

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C07K 16/00

C12N 15/00 C12N 15/63

G01N 33/53



[12] 发明专利申请公开说明书

[21] 申请号 02818596.X

[43] 公开日 2004 年 12 月 29 日

[11] 公开号 CN 1558916A

[22] 申请日 2002.7.31 [21] 申请号 02818596.X

[30] 优先权

[32] 2001. 8. 1 [33] US [31] 09/921,144

[86] 国际申请 PCT/US2002/024582 2002.7.31

[87] 国际公布 WO2003/012069 英 2003.2.13

[85] 进入国家阶段日期 2004.3.23

[71] 申请人 埃比玛克思公司

地址 美国加利福尼亚州

[72] 发明人 王材力 钟平宇 刘胜江 罗培志

李胜峰 王欣慰

[74] 专利代理机构 中原信达知识产权代理有限责
任公司

代理人 丁业平 王维玉

权利要求书 12 页 说明书 62 页 序列表 13 页
附图 31 页

[54] 发明名称 产生嵌合异源多聚体的组合物和方法

[57] 摘要

本发明提供一种把单体多肽特异性组装成异源多聚体的方法。尤其可用本技术来获得具遗传多样性的异源多聚体如抗原结合单位库。本发明也提供由本发明所述的技术组装的非单链和单链抗原结合单位。本发明也提供重组多核苷酸，载体，宿主细胞以及制造受试抗原结合单位的试剂盒。本发明进一步提供使用受试抗原结合单位的方法。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种非单链抗原结合单位，包括：

(a)轻链(L)多肽，它包含的轻链(L)可变区和第一异源二聚化序列
5 框内融合在一起；

(b)重链(H)多肽，它包含的重链(H)可变区和第二异源二聚化序列
框内融合在一起；

其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和
力和二聚化；以及其中至少一种异源二聚化序列在生理缓冲条件和/或
10 生理体温下基本上不能形成同源二聚体。

2. 根据权利要求 1 所述的非单链抗原结合单位，其中第一和第二
异源二聚化序列在生理缓冲条件和/或生理体温下基本上都不能形成
同源二聚体。

3. 一种非单链抗原结合单位，包括：

(a)轻链(L)多肽，它包含的轻链(L)可变区和第一异源二聚化序列
框内融合在一起；

(b)重链(H)多肽，它包含的重链(H)可变区和第二异源二聚化序列
20 框内融合在一起；

其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和
力和二聚化，所述第一和第二异源二聚化序列包含能介导受体异源二
聚化的异源二聚受体序列。

4. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中第一
25 和第二异源二聚化序列形成卷曲螺旋二聚体。

5. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中 L 链
和 H 链多肽通过这两段异源二聚化序列的非共价配对亲和力二聚化。

6. 根据权利要求 4 所述的非单链抗原结合单位，其中 L 链多肽进一步包括其两侧翼为 L 链可变区和第一异源二聚化序列的 flexon。

5 7. 根据权利要求 4 所述的非单链抗原结合单位，其中 H 链多肽进一步包括其两侧翼为 H 链可变区和第二异源二聚化序列的 flexon 序列。

8. 根据权利要求 4 所述的非单链抗原结合单位，其中第一和第二异源二聚化序列均与至少一个半胱氨酸残基连接。

10 9. 根据权利要求 4 所述的非单链抗原结合单位，其中抗原结合单位是多价的。

15 10. 根据权利要求 4 所述的非单链抗原结合单位，其中抗原结合单位是多特异性的。

11. 根据权利要求 10 所述的非单链抗原结合单位，其中抗原结合单位是双特异性的。

20 12. 根据权利要求 10 所述的非单链抗原结合单位，其中抗原结合单位是三特异性的。

13. 根据权利要求 4 所述的非单链抗原结合单位，其中 L 链多肽包括衍生自人轻链的序列。

25 14. 根据权利要求 4 所述的非单链抗原结合单位，其中 H 链多肽包括衍生自人重链的序列。

30 15. 根据权利要求 4 所述的非单链抗原结合单位，其中抗原结合单位交联到化学官能基团上。

16. 根据权利要求 15 所述的非单链抗原结合单位，其中所述官能基团选自信号肽，增强免疫反应活性的试剂，利于和固相支持物偶联的试剂，疫苗载体，生物反应效应物，毒素，可检测标记，顺磁标记和药物。

17. 根据权利要求 4 所述的非单链抗原结合单位，其中第一和第二异源二聚化序列分别来源于 GABA_B 受体 1 和 GABA_B 受体 2 的 C 端序列。

18. 根据权利要求 4 所述的非单链抗原结合单位，其中第一异源二聚化序列包括 GABA_B 受体 1 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.2 中描述的可比长度的线性肽序列基本相同；以及其中第二异源二聚化序列包括 GABA_B 受体 2 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.4 中描述的可比长度的线性肽序列基本相同，其中第一和第二异源二聚化序列都连接在半胱氨酸残基上。

19. 根据权利要求 4 所述的非单链抗原结合单位，其中第一异源二聚化序列包括 GABA_B 受体 1 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.4 中描述的可比长度的线性肽序列基本相同；以及其中第二异源二聚化序列包括 GABA_B 受体 2 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.2 中描述的可比长度的线性肽序列基本相同，其中第一和第二异源二聚化序列都连接在半胱氨酸残基上。

20. 一种单链抗原结合单位，包括轻链(L)可变区和重链(H)可变区，它们由第一和第二异源二聚化序列连接，而这两段序列从一个区的 C 端跨越到另一个区的 N 端，其中这两个可变区通过第一和第二异源二聚化序列的配对亲和力形成一个分子内二聚体；以及其中在生理缓冲条件和/或生理体温下，至少一种异源二聚化序列基本上不能形成同源二聚体。

21. 根据权利要求 20 所述的单链抗原结合单位，其中第一和第二异源二聚化序列在生理缓冲条件和/或生理体温下基本上都不能形成同源二聚体。

5

22. 一种单链抗原结合单位，包括轻链(L)可变区和重链(H)可变区，它们由第一和第二异源二聚化序列连接，而这两段序列从一个区的 C 端跨越到另一个区的 N 端，其中两个可变区通过第一和第二异源二聚化序列的配对亲和力形成一个分子内二聚体；所述第一和第二异源二聚化序列包含能介导受体异源二聚化的异源二聚受体序列。

10

23. 根据权利要求 20 或 22 所述的单链抗原结合单位，其中第一和第二异源二聚化序列形成卷曲螺旋二聚体。

15

24. 根据权利要求 20 或 22 所述的单链抗原结合单位，其中第一和第二异源二聚化序列通过非共价配对亲和力二聚化。

25. 根据权利要求 23 所述的单链抗原结合单位，其中抗原结合单位交联到化学官能基团上。

20

26. 根据权利要求 23 所述的单链抗原结合单位，其中 L 链可变区包括衍生自人轻链的序列。

27. 根据权利要求 23 所述的单链抗原结合单位，其中 H 链可变区包括衍生自人重链的序列。

25

28. 根据权利要求 23 所述的单链抗原结合单位，其中第一和第二异源二聚化序列分别来源于 GABA_B 受体 1 和 GABA_B 受体 2 的 C 端序列。

30

29. 根据权利要求 23 所述的非单链抗原结合单位，其中第一异源二聚化序列包括 GABA_B 受体 1 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.2 中描述的可比长度的线性肽序列基本相同；以及其中第二异源二聚化序列包括 GABA_B 受体 2 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.4 中描述的可比长度的线性肽序列基本相同，其中第一和第二异源二聚化序列都连接在半胱氨酸残基上。

30. 根据权利要求 23 所述的非单链抗原结合单位，其中第一异源二聚化序列包括 GABA_B 受体 1 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.4 中描述的可比长度的线性肽序列基本相同；以及其中第二异源二聚化序列包括 GABA_B 受体 2 中至少 30 个氨基酸残基的多肽，该多肽和 SEQ ID NO.2 中描述的可比长度的线性肽序列基本相同，其中第一和第二异源二聚化序列都连接在半胱氨酸残基上。

31. 一种重组多核苷酸，含有编码权利要求 1 所述的 L 链多肽的编码序列。

32. 一种重组多核苷酸，含有编码权利要求 1 所述的 H 链多肽的编码序列。

33. 一种重组多核苷酸，含有编码权利要求 1 所述的 L 链多肽的第一编码序列和编码权利要求 1 所述的 H 链多肽的第二编码序列。

34. 一种重组多核苷酸，含有编码权利要求 3 所述的 L 链多肽的编码序列。

35. 一种重组多核苷酸，含有编码权利要求 3 所述的 H 链多肽的编码序列。

36. 一种重组多核苷酸，含有编码权利要求 3 所述的 L 链多肽的

第一编码序列和编码权利要求 3 所述的 H 链多肽的第二编码序列。

37. 一种重组多核苷酸，含有编码权利要求 20 所述的单链抗原结合单位的编码序列。

5

38. 一种重组多核苷酸，含有编码权利要求 22 所述的单链抗原结合单位的编码序列。

10

39. 一种载体，含有权利要求 31-38 中任一项所述的重组多核苷酸。

40. 根据权利要求 39 所述的载体，其中载体是表达载体。

41. 根据权利要求 39 所述的载体，其中载体是噬菌体展示载体。

15

42. 一种编码抗原结合单位库的表达载体的选择文库，包含多于 1 个的权利要求 39 所述的载体。

20

43. 根据权利要求 39 所述的选择文库，其中载体是噬菌体展示载体。

44. 一种宿主细胞，含有权利要求 31-38 中任一项所述的重组多核苷酸。

25

45. 根据权利要求 44 所述的宿主细胞，其中编码 L 链多肽的重组多核苷酸和编码 H 链多肽的重组多核苷酸出现在单一载体中。

46. 根据权利要求 44 所述的宿主细胞，其中编码 L 链多肽的重组多核苷酸和编码 H 链多肽的重组多核苷酸出现在不同载体中。

30

47. 根据权利要求 44 所述的宿主细胞，其中宿主细胞是真核细胞。

5 48. 根据权利要求 44 所述的宿主细胞，其中宿主细胞是原核细胞。

49. 一种产生非单链抗原结合单位的方法，包括如下步骤：

10 (a)在宿主细胞中表达编码轻链(L)多肽的第一重组多核苷酸，所述轻链多肽包括框内融合到第一异源二聚化序列上的 L 链可变区，和表达编码重链(H)多肽的第二重组多核苷酸，所述重链多肽包括框内融合到第二异源二聚化序列上的 H 链可变区，其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，以及其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体；和可选择地

