagc ctggtatcaa cagaaaccag
 2461 gaaaagctcc gaaactactg atttactcgg catccttccct ctactctgga gtccctctc
 2521 gcttctctgg atccagatct gggacggatt tcactctgac catcagcagt ctgcagccgg
 2581 aagactctgc aacttattac tgtcagcaac attatactac tcctcccagc tctcgacagg
 2641 gtaccaaggt ggagatcaaa cgtacggtgg cggcgcctac tgttctcctc ttcccgcctc
 2701 ctgatgagca gcttaagtct ggaactgcct ctggtgtgtg cctgctgaat aactctctc
 2761 ccagagaggg caaagtacag tggaaaggtg ataacgcctt ccaatcgggt aactcccagg
 2821 agagtgtcac agagcaggac agcaaggaca gcacctacag cctcagcagc accctgacgc
 2881 tgagcaaacg agactacgag aaacacaaag tctacgcctg cgaagtcacc catcaggcc
 2941 tgagctogcc cgtcacaaaag agcttcaaca ggggagagtg ttaataagcc gcgcgcctga
 3001 tcctctacgc cggacgcctc gtggccctag tacgcaagtt cacgtaaaaa gggtatctag
 3061 aggttgagggt gattttatga aaaagaatat cgcatttctt ctgcatcta tgttcgtttt
 3121 ttctattgct acaaatgcat acgctgaggt tcaactagtg gagtctggcg gtggcctggt
 3181 gcagccaggg ggtcactctc gtttgcctg tgcaactctt ggcttcaaca ttaaagacac
 3241 ctatatacac tgggtgcctc agcccccggc taaggcctg gaatgggttg caaggattta
 3301 tcctacgaat ggttatacta gatatgccga tagcgtcaag ggccggttca ctataagcgc
 3361 agacacatcc aaaaaacacag cctacctgca gatgaacagc ctgctgtctg aggcactgc
 3421 cgtctattat tgttctagat ggggagggga cgcttctat gctatggact actggggtca
 3481 aggaaccctg gtcaccgtct cctcggcctc gaccaagggc ccatcggctt tcccctggc
 3541 accctcctcc aagagcacct ctgggggcac agcggccctg ggctgcctgg tcaaggacta
 3601 cttcccggaa ccggtgacgg tgtcgtggaa ctcaggtgcc ctgaccagcg gcgtgcacac
 3661 cttcccggct gtcctacagt cctcaggact ctactccctc agcagcgtgg tgaccgtgcc
 3721 ctccagcagc ttgggcaacc agacctacat ctgcaacgtg aatcacaagc ccagcaacac
 3781 caaggtggac aagaaagttg agcccaaatc ttgtgcggcc gcagaacaaa aactcatctc
 3841 agaagaggat ctgaatgggg ccgcatagac tgttgaagt tgtttagcaa aacctcatac
 3901 agaaaattca tttactaacg tctggaaaga cgacaaaact ttagatcgtt acgctaacta
 3961 tgagggtctg ctgtggaatg ctacagggcgt tgtggtttgt actggtgacg aaactcagtg

```

4021 ttacggtaca tgggttccta ttgggcttgc tatccctgaa aatgaggggt gtggctctga
4081 ggggtggcgg tctgaggggt gcggttctga ggggtggcgg actaaacctc ctgagtacgg
4141 tgatacacct attccgggct atacttatat caaccctctc gacggcactt atccgcctgg
4201 tactgagcaa aaccccgcta atcctaatec ttctcttgag gagtctcagc ctcttaatac
4261 tttcatgttt cagaataata ggttccgaaa taggcagggt gcattaactg tttatacggg
4321 cactgttact caaggcaactg accccgttaa aacttattac cagtacactc ctgtatcatc
4381 aaaagccatg tatgacgctt actggaacgg taaattcaga gactgcgctt tccattctgg
4441 ctttaatgag gatccattcg tttgtgaata tcaaggccaa tcgtctgacc tgcctcaacc
4501 tcctgtcaat gctggcggcg gctctgggtg ttggtctgggt ggcggctctg aggggtggcg
4561 ctctgaggggt ggcggttctg aggggtggcg ctctgaggggt ggcggttccg gtggcggtc
4621 cggttccgggt gattttgatt atgaaaaaat ggcaaacgct aataaggggg ctatgaccga
4681 aatgcccgat gaaaacgcgc tacagtctga cgctaaaggc aaacttgatt ctgtcgctac
4741 tgattacgggt gctgctatcg atggtttcat ttggtgacgtt tccggccttg ctaatggtaa
4801 ttggtgctact ggtgattttg ctggctctaa ttcccaaagtg gctcaagtgc gtgacgggtga
4861 taattcacct ttaatgaata atttccgtca atatttacct tctttgctc agtcggttga
4921 atgtgcacct tatgtctttg gcgctggtaa accatatgaa tttctattg attgtgacaa
4981 aataaactta ttccgtgggt tctttgcggt tcttttataat gttgccacct ttaagtatgt
5041 attttcgacg tttgctaaca tactgcataa ggagctctaa taagaattca ctggccgtcg
5101 ttttacaacg tcgtgactgg gaaaacctcg gcgttaccca acttaatcgc cttgcagcac
5161 atcccccttt cgccagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac
5221 agttgcgcag cctgaatggc gaatggcgcc tgatgcggtg ttttctcctt acgcatctgt
5281 gcggtatattc acaccgcacg tcaaagcaac catagtacgc gccctgtagc ggcgcattaa
5341 gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc
5401 ccgctccttt cgctttcttc ccttctcttc tcgccacgtt cgccggcttt ccccgcaag
5461 ctctaaatcg ggggtccct ttagggttcc gatttagtgc tttacggcac ctgcacocca
5521 aaaaacttga tttgggtgat ggttcacgta gtgggcccac gccctgatag acggtttttc
5581 gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa
5641 cactcaaccc tatctcgggc tattcttttg atttataagg gattttgccc atttcggcct
5701 attgggtaaa aaatgagctg atttaacaaa aatttaacgc gaattttaac aaaatattaa
5761 cgtttacaat tttatgggtg actctcagta caatctgctc tgatgccgca tagttaagcc
5821 agccccgaca cccgccaaca ccgctgacg cgcctgacg ggcttctctg ctccccgcat
5881 ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg ttttcaccgt
5941 catcaccgaa acgcgcga

```

SEQ ID No. 47

```

1 cgtacggtgg cggcgccatc tgtcttcac tcccgccat ctgatgagca gcttaagtct
61 nbnbnbnbnbg gaactgcctc tgttgtgtgc ctgctgaata acttctatcc cagagagggc
121 aaagtacagt ggaagtgga taacgcctc caatcgggta actcccagga gagtgtcaca
181 gagcaggaca gcaaggacag cactacagc ctacagcagca ccctgacgct gnnbnbnbn
241 nbnbtacnbn nbnnbaaagt ctacgctgtc gaagtcacc atcagggocct gagctcgccc
301 gtcacaaaaga gttcaacag gggagagtgt taataaggcg cgcc

```

SEQ ID No. 48

```

1 cgtacggtgg cggcgccatc tgtcttcac tcccgccat ctgatgagca gcttaagtct
61 nbnbnbnbnbn nbggaactgc ctctgttgtg tgctgctga ataacttcta tcccagagag
121 gccaaagtac agtggaagggt ggataacgcc ctccaatcgg gtaactccca ggagagtgtc
181 acagagcagg acagcaagga cagcacctac agcctcagca gcacctgac gctgnbnbnbn
241 nbnbnbtacn nbnnbnbaa agtctacgcc tgcgaaagtca ccatcaggg cctgagctcg
301 ccgctcaciaa agagcttcaa caggggagag tgtaataaag gcgcgcc

```

SEQ ID No. 49

```

1 cgtacggtgg cggcgccatc tgtcttcac tcccgccat ctgatgagca gcttaagtct
61 nbnbnbnbnbn nbnnbggaac tgcctctgtt gtgtgcctgc tgaataactt ctatcccaga
121 gaggccaaag tacagtggaa ggtggataac gccctccaat cgggtaactc ccaggagagt
181 gtcacagagc aggacagcaa ggacagcacc tacagcctca gcagaccct gacgctgnbn
241 nbnbnbnbnbt acnbnbnbn baaagtctac gcctgcgaag taccatca gggcctgagc
301 tcgcccgtca caaagagctt caacagggga gagtgttaat aaggcgcgc