15 (b)分离在宿主细胞中表达的抗原结合单位。

50. 根据权利要求 49 所述的方法，其中第一和第二异源二聚化序列在生理缓冲条件和/或生理体温下基本上都不能形成同源二聚体。

20 51. 一种产生非单链抗原结合单位的方法，包括如下步骤：

(a)在宿主细胞中表达编码轻链(L)多肽的第一重组多核苷酸，所述轻链多肽包括融合到第一异源二聚化序列上的 L 链可变区，和表达编码重链(H)多肽的第二重组多核苷酸，所述重链多肽包括融合到第二异源二聚化序列上的 H 链可变区，其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，所述第一和第二异源二聚化序列包含能介导受体异源二聚化的异源二聚受体序列；和可选择地

25 (b)分离在宿主细胞中表达的抗原结合单位。

30 52. 根据权利要求 49 或 51 所述的方法，其中在步骤(a)中表达的非单链抗原结合单位展示在宿主细胞表面。

53. 根据权利要求 49 或 51 所述的方法，其中在步骤(a)中表达的非单链抗原结合单位展示在噬菌体粒子上。

5 54. 根据权利要求 49 或 51 所述的方法，其中宿主细胞是真核细胞。

55. 根据权利要求 49 或 51 所述的方法，其中宿主细胞是原核细胞。

10 56. 根据权利要求 49 或 51 所述的方法，其中第一和第二异源二聚化序列形成卷曲螺旋二聚体。

15 57. 根据权利要求 49 或 51 所述的方法，其中 L 链和 H 链多肽通过非共价配对亲和力二聚化。

58. 根据权利要求 56 所述的方法，其中 L 链多肽进一步包括其两侧翼为 L 链可变区和第一异源二聚化序列的 flexon。

20 59. 根据权利要求 56 所述的方法，其中 H 链多肽进一步包括其两侧翼为 H 链可变区和第二异源二聚化序列的 flexon。

60. 根据权利要求 56 所述的方法，其中第一和第二异源二聚化序列与至少一个半胱氨酸残基连接。

25 61. 根据权利要求 56 所述的方法，其中非单链抗原结合单位是多价的。

30 62. 根据权利要求 56 所述的方法，其中非单链抗原结合单位是多特异性的。

63. 根据权利要求 62 所述的方法，其中非单链抗原结合单位是双特异性的。

5 64. 根据权利要求 62 所述的方法，其中非单链抗原结合单位是三特异性的。

65. 根据权利要求 56 所述的方法，其中 L 链多肽包括衍生自人轻链的序列。

10

66. 根据权利要求 56 所述的方法，其中 H 链多肽包括衍生自人重链的序列。

67. 一种产生非单链抗原结合单位的方法，包括如下步骤：

15

(a)制备编码轻链(L)多肽的第一重组多核苷酸，所述轻链多肽包括框内融合到第一异源二聚化序列上的 L 链可变区，和编码重链(H)多肽的第二重组多核苷酸，所述重链多肽包括框内融合到第二异源二聚化序列上的 H 链可变区，其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，以及其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体；和

20

(b)允许第一和第二多肽经过第一和第二异源二聚化序列的配对亲和力进行二聚化。

68. 根据权利要求 67 所述的方法，其中步骤(b)包括在体外二聚化第一和第二多肽。

25

69. 一种产生单链抗原结合单位的方法，包括如下步骤：

(a)在宿主细胞中表达含有编码权利要求 20 或 22 所述的单链抗原结合单位的编码序列的多核苷酸；和选择性的

30

(b)分离在宿主细胞中表达的单链抗原结合单位。

70. 根据权利要求 69 所述的方法，其中多核苷酸包含在噬菌体展示载体中。

5 71. 一种在宿主细胞的表面上展示含有至少两种多肽的嵌合异源多聚体的方法，包括在宿主细胞中表达

(a)编码第一条多肽的第一重组多核苷酸，所述第一条多肽和第一异源二聚化序列和表面呈递序列框内融合在一起；

10 (b)编码第二条多肽的第二重组多核苷酸，所述第二条多肽和第二异源二聚化序列框内合在一起；

其中第一和第二条多肽通过第一和第二异源二聚化序列的配对亲和力和力进行二聚化；其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体。

15 72. 根据权利要求 71 所述的方法，其中第一和第二异源二聚化序列在生理缓冲条件和/或生理体温下基本上都不能形成同源二聚体。

73. 根据权利要求 71 所述的方法，其中第一和第二异源二聚化序列形成卷曲螺旋二聚体。

20

74. 根据权利要求 71 所述的方法，其中第一和第二条多核苷酸由单一噬菌体展示载体表达。

25 75. 根据权利要求 71 所述的方法，其中第一和第二条多核苷酸由不同噬菌体展示载体表达。

76. 根据权利要求 71 所述的方法，其中宿主细胞是原核细胞。

77. 根据权利要求 71 所述的方法，其中宿主细胞是真核细胞。

30

78. 根据权利要求 71 所述的方法，其中嵌合异源多聚体是非单链抗原结合单位。

5 79. 一种嵌合异源多聚体，根据权利要求 71 所述的方法展示在宿主细胞的表面上。

80. 一种能鉴定和所需抗原免疫反应的非单链抗原结合单位的方法，包括下列步骤：

10 (a)制备遗传多样性的抗原结合单位库，其中所述库包括多于一种的权利要求 1 或 3 所述的抗原结合单位；

(b)将抗原结合单位库和所需抗原接触；

(c)检测抗原结合单位和抗原之间的特异性结合，从而鉴定出能和所需抗原进行免疫反应的抗原结合单位。

15 81. 根据权利要求 80 所述的方法，其中抗原结合单位库的制备是通过表达编码多个抗原结合单位的载体文库来进行的。

82. 根据权利要求 80 所述的方法，其中载体文库包括多个噬菌体载体。

20

83. 一种能鉴定和所需抗原免疫反应的非单链抗原结合单位的方法，包括下列步骤：

(a)制备遗传多样性的抗原结合单位库，其中所述库包括至少一种的权利要求 20 或 22 所述的抗原结合单位；

25 (b)将抗原结合单位库和所需抗原接触；检测抗原结合单位和抗原之间的特异性结合，从而鉴定出能和所需抗原进行免疫反应的单链抗原结合单位。

30 84. 根据权利要求 83 所述的方法，其中抗原结合单位库的制备是通过表达编码多个抗原结合单位的载体文库来进行的。

85. 根据权利要求 83 所述的方法，其中载体文库包括多个噬菌体载体。

5 86. 一种试剂盒，以合适包装含有权利要求 39 所述的载体。

87. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中所述第一和第二异源二聚化序列包括生长因子受体的异源二聚受体序列。

10 88. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中所述第一和第二异源二聚化序列包括 G-蛋白偶联受体的异源二聚受体序列。

15 89. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中所述第一和第二异源二聚化序列包括神经递质的异源二聚受体序列。

90. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中所述第一和第二异源二聚化序列包括核激素受体的异源二聚受体序列。

20 91. 根据权利要求 1 或 3 所述的非单链抗原结合单位，其中抗原结合单位是 ccFv 片段。

92. 根据权利要求 1 所述的非单链抗原结合单位，其中生理体温大约是 37°C。

25 93. 根据权利要求 1 所述的非单链抗原结合单位，其中所述第一和第二异源二聚化序列等摩尔混合时，基本上不能形成同源二聚体。

产生嵌合异源多聚体的组合物和方法

5 参照相关申请

本申请是在审美国专利申请 09/921,144 的部分继续, 该专利申请提交于 2001 年 8 月 1 日, 其全文在此引作本申请的参考。

技术领域

10 本发明涉及的是免疫学领域。具体而言, 本发明涉及利用独特的异源二聚化序列产生嵌合异源多聚体如非单链抗原结合单位。本发明也涉及由受试异源二聚化序列稳定的单链抗原结合单位的产生。本发明包括的组合物和方法尤其用于鉴定具主要治疗和/或诊断前景的抗原结合单位。

15

发明背景

抗体或免疫球蛋白是一类能识别并结合特异关连抗原的分子。由于这些抗体具特异性, 因此它们, 尤其是单克隆抗体已被广泛地用于各种不同人类疾病的诊断和治疗中。

20

脊椎动物系统中基本的免疫球蛋白(Ig)由两条相同轻链(L)多肽(大约 23kDa)和两条相同重链(H)多肽(大约 53-70kDa)组成。这 4 条链由二硫键连接成“Y”构型。Y 构型的基部, 是由共价二硫键连接的两条 H 链。L 链和 H 链形成不同的结构域。L 链有两个结构域, 分别是 C 区(CL)和 V 区(VL)。H 链有四个结构域, 一个在 V 区(VH), 而其他三个结构域(CH1, CH2, CH3)则在 C 区。抗体包含两条臂(每条臂就是一个 Fab 片段), 每条臂有 1 个 VL 区和 1 个 VH 区, 彼此相连。就是这对 V 区(VL 和 VH), 它们在抗体中是互不相同的(是因为氨基酸序列发生了变化), 此外, 它们一起负责识别抗原, 并提供抗原结合位点。更具体地说, 每个 V 区由 3 个互补决定区(CDR)构成, 它们被 4 个构