```

SEQ ID No. 50

```
      5      10      15      20      25      30
1  R T V A A P S V F I F P P S D E Q L K S X X X G T A S V V C
31 L L N N F Y P R E A K V Q W K V D N A L Q S G N S Q E S V T
61 E Q D S K D S T Y S L S S T L T L X X X X Y X X X K V Y A C
91 E V T H Q G L S S P V T K S F N R G E C
```

SEQ ID No. 51

```
      5      10      15      20      25      30
1  R T V A A P S V F I F P P S D E Q L K S X X X X G T A S V V
31 C L L N N F Y P R E A K V Q W K V D N A L Q S G N S Q E S V
61 T E Q D S K D S T Y S L S S T L T L X X X X Y X X X K V Y A
91 C E V T H Q G L S S P V T K S F N R G E C
```

SEQ ID No. 52

```
      5      10      15      20      25      30
1  R T V A A P S V F I F P P S D E Q L K S X X X X X G T A S V
31 V C L L N N F Y P R E A K V Q W K V D N A L Q S G N S Q E S
61 V T E Q D S K D S T Y S L S S T L T L X X X X Y X X X K V Y
91 A C E V T H Q G L S S P V T K S F N R G E C
```

附图说明

图 1: SEQ ID No. 1 和 SEQ ID No. 2 的氨基酸序列的序列比对。

图 2: 编码抗 TNF- α 骆驼 VHH 结构域的合成基因的编码序列, 以及以一个字母的氨基酸代号来表示的翻译。用于克隆的限制性位点在该核苷酸序列中用下划线表示。(插入的氨基酸没在该序列中给出。)

图 3 显示 PCR 反应和连接步骤的示意图。水平箭头显示 PCR 引物的位置和方向, 垂直线分别显示 NcoI、XhoI 和 NotI 位点的位置(从左到右)。

图 4: 用于构建实施例 6 中所述文库的 PCR 反应流程图。

图 5: scFV 2F5 野生型合成基因文库 1 (VH-接头-VL)和翻译, 2 个环被随机化。

图 6: scFV 2F5 野生型合成基因文库 2 (VH-接头-VL)和翻译, 2 个环被随机化

图 7: scFV 2F5 野生型合成基因文库 3 (VH-接头-VL)和翻译, 2 个环被随机化

图 8: scFV 2F5 野生型合成基因和翻译(VH-接头-VL)

```

Query 2  VQLQESGGGLVQPPGSLRLSCAASGRTFSDHSGYTYTIGWFRQAPGKEREFVARIYWSSG 61
          VQL  EGGG  VQ  GGLRLSCAASG  S  +      +GWERQAPGKERE  VA  +  +  G
Sbjct 2  VQLVESGGGSVQAGGSLRLSCAASGYIASINY-----LGWFRQAPGKEREGVAAVSPAGG 56

Query 62  NTTYADSVKGRFAISRDIAKNTVDLTMNLEPEDTAVYYC-AARDGIPTSRSVESYNYWG 120
          YYADSVKGRF  +S  D  A+NTV  L  MN+L+PEDTA+YYC  AAR  G      +  YNYWG
Sbjct 57  TPYYADSVKGRFTVSLDNAENTVYLQMNLSLKPEDTALYYCAAAARQGWYIPLNSYGNYWG 116

Query 121  QGTQTVTVSS 129
          QGTQTVTVSS
Sbjct 117  QGTQTVTVSS 125

```

图 1

```

+3 P W Q V Q L Q E S G G L Y Q P C G S L R L S C A A S G R T F S D
1  CCATGGCAG TTAGCTGCA GGAAGGGGT GGCGGCTGG TCGACCTGG CGGACCCCTG CGTCTAGCT GTGCGGCAG CCGCGTACC TTTAGCGCC
  GGTACCGTTC AAGTCGAGT CCTTTCGCCA CCGCCGGACC AGGTCCGACC GCGCTGGAC GCGAATCGA CACCCCGTTC GCGCGCAGG AATTCGGCTGG
+3 H S G Y T Y T I G W F R Q A P G K E R E F V A R I Y W S S G N T Y Y
101 ATAGGGGTA CACCTATACC ATTGGCTGGT TTCGTCAGGC GCCAGGAAA GAACGTGAT TTGTGGCGC TATTTACTGG AGCAGGCGCA ATACTTATTA
  TATCCCGAT GTGGATGG TAACCGACCA AAGCAATCCG CGGTCCITTT CTTCACITTA AHCACCCGCG ATAAATGACC TCGTCCCGT TATGGATAT
+3 A D S V K G R F A I S R D I A K N T V D L T M N N L E P E D T A V
201 TGGGATAGC GTGAAAGCC GTTCCGGAT TAGCCGGAC ATTGCCAGA ACACGGTGA TCTTACGATG AACCAACTGG AGCCCGAAGA CACAGCCGTG
  ACGCCTATCG CACTTCCGG CAAAGCGTA ATCGGGCTG TACCGTTCT TGTCCTACT AGRATGCTAC TTGTTGGACC TCGGGCTTCT GTGTGGGCAC
+3 Y Y C A A R D G I P T S R S V E S Y N Y M G Q G T Q V T V S S A A
301 TATTAATGCG CGCTCGGA TGGCAITCCG ACCAGCCGTA GGTGGANAG CTACATATAC TGGGCGCAGG GACCCAGGT GACCTCAGC TCTGGGCGG
  ATAAATACGC GCCGAGCCCT ACCGTAAAGC TGGTGGGCAI GCGACCTTC GRATTTAATG ACCCGGTYCC GGTGGGTCCA CTGGCAGTGG AGACCGCGGC
+3
401 C