25

30

架区(FR)分开。这些 CDR 是可变区中变异程度最高的部分，它们在抗原结合功能中起关键作用。这些 CDR 区由许多潜在的胚系序列经过复杂的过程，包括重组，突变和选择形成的。

5 近年的研究表明，完整抗体的一些片段也具有结合抗原的功能。示例性的抗原结合片段有(i)由 VL, VH, CL 和 CH1 区组成的 Fab 片段；(ii)由 VH 和 CH 区组成的 Fd 片段；(iii)由 VH 区组成的 dAb 片段(Ward, E.S. 等，自然，341，544-546(1989))；(iv)分离的 CDR 区；和(v)F(ab')₂ 片段，它是含有 2 个 Fab 片段的二价分子，通过铰链区的二硫桥连接在一起；和(vi)由抗体单一臂 VL 和 VH 区组成的 Fv 片段。Fv 片段是和抗原高亲和性结合的最小功能单位。

10

抗体领域的一个主要挑战一直是重构许多不同的免疫球蛋白库，用来模仿人免疫系统中的免疫球蛋白池。这种库的复杂度一般是 10⁸-10¹³ 种不同的免疫球蛋白。这种库的产生将非常利于能特异地和治疗靶点相互作用的免疫球蛋白的鉴定和产生。然而，由于一直以来，不能对最小的功能单位，即 Fv 片段稳定装配方法，因此阻碍了这种库的设计和制造。本领域一个众所周知的问题是，单独表达 VH 和 VL 区时，相互作用能非常低(Glockshuber 等，(1990)Biochemistry 29(6):1362-1367)。这两种成份在蛋白浓度很低时解离，而且，在生理体温下应用很不稳定。另一个长期认识的技术障碍是大蛋白，如完整的抗体(尽管十分稳定)在宿主细胞中的表达达不到可测量的水平，这样使得高度不同的抗体库的构建十分困难。

15

20

25 最近已开发了 3 种能获得稳定 VL 和 VH 复合物的方法。然而，这些方法各有许多内在局限，没有一种方法能完全克服前文提及的技术障碍。第一种方法用肽接头把 VL 和 VH 连接成单链("scFv")(Huston 等(1988), Proc. Natl. Acad. Sci. U.S.A. 85:5879-5883)。尽管获得的 scFv 基本上表示出了抗原结合活性，但不是所有抗体均能制成单链而仍保持高的亲和力(Huston 等(1988), Proc. Natl. Acad. Sci. U.S.A. 85:5879-

30

5883; Stemmer 等(1993)Biotechniques14(2):256-265)。这在一定程度上是由于连接肽序列对抗原结合位点干扰所致。第二种方法是在 VL 和 VH 区插入一对半胱氨酸残基, 获得一种二硫键稳定的 Fv(“dsFv”)(Brinkmann 等(1993), Proc. Natl. Acad. Sci. U.S.A. 90(16): 7538-7542)。然而, 在许多宿主细胞中, 插入的二硫键在还原条件下是不稳定的。例如, 在大肠杆菌的胞质溶胶中, 分子间的二硫键经常不足以稳定 VL 和 VH 复合物。此外, 这种方法通常要求 V 区的三维结构构象, 以确保一对半胱氨酸残基插入合适的位置, 而不破坏结合活性。由于大部分现有抗体的三维结构未知, 因此, 这种方法的实际用处不大, 它特别不适合抗体文库的构建, 尤其是不适合来自 B 细胞的抗体库的构建。第三种方法利用的是 CH1 和 CL 区中的天然二硫键来稳定 VL 和 VH 区。这种方法通过将连接 CH1 和 CL 区的二硫键嫁接到 VL 和 VH 区的 C 端, 以重构 Fab 片段。尽管获得的 Fab 片段一般比 scFv 更稳定, 结合亲和力更高, 但由于其尺寸太大, 因此 Fab 并不能最适于高水平表达和抗体库构建。

某些能形成卷曲螺旋结构的二聚化序列也能用来装配多价抗体。具体来说, 美国专利 NO5932448 描述了由 Fos 和 Jun 亮氨酸拉链连接的双特异的 $F(ab')_2$ 异源二聚体。Fos 和 Jun 亮氨酸拉链是充分表征的序列, 已知优选形成异源二聚体。然而, 它们在生理缓冲条件和/或生理体温下仍然有形成同源二聚体的倾向(O'Shea 等, (1992)细胞 68: 699-708; Vidal 等, (1996)Proc. Natl. Acad. Sci. U.S.A.)。事实上, Jun/Jun 同源二聚体是如此稳定以至在体外, 首先必须通过加热或用 2-巯基乙醇胺还原解离 Jun/Jun 同源二聚体, 才能形成 Fos/Jun 异源二聚体(参看美国专利 No5910573 第 7 栏第 35-37 行; 美国专利 5932448 第 16 栏第 15-30 行)。然而, 在体内实验时, Fos 和 Jun 都会产生可检测量的同源二聚体(例如参看美国专利 No5932448 第 15 栏第 41-43 行; 和 Vidal 等(1996)Proc. Natl. Acad. Sci. U.S.A.)。尽管单一种抗体的产生不必要考虑某些同源二聚化的倾向, 但这种倾向对抗体库的构建是个问题, 其中 VL 和 VH 之间的高效异源二聚体化是必需的。

除了 Fos 和 Jun 亮氨酸拉链, Housotn 等在美国专利 No5824483 中描述了卷曲螺旋二聚化肽组合文库的构建。Houston 等人提出这种文库对于多肽的鉴定是有用的, 而这种多肽能特异性地和选择的大分子配体如抗体相互作用 (参看跨越 8 和 9 页的最后一段)。很明显, Houston 等人考虑的是能和靶向抗体结合的“抗原肽”的选择, 而不是靶抗体的构建和选择。由于目的完全不同, Houston 等对卷曲螺旋序列用于产生稳定的抗原结合单位并没有描述, 甚至连暗示都没有。

因此仍很有必要对产生稳定抗原结合单位及其文库的组合物和方法进行改进, 从而影响治疗用抗原结合单位鉴定。一种理想的抗原结合单位应该比 Fv 片段更稳定, 但优选的是比 Fab 片段更小的, 这样就能大规模地制造及有效地展示。这种抗原结合单位也能用作构建多价和/或多特异性抗体的构件块。本发明满足了这种需要, 并提供了相关的优点。

发明概述

本发明的一个主要方面是设计一种技术, 它能把单体多肽特异装配成稳定的异源多聚体。这种产生异源多聚体的技术能便利地产生高通量的功能异源多聚体, 同时也能避免不需要的同源二聚体的装配。这种方法在产生具遗传多样性异源多聚体库如抗原结合单位文库时特别有用。这种技术能轻易地应用到不同的“遗传包装展示”技术中, 而这些技术使具所需结合特异性的抗原结合单位的选择更便利, 在美国专利 Nos 6248516, 5969108, 5885793, 5837500, 5571698, 5223409, 5514548, WO9005144, EP0368684, WO9201047, WO09311236 和 WO9708320 中详细介绍了这些遗传包装展示技术。

可以用一对完全不同的异源二聚化序列的配对亲和力来装配和固定受试的抗原结合单位。这对序列是完全不同的, 因为这对异源二聚体中的至少一个成员在生理缓冲条件和/或生理体温下基本上是不能形

成同源二聚体。在某些实施方案中，固定化的抗原结合单位不仅分子量比 Fab 片段小，而且也表示出必需的结合特异性和亲和力。此外，本发明的某些非单链抗原结合单位的结合亲和力比相应的常规单链抗体(scFv)要高。抗原结合单位特别适合抗体文库的构建及展示。受试的抗原结合单位的某些构型可以方便地用作多价和多特异性的免疫球蛋白的结构单位。

具体而言，本发明提供的非单链抗原结合单位包括：(a)轻链(L)多肽，它包含的轻链(L)可变区和第一异源二聚化序列融合在一起；(b)重链(H)多肽，它包含的重链(H)可变区和第二异源二聚化序列融合在一起；其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化；同时，其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体。优选的是第一和第二异源二聚化序列在生理缓冲条件和/或生理体温下基本上都不能形成同源二聚体。

另一方面，本发明提供的非单链抗原结合单位包括：(a)轻链(L)多肽，它包含的轻链(L)可变区和第一异源二聚化序列融合在一起；(b)重链(H)多肽，它包含的重链(H)可变区和第二异源二聚体序列融合在一起；其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，这两段异源二聚化序列衍生自异源二聚受体。一方面，包括异源二聚受体序列的第一和第二异源二聚化序列能介导受体的异源二聚化。另一方面，受试的异源二聚化序列形成卷曲螺旋二聚体。再一方面，L 链和 H 链经过这两段异源二聚化序列的非共价配对亲和力二聚化。优选的，L 链或 H 链进一步还包括其两侧翼为可变区和异源二聚化序列的 Flexon。L 链和 H 链多肽均来源于人的 L 链和 H 链。为了进一步固定异源二聚化抗原结合单位，可以引入半胱氨酸残基，从而可以在第一和第二异源二聚化序列之间产生二硫键。非单链抗原结合单位可以是单价或多价的。它们可以是单特异性的或多特异性的。优选的多特异性的 Abus 是二特异性的，三特异性的和四特异性

的分子。

在一个单独的实施方案中，本发明提供的单链抗原结合单位包括轻链(L)可变区和重链(H)可变区，它们由第一和第二异源二聚化序列连接，而这两段序列从一个区的C端跨越到另一个区的N端，其中两个可变区通过这两段异源二聚化序列的配对亲和力形成一个分子内二聚体；而且其中在生理缓冲条件和/或生理体温下，至少一种异源二聚化序列基本上不能形成同源二聚体。优选的是这两段异源二聚化序列在生理缓冲条件和/或生理体温下基本上都不能形成同源二聚体。

另一方面，本发明提供了一种单链抗原结合单位，其中VL和VH区通过来源于异源二聚受体的第一和第二异源二聚化序列的配对亲和力形成一个分子内二聚体。一方面，包括异源二聚受体序列的第一和第二异源二聚化序列能介导受体的异源二聚化。

再一方面，第一和第二异源二聚化序列形成卷曲螺旋二聚体。另一方面，第一和第二异源二聚化序列经过非共价配对亲和力二聚化。VL区和VH区均可能分别来源于人L区和H区的相应序列。

无论非单链还是单链抗原结合单位均可交联到一个具化学官能的基团。具这种官能的基团的例子包括，但不局限于信号肽，增强免疫反应活性的试剂，利于和固相支持物偶联的试剂，疫苗载体，生物反应调节物，毒素，可检测标记，顺磁标记和药物等。

受试的抗原结合单位中包含的优选异源二聚化序列分别来源于GABA_B受体1和GABA_B受体2的C端序列。更优选的是，第一异源二聚化序列连接到一个半胱氨酸残基上，该异源二聚化序列包括GABA_B受体1中至少30个氨基酸残基的多肽，这段多肽和SEQ ID NO.2中描述的可比长度的线性肽序列基本相同；而第二异源二聚化序列连接到一个半胱氨酸残基上，该异源二聚化序列包括GABA_B受体2

中至少 30 个氨基酸残基的多肽，这段多肽和 SEQ ID NO.4 中描述的可比长度的线性肽序列基本相同。或者，第一异源二聚化序列连接到一个半胱氨酸残基上，该异源二聚化序列包括 GABA_B 受体 2 中至少 30 个氨基酸残基的多肽，这段多肽和 SEQ ID NO.4 中描述的可比长度的线性肽序列基本相同；而第二异源二聚化序列连接到一个半胱氨酸残基上，该异源二聚化序列包括 GABA_B 受体 1 中至少 30 个氨基酸残基的多肽，这段多肽和 SEQ ID NO.2 中描述的可比长度的线性肽序列基本相同。

本发明提供的重组多核苷酸包括编码非单链抗原结合单位 L 和/或 H 多肽的编码序列。本发明也提供了包括编码单链抗原结合单位 VL 和/或 VH 区的编码序列的重组多核苷酸。此外，也提供了含有本发明提及的任一重组多核苷酸的载体。载体可以是表达载体，如噬菌体展示载体。本发明进一步提供了编码抗原结合单位库的表达载体可选择文库，这个选择文库包括多于 1 个的受试载体。优选的是这个可选择文库包括许多噬菌体展示载体。

本发明也提供了包括受试重组多核苷酸的宿主细胞。编码 L 链多肽的重组多核苷酸和编码 H 链多肽的多核苷酸可以出现在同一个载体或不同的载体中。宿主细胞可以是真核的或原核的。

在另一个实施方案中，本发明提供了生产非单链抗原结合单位的方法。这方法涉及如下步骤：(a)在宿主细胞中表达编码轻链(L)多肽的第一重组多核苷酸，所述轻链多肽包括融合在第一异源二聚化序列上的 L 链可变区，和表达编码重链(H)多肽的第二重组多核苷酸，所述重链多肽包括融合在第二异源二聚化序列上的 H 链可变区，其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，以及其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体；和可选择地(b)分离在宿主细胞中表达的抗原结合单位。

产生的抗原结合单位也可能包括来源于异源二聚受体的异源二聚化序列。此外，在步骤(a)中表达的非单链抗原结合单位可以在宿主细胞表面展示。在步骤(a)中表达的非单链抗原结合单位优选展示在噬菌体粒子表面。

在又一个实施方案中，本发明提供了生产非单链抗原结合单位的方法，所述方法包括如下步骤：(a)制备编码轻链(L)多肽的第一重组多核苷酸和编码重链(H)多肽的第二重组多核苷酸，所述轻链多肽包括融合在第一异源二聚化序列上的 L 链可变区，所述重链多肽包括融合在第二异源二聚化序列上的 H 链可变区，其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，以及其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体；和(b)允许第一和第二多肽经过第一和第二异源二聚化序列的配对亲和力进行二聚化。二聚化步骤可以离体或在体进行。

本发明也包括产生单链抗原结合单位的方法。所述方法涉及如下步骤：(a)在宿主细胞中表达含有编码受试单链抗原结合单位的编码序列的多核苷酸；(b)可选择地分离在宿主细胞中表达的单链抗原结合单位。

本发明进一步包括在宿主细胞表面展示嵌合异源多聚体的方法，其中嵌合异源多聚体包括至少两种多肽。这个方法包括在宿主细胞中表达(i)编码第一条多肽的第一重组多核苷酸；而第一条多肽与第一异源二聚化序列和表面呈递序列融合在一起；(ii)编码第二条多肽的第二重组多核苷酸，第二条多肽与第二异源二聚化序列融合在一起；其中第一和第二条多肽通过第一和第二异源二聚化序列的配对亲和力进行二聚化；其中至少一种异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体。一方面，第一和第二条多核苷酸可以通过单个噬菌体展示载体表达。另一方面，这两条多核苷酸也可

以分别通过噬菌体展示载体表达。嵌合的异源多聚体优选的是本发明的非单链抗原结合单位。

5 本发明也包括一种能鉴定和所需抗原发生免疫反应的非单链抗原结合单位的方法。该方法包括下列步骤：(a)制备遗传多样性的抗原结合单位库，其中这库包括多于一种的受试抗原结合单位；(b)把抗原结合单位库和所需抗原接触；和(C)检测抗原结合单位和抗原之间的特异结合，从而鉴定出能和所需抗原进行免疫反应的抗原结合单位。在该实施方案的一个方面，通过表达编码多个抗原结合单位的载体文库制
10 备抗原结合单位库。载体文库优选的是包括多个噬菌体载体。

最后，本发明提供了以合适包装包括本发明载体的试剂盒。

发明中使用的缩略词的解釋

- 15 1. Nsc: 非单链
2. Sc: 单链
3. Abu: 一个抗原结合单位
4. Abus: 多个抗原结合单位
4. L chain: 轻链
20 5. H chain: 重链
6. VL: 轻链可变区
7. VH: 重链可变区

附图简述

25 图 1 是一个示意图，它描述的是各种不同的抗原结合单位。

图 2 是用来构建受试 Abus 的 GABA_B 受体 1 和 2 的核苷酸和氨基酸序列。卷曲螺旋序列来源于人 GABA_B-R1 和 GABA_B-R2 受体。如图 2 上面一组序列所示，GABA_B 受体 1 的编码氨基酸序列是以 EEKS 开始，以 QLQS 结束。如图 2 下面一组序列所示，GABA_B 受体 2 的编码氨基酸序列是以 TSRL 开始，以 QLQD 结束。可弯曲的间隔序列
30

SerArgGlyGlyGlyGly 加到 R1 和 R2 异源二聚化序列的氨基端, 帮助形成功能性 Fv 异源二聚体。为了进一步稳定异源二聚体, 我们导入了 ValGlyGlyCys 间隔序列, 以利用半胱氨酸残基(SEQ ID NOS.2 和 4)间形成的二硫键把异源二聚体的卷曲螺旋对连锁起来。GGGG 间隔序列的 N 末端的 SerArg 编码序列提供了 Xba I 或 Xho I 位点, 用来将 GR1(来自 GABAB1 受体的异源二聚化序列)和 GR2(来自 GABAB2 受体的异源二聚化序列)分别融合到 VH 和 VL 片段的羧基端。

图 3A 是两个表达载体 pABMX1 和 pABMX2 的示意图。PABMX1 和 pABMX2 均是从 pbluescriptSK(+)衍生而来的, 包括抗生素筛选用的氨苄青霉素抗性基因(Amp), 一个质粒复制起点(colE1ori), 一个 f1 噬菌体复制起点(f1 ori), 和 lac 启动子/lacO1 驱动的表达盒(对 pABMX1 是 plac-RBS-p8 前导序列-DH 标志, 对 pABMX2 是 plac-RBS-pelB 前导序列-DH 标志)。异源序列被表达成 DH-标志的融合蛋白(HA 和 6×His 标志), 并被信号肽 p8 前导肽或 pelB 前导肽导入周质空间, 在此将前导序列剪切掉。

图 3B 是 pABMX1 和 pABMX2 中 lac 启动子后 Age I 和 Bgl II 位点之间的序列(SEQ ID NOS.5-8), 可以利用 HindIII/Xba I 或 HindIII/Not I 或 Xba I /Not I 位点来插入在 pABMX1 载体中表达的异源序列。在 pABMX2 载体中还包括克隆位点 Nco I , Pst I , Xba I 和 Not I 。