```

图 2

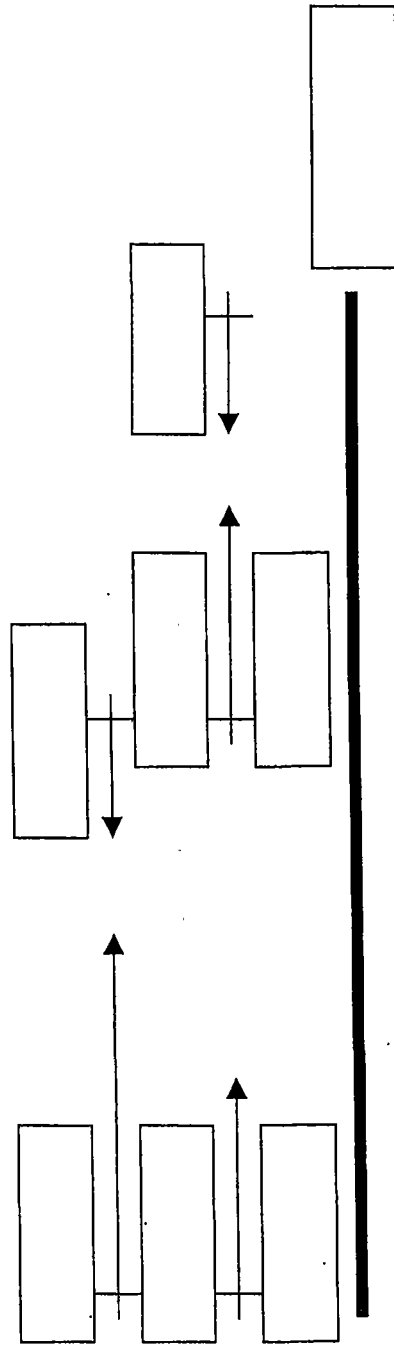


图3

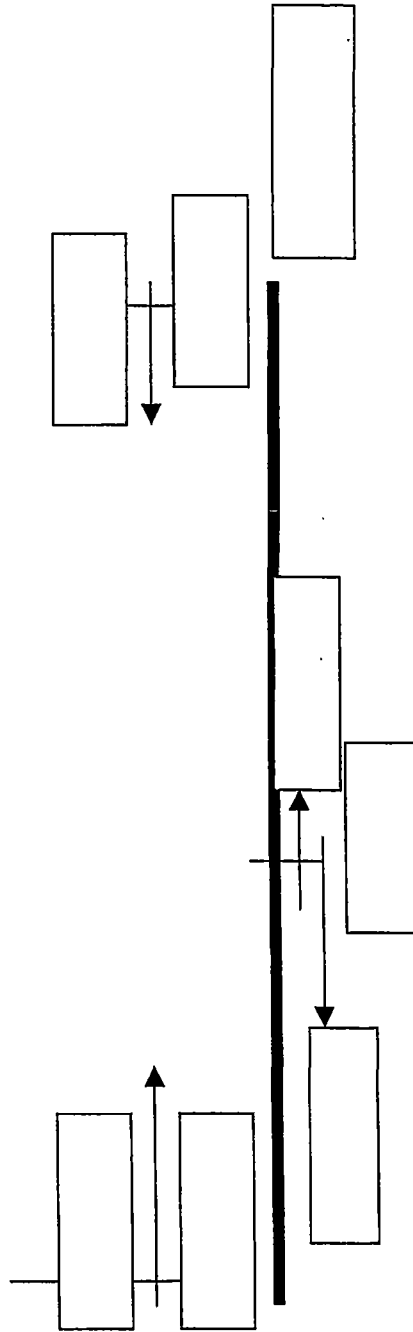


图 4

```

+3 P W R I T L K E S G P P L V K P T Q T L T L T C S F S G F S L S D
1 CCATGGCGGA TCACCTGAA AGAGAGTGA CCCCCTGG TGAACCTAC CCAGACCTG ACCTGACTT GCTCAITTAG CCGCTTAGC CTGAGCGGATT
GGTACCGCCT AGTGGGACTT TCTCTACCT GGGGGGACC ACTTTGGATG GGTCTGGAC TGAGACTGAA CAGATRAATC GCCGAATCG GACTCGCTAA

+3 F G V G V G I R Q P P G K A L E W L A I I Y S D D D K R Y S P X X
101 TTGGCTCGG CGTTGGTGG ATTCGCCAGC CTCGGGCAA AGCCCTGAA TGGCTGGCA TCATCTACT CGATGATGAC AAGCTTATA GCCCNSNN
AACCCGACC GCACCAACC TAAGCGGTG GAGGCCGTT TCGGACCTT ACCGACCGT AGTAGATGAG GCTACTACTG TTCGCAATAT CCGGGAGCGA

+3 X X X L T I T K D T S K N Q V V L V M X X V S P V D T A T Y F C A
201 SNNNSNNS CTGACCAICA CCAAGATAC GAGCAAGAAC CAGGTGTTT TGGTAATGN SNNSTGAGC CCCGTGACA CCGGACTTA TTTCGTGCC
CTTATGGCA GACTGGTAGT GGTTCATG GTTTCATG GTCCACAAA ACCATTACTG GGCACACTCG GGCACGCTGT GCGCTGAT AAAGACACGG

+3 H R R G P T L F G V P I A R G P V N A M D V W G Q G I T V T I S
301 CATCGTCGT GTCCGACCAC CCTGTTGGT GTCCGATG CAGCGGTCC CGTGAATGG ATGGATGTT GGGGGCAGG GATTACCGTG ACCATTTCAI
GTAGCAGCAC CAGGCTGTTG GGACAAACCA CAGGCTAAC GTCCGCCAG GCACTACGC TACCTACACA CCCCCCTCC CTAATGGCAC TGGTAAAGTA

+3 S G G G S G G G S G G S A L O L T O S P S S L S A S V G D R
401 CCGGTGAGG TGTAGTGA GGGGTGGT CAGCGGTG CCGCTCCGC TTACACTGA CCGAGACCC GTCAGTTTG AGCCAGCGG TGGGCGATCG
GGCCACTCC ACCATCACT CCCCACCCA GTCCGCCACC GCGGAGGG ANGTGACT GCCTCTCGG CAGATCAAC TCGCTTCG ACCCGTAGC

+3 I T I T C R A S Q G V T S A L A W Y R Q K P G S P P Q L L I Y D A
501 TATTACAAT ACCTGTCGG CGAGCCRAGG TGTACTCC CCGTGGCT GGTATCGTCA GAAACCCGG AGCCCGCAC AGTTGTTGAT CTACGATCG
ATAATGTAA TGGACAGCCC GTCGGTTC ACAATGGAG CCGGACCGGA CCAVACAGT CTTTGGGCC TCGGGCGGTG TCAACAAC TA GTCTACCG

+3 S S L E S G V P S R F S G S G S G T E F T L T I S T L R P E D F A
601 TCCTACTGG AATCAGGGT CCTAGCCGC TTTCCGGT CCGCAGCG CACGGAAT ACANTGACA TAAGCACCT GCGTCCGGA GATTTGCA
AGGAGTACC TTAGTCCCA GGTATCGCG AAGGCCCA GCGCTGCC GTGCTTAA TGTACTGTT ATTCGTGGA CCGAGGCCCT CTAACAACGGT

+3 T Y Y C Q Q L H F Y P H T F G G G T R V D V R R T V A A A A
701 CCTATTATTG CCAACAGCTG CACTTTATC CCCATACCT CCGTGGGGG ACGCGGTTG ACCTGCTG TACCGTAGCT GCTGGCGCG C
GGATAATAC GGTTCGAC GTGAARATG GGTATGHA GCCACCCCC TGGCCCAAC TGCACGCAGC ATGGCATCGA CGAGCCCGC G