图 4A 是用来展示抗原结合单位的噬菌粒载体 pABMD1 和 pABMD2 的示意图。pABMD1 和 pABMD2 分别衍生自 pABMX1 和 pABMX2。它们包括载体 pABMX1 和 pABMX2 的所有功能元件和丝状噬菌体的 pIII基因。PIII基因紧邻 DH 标志的 3'端插入。Lac 启动子驱动异源序列表达成 pIII衣壳融合蛋白, 然后在辅助噬菌体如 KO7(Amersham pharmacia Biotech)或 R408(Stratagene)的超感染下展示在噬菌体粒子表面。这个载体也可在非抑制性菌株中用于表达可溶性蛋白。

图 4B 是 pABMD1 和 pABMD2 中 lac 启动子后的 Age I 和 Sal I 位点之间的序列(SEQ ID NOS.9-12)。

图 5A 是载体 pABMX5 和 pABMX6 的示意图。PABMX5 和

pABMX6 分别衍生自 pABMX1 和 pABMX2。在 pABMX5 和 pABMX6 中分别掺入了不同的前导序列。异源序列如 VH 基因插入的亚克隆位点在这两个载体中也有差异。pABMX5 包括 p8 前导序列，而 pABMX9 则包含 pelB 前导序列。Lac 启动子驱动下的两种蛋白表达盒通过基因工程手段导入了这两个载体。第一个表达盒用来表达 VH-GR1(GABA_B 受体 1 的 VH 异源二聚化序列)，而第二个表达盒用来表达 VL-GR2(GABA_B 受体 2 的 VL 异源二聚化序列)。DH 标志融合到 GR2 区，以便于纯化产生的异源二聚体。

图 5B 描述的是载体 pABMX5 和 pABMX6 中前导序列和 DH 标志之间的序列(SEQ ID NOS.13-16)。此外，核糖体结合位点，DH 标志，用于插入 VH、VL、GR1 和 GR2 的亚克隆位点也都标出来了。

图 6A 是噬菌粒 pABMD5 和 pABMD6 的示意图，它们用来在噬菌体粒子中表达和展示 ccFv。pABMD5 和 pABMD6 分别衍生自 pABMX5 和 pABMX6。来源于丝状噬菌体的 pIII 基因被紧接 DH 标志之后插入。VL-GR2 蛋白连接到 pIII 衣壳蛋白上，以便于 ccFv 异源二聚体的展示。

图 6B 描述的是载体 pABMD5 和 pABMD6 中前导序列和 pIII 标志之间的序列(SEQ ID NOS.13-16)。此外，核糖体结合位点，DH 标志，部分 pIII，用于插入 VH、VL、GR1 和 GR2 的亚克隆位点也都标出来了。

图 7 描述的是用来在酵母中表达 ccFv 片段的载体 pAMEX7。

图 8 描述的是利用 AM2-scFv 片段进行 ELISA 测定的结果，AM2-scFv 片段是用载体 pABMX1 表达的。结果表明结合到抗原 AM2 上的 AM2-scFv 是剂量依赖的。

图 9 描述的是利用展示在噬菌体粒子上的 AM2-scFv 片段进行 ELISA 测定的结果。结果证明用噬菌粒载体 pABMD1 在噬菌体粒子中能装配出功能性 scFv 片段。

图 10A 描述的是还原或非还原条件下在大肠杆菌中表达的 AM1-ccFv 进行 SDS-PAGE 分析结果。结果证明在大肠杆菌中成功地表达和装配了 ccFv 异源二聚体。

图 10B 描述的是利用在大肠杆菌中表达的可溶性 AM1-ccFv 进行 ELISA 进行分析的结果。结果表明，装配成功的功能性 ccFv 具有预期的相应抗原结合特异性。

5 图 11A 描述的是 AM1-ccFv 表达噬菌体和 AM1-scFv 表达噬菌体的抗原结合能力的比较。结果表明展示 AM1-ccFv 片段的噬菌体粒子的结合能力比展示常规 scFv 片段的噬菌体粒子的结合能力要稍高些。

图 11B 描述的是 AM2-ccFv 表达噬菌体和 AM2-scFv 表达噬菌体的抗原结合能力的比较。结果表明展示 AM2-ccFv 片段的噬菌体粒子的结合能力比 AM2-scFv 表达噬菌体的结合能力要高大约 1 个数量级。

10 图 12 描述的是三种多价的 Abu 构型，每种构型包括超过 1 个的 ccFv 基本单位。

图 13 描述的是四种二价的 Abu 构型，每种构型包括一个 ccFv 基本单位和一个 scFv 或 dsFv 片段。

15 图 14 描述的是三种三价的 Abu 构型，每种构型包括一个或多个 ccFv 基本单位，一个或多个 scFv 或 dsFv 片段。

图 15 描述的是四种双特异的 Abu 构型，每种构型包括一个或多个具不同结合特异性的 ccFv 基本单位和/或 scFv 或 dsFv 片段。

图 16 描述的是另外三种双特异性的 Abu 构型。

20 图 17 描述的是三种三特异性的 Abu 构型，每种构型包括至少一个 ccFv 基本单位以及至少一个 scFv 或 dsFv 片段。

图 18 描述的是两例单链 Abus，在这些 Abus 中，异源二聚化序列以平行或反式平行构型排列。

图 19 是展示在原核或真核细胞表面上的 ccFv 的示意图。上面一幅图描述的是由噬菌体粒子展示的 ccFv，其粘附在宿主细胞表面。

25 图 20 描述的是利用 ELISA 对模式噬菌体文库进行淘选的结果，从淘选“模式文库”中涉及 AM2 的固定化蛋白抗原。“模式文库”包含 AM2-ccFv 噬菌体和不相关噬菌体、AM1-ccFv 的混合物，两者的比例分别是 $1:10^6$ 和 $1:10^7$ 。因为 ELISA 只能检测到 AM2-ccFv 的反应活性，因此 ELISA 的读数只代表混合物中 AM2-ccFv 群体。第一轮淘选后，没有检测到 ELISA 信号，表明出现在混合物中的 AM2-ccFv

30

部分仍然很低。然而，两轮淘选后，检测到的 AM2-ccFv 部分开始占优，ELISA 信号的显著增长便能证明。

图 21 描述的是用 PCR 分析来证实图 12 所示的模式文库淘选后 AM2-ccFv 的富集。在淘选之前，可以在每 10^7 不相关噬菌体中检测到 1 个 AM2-ccFv 噬菌体(0.00001%)。第一轮淘选后，AM2-ccFv 出现率达到 4.4%，而第二轮淘选后则达到 100%。

图 22 描述的是 AM2-ccFv 的 VH(重链)CDR3 的设计。用位于两个亚克隆限制性位点侧翼的两条 PCR 引物扩增变性的寡聚 DNA。首先用 PCR 扩增寡 DNA，然后用限制性酶进行消化，接着亚克隆取代 AM2-ccFv VH 中的野生型 CDR3。

图 23 描述的是从第 5 轮(上图)和第 7 轮(下图)淘选中随机挑选的单个克隆的 AM2-ccFv 噬菌体 ELISA 结果。每个柱子代表的是每个特定 AM2-ccFv 变体在 OD405 的读数。

图 24 描述的是从 AM2-ccFv CDR3 VH 文库淘选中筛选的变体的序列。最上面所示为文库的氨基酸序列(X 表示从图 22 所述的变性寡聚 DNA 中衍生出来的残基的多种选择)；下面所示的是在每个“X”位置从淘选中选择的残基及它们的撞击速率(特定残基的发生速率)。通过文库设计没有改变的位置用空白表示。衍生自淘选的共有序列如最下面所示。

图 25 描述的是 AM2-ccFv 变体和野生型 AM2-ccFv 的 K_{off} 速率的比较。曲线的斜率表示的是检测的抗体从抗原上解离有多快。

实施本发明的方式

在整个公开内容中，不同的出版物，专利以及出版的专利说明书被识别引用引作参考。这些出版物，专利和出版的专利说明书的公开内容在这里一并引作本公开内容的参考。

总的技术

除非另有说明，本发明的实施将采用来自免疫学，生物化学，化学，分子生物学，微生物学，细胞生物学，遗传学和重组 DNA 的常

规技术，它们都是本领域的技术。参考，如 Matthews, PLANT VIROLOGY，第三版(1991)；Sambrook, Fritsch 和 Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 第二版(1989)；CURRENT PROTOCOLS IN MOLECULAR BIOLOGY(F.M. Ausubel 等编辑(1987)；METHODS IN ENZYMOLOGY 系列(Academic 出版社公司)；PCR 2: A PRACTICAL APPROACH(M.J. MacPherson, B.D. Hames 和 G.R. Taylor 编辑(1995)，Harlow 和 Lane 编辑(1988)ANTIBODIES, A LABORATORY MANUAL, 和 ANIMAL CELL CULTURE(R.I. Freshney 编辑(1987))。

10

说明书和权力要求书中使用的单数形式“一个”，“一种”和“所述”，如果上下文没有清楚地表示，包括复数关系。例如，术语“一个细胞”指的是多个细胞，它包括细胞混合物。

15

定义

在这里交替使用的术语“多肽”，“肽”和“蛋白”指的是任意长度的氨基酸多聚体。这个多聚体可能是线性，环形或分支的，它可能包含修饰的氨基酸，也可能被非氨基酸所打断。这些术语也包括那些经过修饰的氨基酸多聚体，如经过硫化，糖基化，脂化，乙酰化，磷酸化，碘化，甲基化，氧化，蛋白水解加工，磷酸化，异戊二烯化，外消旋化，硒化，转移 RNA 介导的向蛋白添加氨基酸，如精氨酰化，遍在蛋白化或任何别的操作，如和标记组分交联。这里使用的术语“氨基酸”指的是天然和/或非天然的或合成的氨基酸，包括甘氨酸和 D 或 L 光学异构体和氨基酸类似物和拟肽(peptidomimetics)。

25

多肽或氨基酸序列“衍生”自一种已经命名的蛋白质指的是多肽的起源。优选的是，多肽具有和序列编码的多肽或其部分基本一样的氨基酸序列，其中所述部分的组成为至少 10-20 个氨基酸，优选的是至少 20-30 个氨基酸，更优选的是至少 30-50 个氨基酸，或者它和这个序列编码的多肽通过免疫学方法可加以识别。这个术语也包括从已

30

经命名的核酸序列表达的多肽。

“嵌合”蛋白包括至少一种融合多肽，该融合多肽所含有的区在序列中的位置和天然多肽不同。这些区域正常情况下存在于不同的蛋白质中，并在融合多肽中结合在一起了；或者在正常情况下它们位于同一蛋白质中，但在融合多肽中有了新的排列。嵌合蛋白的产生方式有，如化学合成，或制造和翻译其中肽区按所需关系加以编码的多核苷酸。

这里使用的“多聚体蛋白”指的是一种球蛋白，它包含多于一种单独的多肽或蛋白链，它们在体外或体内相互结合形成单一球蛋白。由 1 个以上同种多肽组成的多聚体蛋白形成“同源多聚体”。另外，由 1 个以上不同序列的多肽组成的多聚体蛋白也可形成“异源多聚体”。因此，“异源多聚体”是一种包括至少第一多肽和第二多肽的分子，其中第二多肽和第一多肽在氨基酸序列上至少有一个氨基酸残基的不同。异源多聚体可包括由第一和第二多肽组成的“异源二聚体”，或者可形成更高级的三级结构，其中出现了多于两个的多肽。异源多聚体的示例性结构包括异源二聚体(如 Fv 和 Fab 片段; diabodies; GABA_B 受体 1 和 2 复合物)，三聚体 G 蛋白，异源四聚体(如 F(ab')₂ 片段)以及进一步包括寡聚结构。

嵌合异源多聚体的“第一条重组多肽”指的是任何可以通过两条二聚化序列的配对亲和力与第二条重组多肽结合或曾结合在一起的多肽，其中二聚化序列分别和第一和第二多肽相连接。第一和第二多肽包含的序列优选来自免疫球蛋白的轻链或重链。更优选的是，第一和第二多肽形成 Nsc Abu，其给所需抗原赋予结合特异性。

“第一异源二聚化序列”指的是任何二聚化序列，它和“第二异源二聚化序列”结合或曾结合在一起，其中第二异源二聚化序列在氨基酸序列上至少有一个氨基酸残基不同。“一对异源二聚化序列”指

的是能形成异源二聚体的两段异源二聚化序列。

这里使用的“抗体”指的是免疫球蛋白分子和免疫球蛋白分子中具免疫活性的部分，即含能特异地结合抗原(和抗原免疫反应)的抗原结合位点的分子。从结构上说，天然存在的最简单的抗体(如 IgG)包括 4 条多肽链：2 条重链(H)和 2 条轻链(L)，它们通过二硫键交连在一起。免疫球蛋白代表一个大的分子家族，包括若干类分子，诸如 IgD, IgG, IgA, IgM 和 IgE。例如，术语“免疫球蛋白分子”包括例如杂合抗体或改型抗体及其片段。抗体的抗原结合功能可以由天然存在抗体的片段来执行。这些片段统称为“抗原结合单位”(“Abus”)。从更广泛的范围来说，Abus 基于其分子结构可分为“单链”(“Sc”)和“非单链”(“Nsc”)。

术语“抗体”和“Abus”也包括不同种属来源的免疫球蛋白分子，所述种属包括无脊椎动物和脊椎动物。当“人”这个术语用在抗体或 Abu 上时，指的是由人基因或基因片段表达的免疫球蛋白分子。术语“人源化”应用到非人抗体(如啮齿类或灵长类)时，指的是杂合的免疫球蛋白，免疫球蛋白链或其片段，它们含有极少的来源于非人免疫球蛋白的序列。人源化抗体绝大部分是人免疫球蛋白(受体抗体)，在这种人源化抗体中，受体的互补决定簇(CDR)中的残基被非人物种(供体抗体)，如小鼠，大鼠，兔子或灵长类动物中具所需特异性，亲和力和容量的 CDR 中的残基取代。在一些情况下，人免疫球蛋白的 Fv 框架区(FR)残基被相应的非人残基取代。此外，人源化抗体也包括一些既没在受体抗体中发现，也没在输入的 CDR 或框架序列中发现的残基。当抗体进入人体时，应对抗体进行修饰，以进一步精制和优化抗体效用，减少免疫原性。总之，人源化抗体基本上包括所有那些至少含一个，典型是 2 个可变区的抗体，在这些抗体中，所有或基本上所有的 CDR 区对应于非人免疫球蛋白的那些 CDR，而所有或基本上所有的 FR 区是人免疫球蛋白序列的 FR。人源化抗体也包括至少一部分免疫球蛋白恒定区(Fc)，典型的是人免疫球蛋白恒定区。

“非单链抗原结合单位”(Nsc Abus)是含有轻链多肽和重链多肽的异源多聚体。Nsc Abus 的例子包括，但不局限于(i)由本发明公开的异源二聚化序列稳定的 ccFv 片段(图 1); (ii)任何含有本发明描述的至少一个 ccFv 片段的其他单价和多价分子; (iii)由 VL, VH, CL 和 CH 区组成的 Fab 片段; (iv)由 VH 和 CH1 区组成的 Fd 片段; (v)由任何抗体单臂的 VL 和 VH 区组成的 Fv 片段; (vi)F(ab')₂ 片段，它是一个二价片段，在铰链区含有两个由二硫键连接的 Fab 片段; (vii)diabody 和(viii)任何别的在 Little 等(2000)Immunology Today 中介绍的 Nsc Abus。

正如上面提及的那样，Nsc Abus 可以是单价的，也可以是多价的。然而，前者每个抗原结合单位只有一个结合位点，而后者包含多个能结合相同或不同抗原的结合位点。根据结合位点数目，Nsc Abus 可以是二价的(有两个抗原结合位点)，三价的(有三个抗原结合位点)，四价的(有四个抗原结合位点)等等。

多价的 Nsc Abus 可以根据它们的结合特异性进一步分类。“单特异性”的 Nsc Abu 是一种能结合一或多个相同抗原的分子。“多特异的”Nsc Abu 是一种对至少两种不同抗原具结合特异性的分子。尽管正常情况下，这种分子只能结合两种不同的抗原(即双特异的 Abus)，带有附加特异性的抗体如三特异性抗体在本发明中使用，也包括在这种表达中(如参看图 15-17)。双特异性抗原结合单位的例子包括那些一条臂抗肿瘤细胞抗原，而另一条臂抗细胞毒性触发分子的抗原结合单位，如抗 FcγRI/抗 CD15，抗 p185^{HER2}/FcγRIII(CD16)，抗-CD3/抗恶性 B 细胞(1D10)，抗 CD3/抗-p185^{HER2}，抗 CD3/抗-p97，抗 CD3/抗肾细胞癌，抗 CD3/抗 OVCAR-3，抗 CD3/L-D1(抗结肠癌)，抗-CD3/抗促黑激素类似物，抗 EGF 受体/抗 CD3，抗 CD3/抗 CAMA1，抗 CD3/抗 CD19，抗 CD3/MoV18，抗神经细胞粘着分子(NCAM)/抗 CD3，抗叶酸结合蛋白(FBP)/抗 CD3，抗泛肿瘤相关抗原(pan carcinoma

associated antigen)(AMOC-31)/抗 CD3; 一条臂特异地结合肿瘤抗原, 而另一条臂结合毒素的双特异 Abus, 如抗皂草素/抗 Id-1, 抗 CD22/抗皂草素, 抗 CD7/抗皂草素, 抗 CD38/抗皂草素, 抗 CEA/抗蓖麻毒蛋白 A 链, 抗 α 干扰素(IFN- α)/抗杂交瘤独特型, 抗 CEA/抗长春花生物碱; 转变酶激活前药的 BsAbs 如抗 CD30/抗碱性磷酸酶(它能催化磷酸丝裂霉素前药向乙醇丝裂霉素的转变); 能用作纤溶蛋白抗原的双特异 Abus 如抗纤维蛋白/抗组织纤溶酶原激活剂(tPA), 抗纤维蛋白/抗尿激酶型纤溶酶激活剂(uPA); 将免疫复合物导向细胞表面受体的双特异抗原结合单位如抗低密度脂蛋白(LDL)/抗 Fc 受体(如 Fc γ R I, Fc γ R II, Fc γ R III); 用于治疗感染病的双特异 Abus 如抗 CD3/抗单纯疱疹病毒(HSV), 抗 T 细胞受体:CD3 复合物/抗流感, 抗 Fc γ R/抗 HIV; 离体和在体检测肿瘤的双特异 Abus 如抗 CEA/抗 EOTUBE, 抗 CEA/抗 DPTA, 抗 P185^{HER2}/抗半抗原; 作为疫苗佐剂的 BsAbs(参看 Fanger 等, 同上); 作为诊断工具的双特异 Abus 如抗兔 IgG/抗铁蛋白, 抗辣根过氧化物酶(HRP)/抗激素, 抗促生长素抑制素/抗 P 物质, 抗 HRP/抗 FITC, 抗 CEA/抗 β -半乳糖苷酶(参看 Nolan 等, 同上)。三特异性抗体的例子包括抗 CD3/抗 CD4/抗 CD37, 抗 CD3/抗 CD5/抗 CD37 和抗 CD3/抗 CD8/抗 CD37。