```

图 5

```

+3 P W R I T L K E S G P P X X X P X X X L T L T C S F S G F S L S D
1 CCATGGCGGA TCACCTGAA AGAGAGTGA CCCCCNNSN NSNNSCTNN SNNNSCTG ACCTGACTT GCTCAWTTAG CGGCTTAGC CTGAGCGATT
  GGTACCGCCT AGTGGACTT TCTTCACCT GGGGGGACC ACTTTGGATG GGTCTGGGAC TGAGACTGAA CAGTAAATC GCCGAAATC GACTCGCTAA

+3 F G V G V G W I R O P P G K A L E W L A I I Y S D D D K R Y S P S L
101 TTGGGTCGG CGTTGGTTGG AATCCCAAGC CTCGCGGAA AGCCCTGAA TGCTGGCCA TCATCTACTC CGATGATGAC AAGCCTTATA GCCCTCCTCCT
  AACCCGAGCC GCACCAACC TAAGCGGTG GAGGGCCGTT TCGGGACCTT ACCGACCGGT AGTAGATGAG GCTACTACTG TTCGCAATAT CGGGAGGCGA

+3 N T R L T I T K D T S K N Q V V L V M X X V S P V D T A T Y F C A
201 GAATACCCCT CTGACATCA CCAAGATAC GAGCAAGAC CAGGTGGTTT TGATAATGN SNNSTGAGC CCCGTGACA CGCGACTTA TTTCTGTGCC
  CTTATGGGCA GACTGGTAGT GGTTCATG CTCGTTCTATG GTCCACCAA ACCATTAATG GGCACACTCG GGGCAGCTGT GCGCTGAT AARGCACGG

+3 H R R G P T T L F G V P I A R G P V N A M D V W G Q G I T V T I S
301 CATCGTCGNG TTCGACCAC CCTGTTTGGT GTGCCGATG CAGCGGTCC CBTGATGCG BTGGATGTT GGGGCGAGG GATPACCGTG ACCATTTTCA
  GTAGCAGCAC CAGGCTGTTG GGACAAACCA CAGGCTAAC GTGGGCCAGG GCACTTACGC TACCTACACA CCCCCGTCCC CTAATGGCAC TGGTAAAGTA

+3 S G G G S G G G G S G G A L Q L T O S P S S L S A S V G D R
401 CCGGTGGAG TGGTAGTGA GGGGTGGGT CAGCGGTGG CGGCTCGCC TTACACTGA CGCAGAGCCC GTCCTAGTTG AGCCGACGG TGGCGGATCG
  GGGCACCTCC ACCATCACCT CCCCCACCA GTCCGCCACC GCGGAGCGG AATGTTGACT GCGTCTCGGG CAGATCAAC TCGCGTTCCG ACCCGCTAGC

+3 I T I T C R A S Q G V T S A L A W Y R Q K P G S P P Q L L I Y D A
501 TATTACAAT ACCTGTGGG CGAGCCAAG TGTACCTCC GCCGTGCCCT GGTATGTC A GAACCCGGG AGCCCGCCAC AGTTGTTGAT CTAGCATCG
  ATATGTTAA TGGACAGCC GCTCGTTC ACANTGGAG CCGGACCGGA CCATAGCAGT CTTTGGGCCC TGGGGGGTGC TCAACACTA GATGCTACG

+3 S S L E S G V P S R F S G S G S G T E F T L T I S T L R P E D F A
601 TCCTCACTGG AATCAGGGT CCCTAGCCGC TTTTCCGGT CCGCAGCGG CAGCGAATTT ACATTCACCA TAAGCACCT CCGTCCGAA GATTTTGCCA
  AGGAGTGACC TTAGTCCCA GGGATCGGG AABAGGCCCA GCGGTCGCC GTGCCITAAA TGTAACTGTT ATTCGTGGGA CGCAGGCCCT CTAAACCGGT

+3 T Y Y C Q O L H F Y P H T F G G G T R V D V R R T V A A A
701 CCTATATTG CCACAGCTG CACTTTTATC CCCATACCTT CGGTGGGGG ACGGGGTTC AGTGGGTGC TACCCTAGCT GCTGGGGCCG C
  GGATAATAAC GGTTCGAC GTGAAATAG GGGTATGAA GCCACCCCC TCGCCCCAC ATGGCATGA CGACCCCGC G

```

图 6

```

+3 P W R I T L K E S G P P X X X P X X X L T L T C S F S G F S L S D
1 CCATGGCGGA TCACCTGAA AGAGATGGA CCCCNNNS NNNSCTNN SNNNSCTG ACTCTGACTT GCTCATTTAG CCGCTTTAGC CTGAGGGATT
GGTACCGCCT AGTGGACCTT TCTCTCACCT GGGGGGACC ACTTTGGATG GECTTGGGAC TGAGACTGAA CGAGTAATAC GCCGAAATCG GACTCGCTAA

+3 F G V G V G W I R Q P P G K A L E W L A I I Y S D D D K R Y S P X X
101 TTGGGTCGG CFTTGTGG ATTCCGAGC CTCGCGGAA AGCCCTGGAA TGGTGCCCA TCATCTACTC CGATGATGAC AAGGTTATA GCCCCNNNS
AACCAGCC GCARCRACC TAAGCGTCC GAGGCGCTT TCGGGACCTT ACCGACCGGT ACTAGATGAG GCTACTACTG TTCCGAATAT CCGGGAGCGA

+3 X X X L T I T K D T S K N Q V V L V M X X V S P V D T A T Y F C A
201 SNNNSNNS CTGACCATCA CCAAATAC GAGCAAGAAC CAGGTGGTTT TGGTAATGNN SNNSTGAGC CCGTCTGACA CCGGACTTA TTCTGTGCC
CTTATGGGCA GACTGTAGT GGTTCATG CTCGTTCTTG GTCCACAAA ACCATTACTG GGCACACTCG GGCAGCTGT GCGCTGAAI AARGACAGG

+3 H R R G P T L F G V P I A R G P V N A M D V W G Q G I T V T I S
301 CATCTCTG GTCCGACC CTTGTGGT GTCCGATTG CAGCGGTC CGTGAATCG ATGATGATG TGGGGCAGG GATTACCGTG ACCATTGAT
GTAGCRGAC CAGGCTGTC GGCARACCA CAGGCTAAC CAGCGCAGG GCACTTACGC TACTACACA CCCCCCTCC CTAATGGCAC TGGTAAAGTA

+3 S G G G S G G G S G G G S A L Q L T Q S P S S L S A S V G D R
401 CCGTGGAG TGGTAGTGA GGGGTGGT CAGCGGTGG CAGCGGTGG CCGTCCGC TTACAATGA CCGAGGCC GTCTAGTTG AGCCRAGCG TGGGCGATCG
GCCCCCTCC ACCATCACCT CCCCACCCA CTCGCGCAC GCGAGGCG AATGTTGACT GCGTCTCGG CAGATCABC TCGGTTCC ACCCGTAG

+3 I T I T C R A S Q G V T S A L A W Y R Q K P G S P P O L L I Y D A
501 TATTACAATT ACCTGTCGG CGAGCAAG TGTACTCC GCCTGGCT GGTATCTCA GAACCCGGG AGCCCCAC AGTTGTTGAT CTACGATCGG
ATAATGTTAA TGGACAGCC GCTCGTTC ACAATGGAG GGGACCGGA CCATAGCAGT CTTGGGCC TCGGGCGTG TCAARACTA GATGCTACG

+3 S S L E S G V P S R F S G S G T E F T L T I S T L R P E D F A
601 TCTCACTGG AATCAGGCT CCGTACCGG TTTCCGGT CCGCAGCG CAGGAAAT ACATTGACA TAAGCACCTT GCGTCCGAA GATTTGCCA
AGGATGACC TTAGTCCCA GGCATCGCG AAGAGCCCA GCGCTCCG CTGCGTTAAA TGTRACTGT ATTGTTGGA CCGAGGCTT CTAACCGT

+3 T Y Y C Q Q L H F Y P H T F G G G T R V D V R R T V A A A
701 CCTATTATG CCAACAGCT CACTTTATC CCCATACCTT CCGTGGGGG ACGGGGTTG AGTGGCTCG TACCCTAGCT GCTGGGCGG C
GGATAATAAC GGTGTCGAC GTGAAATAG GCGATGGA GCCACCCCC TCGCCCAAC ATGGCATCGA CGAGCCCGG G