“单链抗原结合单位”(“Sc Abu”)指的是单体 Abu。尽管 Fv 片段两个结构域是由不同基因编码, 但可以通过重组技术, 合成一个连接分子, 通过连接分子使这两个结构域形成一条单一的蛋白链(即在 Bird 等人(1988)科学 242: 423-426 和 Huston 等(1988)PNAS 85: 5879-5883 中介绍的单链 Fv(“svFv”)。别的 Sc Abus 还包括由受试的异源二聚化序列稳定的抗原结合分子(例如, 参看图 18), 和由 VH 结构域和分离的互补决定区(CDR)组成的 dAb 片段(Ward 等(1989)自然, 341: 544-546)。连接肽如(GGGGS)₃, 它连接一个 V 区 C 末端和另一个 V 区氨基末端之间的距离大约是 3.5nm。别的接头也可应用, 而且它们能提供附加功能, 如可以作为粘附药物或固相支持物的工具。一种优选的单链抗原结合单位包括 VL 区和 VH 区, 它们被一对受试的异源

二聚化序列连接和固定在一起。ScFv 可以任何顺序装配，例如 VH-(第一异源二聚化序列)-(第二异源二聚化序列)-VL，或 VL-(第一异源二聚化序列)-(第二异源二聚化序列)-VH。

5 “抗原结合单位库”指的是多个抗原结合单位，其中至少有两个表示出完全不同的结合特异性。具遗传多样性的抗原结合单位库指的是多个抗原结合单位，就彼此而言，如果不是全部，则大部分抗原结合单位表示出唯一的结合特异性。典型的具遗传多样性的库复杂度至少是 10^6 - 10^{13} ，优选的是 10^7 - 10^9 ，更优选的是 10^8 - 10^{10} ，甚至更优选
10 10^8 - 10^{11} 种完全不同的抗原结合单位。

如果抗体或 Abu 结合抗原的亲和性或亲和力比它结合别的对照抗原如多肽或其他物质大时，它就和抗原“特异性结合”或“免疫反应”。

15 当 Abu 被呈递到宿主细胞外表面时，Abu 就是被展示“在宿主细胞表面”。展示的 Abu 可以直接粘附到宿主细胞外表面，或通过宿主细胞结合的遗传包装单位如噬菌体粒子间接粘附到宿主细胞。

“表面呈递序列”指的是能便于异源序列展示的序列。典型的情况下是表面呈递序列呈递在遗传包装单位如噬菌体或细菌的外表面。
20 优选的噬菌体表面呈递序列是 M13 丝状噬菌体的 PIII。

本发明中使用的“抗原”是一种能被抗体识别和特异性结合的物质。抗原可以包括肽，蛋白，糖蛋白，多糖和脂；它们的一部分或它们的组合。
25

正如本发明使用的那样，术语“表面抗原”指的是细胞的质膜成份。它包括膜内在蛋白和外周膜蛋白，构成质膜的糖蛋白，多糖和脂质。“膜内在蛋白”是一类跨膜蛋白，它们跨过细胞质膜的脂双分子层。典型的膜内在蛋白由至少一种“跨膜区段”组成，跨膜区段一般
30

包含的是疏水性氨基酸残基。外周膜蛋白不会进入疏水性脂双分子层内部，它们通过和别的膜蛋白的非共价相互作用结合到膜表面。

5 当术语“膜”，“胞质的”，“核”和“分泌的”应用到细胞蛋白时，特指细胞内蛋白主要的，占优的，或优选的位于其中的细胞外和/或亚细胞位置。

“细胞表面受体”表示的是一套膜蛋白，它们能结合各自的配体。细胞表面受体是一些锚定或插入细胞质膜中的分子。它们构成了一大
10 类蛋白，糖蛋白，多糖和脂质家族。它们不仅可作为质膜的结构组成，而且也能作为控制不同生物功能的调节元件。

“异源二聚受体”包括含有两个蛋白样(Proteinaceous)亚单位的细胞蛋白，它对配体表示出结合亲和力。这两个蛋白样亚单位是两个
15 不一样的分子，在氨基酸序列上至少有一个氨基酸残基的差异。非限制的异源二聚受体如那些能和生长因子(如 heregulin)，神经递质(如 γ -氨基丁酸)和别的有机或无机小分子(如盐皮质激素，糖皮质激素)结合的受体。优选的异源二聚受体是核激素受体(Belshaw 等(1996), Proc. Natl. Acad. Sci. U.S.A. 93(10):4604-4607)，erbB3 和 erbB2 受体复合物，
20 G 蛋白偶联受体和 GABA_B 受体家族，而 G 蛋白偶联受体包括，但不限于阿片样物质(Gomes 等(2000)J. Neuroscience 20(22). RC110); Jordan 等(1999)Nature399:697-700)，毒绳碱，多巴胺，血清素，腺苷/多巴胺。

“结构域”指的是蛋白质的一部分。它在物理结构和功能上与蛋白质或肽的其它部分不一样。从物理结构上定义的结构域包括那些具
25 异常疏水性或亲水性的氨基酸序列，如那些膜结合和胞质结合的序列。例如，结构域也可从基因复制出现的内部同源性来定义。从功能上定义的结构域具不同的生物学功能。例如，受体的配体结合结构域便是结合配体的结构域。抗原结合结构域指的是能结合抗原的抗原结合单位或抗体的一部分。功能上定义的结构域不必由邻近的氨基酸序
30

列编码。功能意义上的结构域可能包含 1 或多个物理结构意义的结构域。例如受体一般分成胞外配体结合结构域，跨膜结构域和胞内效应结构域。“膜锚定结构域”指的是能介导膜结合的蛋白部分。一般而言，膜锚定结构域由疏水性氨基酸残基组成。此外，膜锚定结构域也包含修饰的氨基酸，如粘附到脂肪酸链的氨基酸，它又能将蛋白锚定到膜上。

“宿主细胞”包括单个细胞或细胞培养物，它们可以是或已是受试的载体的受体。宿主细胞包括单个宿主细胞的后代。由于存在自然突变，偶然突变或故意突变，宿主细胞后代没必要和最初的亲本完全一样(在形态学或总 DNA 互补的基因组中)。宿主细胞还包括用本发明载体在体内转染的细胞。

“细胞系”或“细胞培养物”意思是指在体外生长或保存的细菌，植物，昆虫或高等真核生物细胞。细胞后代和亲本可以不完全一样(形态学的，基因型的，或表型的)。

“合成培养基”指的是含有细胞生存和/或生长所必需的营养和激素的培养基，这种培养基的成分是确定的。按惯例，通过添加生长和/或生存所必需的营养和生长因子，按配方配制合成培养基。典型的合成培养基提供了至少一种来自下列一或多个类别的成份：a)所有必需氨基酸，和通常包括 20 个氨基酸加上半胱氨酸的一套基本氨基酸；b)能源，通常是碳水化合物形式，如葡萄糖；c)维生素和/或别的要求低浓度的有机化合物；d)游离脂肪酸；和 e)微量元素，其中，微量元素定义为无机化合物或通常浓度要求比较低的天然存在的元素，一般是在微摩范围。合成培养基也可选择性的补充来自下列任何一个类别的一或多种成份：a)一或多种促有丝分裂剂；b)盐和缓冲液，如钙，镁和磷酸盐；c)核苷和碱基，如腺苷和胸苷，次黄嘌呤，和 d)蛋白和组织水解产物。

本发明使用的术语“分离的”意思是指从成份、细胞等分离出来，在这些成份、细胞中，多核苷酸，肽，多肽，蛋白，抗体或其片段正常状态下是自然结合在一起的。不是天然存在的多核苷酸，肽，多肽，蛋白，抗体或其片段并没必要通过分离来和它们自然存在的对应物相区分，这一点本领域的技术人员都很明白。此外，“浓缩的”，“分离的”或“稀释的”多核苷酸，肽，多肽，蛋白，抗体或其片段和及自然存在的对应物是有区别的，这是因为和天然出现的对应物相比较，每体积分子浓度或数目比“浓缩的”要大，或比“分离的”要小。

可以用绝对值检测富集，如每体积溶液的重量，或可以以相对于混合物中存在的第二个，潜在的干扰物质来检测。在本发明中，富集程度增加越多的实施方案越是优选的。因此，例如，2 倍富集是优选的，那富集 10 倍，100 倍，甚至于 1000 倍都是更优选的。通过一种人工装配方法，如化学合成或重组表达可以提供一种分离的物质。

“连接的”和“融合的”或“融合”在本发明中交替使用。这些术语指的是多于两个的化学元素或成份，通过包括化学交联或重组技术在内的任何方法被连接在一起。“框架内融合”指的是以一种保持原始 OFR 读框正确的方式，将两个或多个开放阅读框(OFRs)连接在一起，形成一个连续更长的开放阅读框。因此，产生的重组融合蛋白是一种单链蛋白，它包含两个或多个区段，它们对应于原始 OFR 编码的多肽(这些 OFR 的区段在正常情况下不是这样连接的)。尽管读框在整个融合区段中会这样被制成连续的，但这些区段在物理结构或空间上还是会被隔开的，如被下文所述的框内接头(如“flexon”)隔开。