```

图 7

```

+3 P W R I T L K E S G P P L V K P T O T L T L T C S F S G F S L S D
1 CCATGGGGA ACCTCCGAA AGAGTGGG CCCCCTGG TGAACCTAC CAGACCCCTG ACCTGACTT GCTCAITTAG CGGCTTAGC CTGAGCGATT
GGTACCCTT AGTGGACTT TCTTCACCT GGGGGGACC ACTTTGGATG GGTCTGGGAC TGAGACTGAA CGAGTAAAC GCCGAATCG GACTCGCTAA

+3 F G V G V G W I R Q P P G K A L E W L A I I Y S D D D K R Y S P S L
101 TTGGCTGG CGTTGGTTGG ATTCGCCAGC CTCCCGCAA AGCCCTGGAA TGGCTGGCCA TCATCTACT CGATGATGAC AAGCGTTATA GCCCTCGCT
AACCGCAGCC GCACCAACC TAAGCGTCC GAGGGCGTT .TCGGACCTT ACCGACCGGT AGTAGATGAG GCTACTACTG TTCCGATAT CGGGAGCGA

+3 N T R L T I T K D T S K N Q V V L V M T R V S P V D T A T Y F C A
201 GAATACCCCT CTGACCAATCA CCAAGATAC GAGCAGAAC CAGGTGGTTT TGGTAATGAC CCGTGTGAC CCGCTGACA CCGCGACTTA TTCTGTGCC
CTTATGGGCA GACTGTAGT GGTTCATG CTGTTCTATG CTCACCAAAA ACCAATCTG GGCACACTG GGCAGCTGT GCGGCTGAAT AAGACACCG

+3 H R R G P T L F G V P I A R G P V N A M D V W G Q G I T V T I S
301 CATCTCTG GTCCACAC CCTGTGTGT GTCCGATG CACGGTCC CGTGAATCG ATGGATGT GGGGGAGG GATACCGTG ACCATTTCAT
GTAGCAGCAC CAGGCTGGT GACARACCA CAGGCTAAC GTGGCCGAG GCATCTAGC TACCTACACA CCCCCTCC CTRATGGCAC TGCTAAGTA

+3 S G G G S G G G S G G S A L O L T O S P S L S A S V G D R
401 CCGGTGAGG TGTATGGA GGGGTGGT CAGCGGTG CCGTCCGC TTACACTGA CCGAGAGCCC GTCTAGTTG AGCGAAGCG TGGGGATCG
GGCCACTCC ACCATCACT CCCCACCA GTCCGACC GCGAGGCG AATGTGACT GCTCTCGGG CAGATCAAC TGCGCTCC ACCCGTAGC

+3 I T I T C R A S O G V T S A L A W Y R Q K P G S P P Q L I Y D A
501 TATTCANAT ACCTGTGGG CGAGCCAGG TGTACTCC GCCCTGGCT GGTATCTCA GAACCCGG AGCCGCCAC AGTTGTGAT CTACGATCG
ATATGTTRA TGGACAGCC GTCGTTCC ACATGGAG CCGACCGGA CCAATAGCT CTTTGGGCC CCGGGGGTG TCAACACTA GATGCTAGC

+3 S S L E S G V P S R F S G S G S G T E F T L T I S T L R P E D F A
601 TCCTACTG RATCAGGT CCCTAGCGG TTTTCGGGT CCGCAGCG CACGCAAT ACATGACA TAAGCACCT GCGTCCGAA GATTTGCA
AGGATGACC TTATGCCCA GGTTCGGG ABAAGCCCA GCGCTGCC GTCGTTAAA TGTACTGT ATTCTGGGA CCGAGGCTT CTRAACGGT

+3 T Y Y C Q Q L H F Y P H T F G G G T R V D V R R T V A A A
701 CCTATTATG CCAACGCT CACTTTATC CCAATACCT CCGTGGGG AGCGGTTG AGTGTGCTG TACCTAGCT GCTCGGCC C
GGATATAC GGTGTGAC GTGAATAG GGCTATGAA GCCACCCC TCGGCCAAC TCCACCGAGC ATGGCATCGA CGACGCCGC G

```

图 8