从多肽的前后意思来看，“线性序列”或“序列”指的是多肽中的氨基酸顺序，方向是从氨基端到羧基端，根据这个方向，序列中互相邻接的残基在多肽的一级结构中是邻近的。“部分序列”指的是多肽的部分线性序列，已知多肽在一个或两个方向上均包含其他残基。

“异源的”意思是指从一种实体衍生出来的，该实体在基因型上截然不同于正被比较的实体的其余部分。例如，从启动子的天然编码序列中去掉启动子，且可操纵地连接到不是天然序列的编码序列上，这种启动子叫异源启动子。术语“异源的”可以应用到多核苷酸、多肽，意思是指多核苷酸或多肽衍生自一种实体，该实体在基因型上截然不同于正被比较实体的其余部分。例如，异源的多核苷酸或抗原可以衍生自不同的物种来源，不同的细胞类型以及不同个体的同类细胞。

术语“多核苷酸”“核酸”“核苷酸”和“寡核苷酸”在本发明中交替使用。它们指的是任何长度的核苷酸的多聚体形式，是去氧核糖核苷酸或核糖核苷酸，或它们的类似物。多核苷酸可具有任何三维结构，可执行任何已知或未知的功能。下面是非限制的多核苷酸例子：基因或基因片段的编码或非编码区，从连锁分析中定义的基因座，外显子，内含子，信使 RNA(mRNA)，转移 RNA(tRNA)，核糖体 RNA，核酶，cDNA，重组多核苷酸，分枝多核苷酸，质粒，载体，分离的任何 DNA 序列，分离的任何 RNA 序列，核酸探针和引物。多核苷酸也可能包括修饰的核苷酸，如甲基化的核苷酸和核苷酸类似物。如果核苷酸结构被修饰的话，会出现在多聚体装配之前或之后。核苷酸序列可以被非核苷酸成份打断。多核苷酸在多聚化后可以被进一步修饰，如和一种标记组分交联。

当“重组的”应用到多核苷酸时，意思指多核苷酸是克隆，限制性酶切和连接步骤以及别的程序的不同组合产物，从而产生一种和天然发现的多核苷酸截然不同的构建体。

术语“基因”或“基因片段”在本发明中交替使用。它们指的是一种多核苷酸，包括至少一个开放阅读框，在转录和翻译后，能够编码一种特定蛋白。基因或基因片段可以是基因组或 cDNA，只要多核苷酸包括至少一种开放阅读框，而开放阅读框覆盖了整个编码区或其

区段。

5 “可操作连接”或“操作性连接”指的是一种并列关系，其中这些描述的成分是一种允许它们按目的方式行使功能的关系。例如如果一个启动子序列能启动编码序列的转录，则这个启动子序列就可操作地连接到编码序列上。

10 “融合基因”是一种基因，它由至少两个连接在一起的异源多核苷酸组成。

基因“数据库”指的是一套贮存的数据，它代表的是包括核苷酸和肽序列的序列集合，而这些序列又代表的是生物参照材料的集合。

15 如本发明所使用的那样，“表达”指的是多核苷酸被转录成 mRNA 的过程和/或转录的 mRNA(也叫转录产物)随即被翻译成肽，多肽或蛋白质的过程。转录产物和编码的多肽被统一叫基因产物。假如多核苷酸衍生自基因组 DNA，真核细胞中的表达还包括 mRNA 的剪切。

20 本发明中使用的“受试者”指的是含有表达的遗传材料的生物实体。生物实体优选的是植物，动物或微生物，而微生物包括细菌，病毒，真菌和原生动物。在体获得的或体外培养的生物实体的组织，细胞和它们的后代也包括在内。

25 “载体”是核酸分子，优选的是能自主复制的，载体能将插入的核酸分子转移进宿主细胞，或者在宿主细胞之间转移。这个术语包括功能主要是将 DNA 或 RNA 插入到细胞中的载体，功能主要是复制 DNA 或 RNA 的载体的复制，和功能是转录和/或翻译 DNA 或 RNA 的表达载体。提供一种以上的上述功能的载体也包括在内。

30 “表达载体”是一种多核苷酸，当导入合适的宿主细胞时，能被

转录和翻译成多肽。一种“表达系统”通常暗含一种合适的宿主细胞，该宿主包括一种能产生所需表达产物的表达载体。

5 “复制子”指的是包括一个复制起点(一般叫 ori 序列)的多核苷酸，它允许多核苷酸在合适的宿主细胞中复制。复制子的例子包括附加体(如质粒)，以及染色体(如核染色体或线粒体染色体)。

本发明的嵌合异源多聚体

10 如上所述，多肽亚基正确装配成稳定复合物是保证多聚体蛋白的生物功能所必需的。因此，本发明一个主要方面是设计一种技术，使得选择的单体多肽能特异性地装配，从而影响异源二聚体的有效产量。实验设计对于异源二聚体如 Abus 的产生和筛选特别有用，异源二聚体的结合特异性取决于特定亚基以特异方式进行的装配。和以前报道的嵌合 Abus 有区别的是，受试的 Abus 有一个或多个下列唯一的特征。首先，Abus 可以通过两条异源二聚化序列的配对亲和力进行重
15 构，这两段序列中至少有一条，优选的是两条都缺乏形成同源二聚体的可检测倾向。不象以前报道的二聚化序列，如 Fos 和 Jun 亮氨酸拉链，已知它们在生理缓冲液条件和/或特定的体温下均能形成同源二聚体(O'shea 等(1992)细胞 68:699-708, Vidal 等(1996)Proc. Natl. Acad. Sci.
20 U.S.A)，受试的异源二聚化序列在特定缓冲液条件和/或特定的体温下基本上不能形成同源二聚体。如下详细介绍的那样，受试的异源二聚化序列和以前采用的序列在结构水平上也有所不同。

25 在一个实施方案中，本发明提供了一种展示在宿主细胞表面上的杂合异源多聚体，其中异源多聚体包括(i)与第一异源二聚化序列和表面呈递序列相融合的第一条多肽；(ii)与第二异源二聚化序列融合的第二条多肽；其中第一和第二条多肽通过两段异源二聚化序列的配对亲和力进行二聚化；其中至少一种异源聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体。

30

在另一个实施方案中，本发明提供的 Nsc Abu 包括：(a)轻链多肽，它包含的轻链可变区和第一异源二聚化序列相连；(b)重链多肽，它包含的重链可变区和第二异源二聚化序列相连；其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，这两段异源二聚化序列中的至少一种在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体。在另一方面，本发明提供了一种 Nsc Abu，其中 L 链和 H 链多肽通过第一和第二异源二聚化序列的配对亲和力二聚化，这两段异源二聚化序列衍生自异源二聚受体。一方面，包括异源二聚受体序列的第一和第二异源二聚化序列能介导受体的异源二聚化。

在一个单独的实施方案中，本发明提供的单链抗原结合单位(Sc Abu)包括轻链(L)可变区和重链(H)可变区，它们由第一和第二异源二聚化序列连接，而这两段序列从一个区的 C 端跨越到另一个区的 N 端，其中这两个可变区通过第一和第二异源二聚化序列的配对亲和力形成一个分子内二聚体，而且在生理缓冲条件和/或生理体温下，这两段异源二聚化序列中的至少一种基本上不能形成同源二聚体。该实施方案的另一方面，本发明提供了一种单链抗原结合单位，其中 VL 和 VH 区通过来源于异源二聚受体的第一和第二异源二聚化序列的配对亲和力形成一个分子内二聚体。一方面，包括异源二聚受体序列的第一和第二异源二聚化序列能介导受体的异源二聚化。

异源二聚化序列的选择

若干因素应用到了具一个或多个上述特征的 Abus 的设计中。首先，异源二聚化序列必须表示出能影响稳定复合物形成的配对亲和力。“稳定”意味着复合物或二聚化足以在复合物或二聚体的形成和随后它的检测和/或纯化间长期保持。这些复合物或二聚体在形成到检测这段时间内，无论存在或引入什么条件，均经受得住，而这些条件能进行试验或反应。优选在生理缓冲液条件和室温到约 37℃范围内的生理体温条件下，形成复合物或二聚体。干扰条件包括冲洗，加热，往反应混合物中添加另外的溶质或溶剂(如变性剂)，以及和另外的反

应物种竞争，它们可以选择性地存在，并移出复合物或二聚体。稳定的复合物或二聚体是不可逆或可逆的，但必须满足本定义的其他要求。因此，在反应混合物中可以形成一种瞬间复合物或二聚体，但是如果这个复合物或二聚体在生理缓冲液条件下自发解离或在检测之前新加入别的条件或引入操作，它就不可能构成稳定复合物。

第二，选择的异源二聚体序列必须表示出配对亲和力，这样就会主要形成异源二聚体，而基本上排除了同源二聚体。在生理缓冲条件和/或生理体温允许下，主要形成的异源多聚体池优选含有至少 60%的异源二聚体，而至少 80%，85-90%，90-95%，甚至于 96-99%的异源二聚体则是更优选的。在本发明的某些实施方案中，至少一种用来重构 Abu 的异源二聚化序列在生理缓冲条件和/或生理体温下基本上不能形成同源二聚体。“基本上不能”意思是指选择的异源二聚化序列在 Kammerer 等人详细介绍的体外沉降实验(1999, Biochemistry, 38:13263-13269)或在体内酵母双杂交分析中单独实验时，检测不到同源二聚体(参考例如 White 等，自然(1998)396: 679-682)。具体而言，Kammerer 等人通过沉降实验已经证明，GABA_B 受体 1 和 2 的异源二聚化序列单独实验时，在生理条件和生理体温下(如在 37°C)是以单体分子量沉降的。当以等摩尔量混合时，GABA_B 受体 1 和 2 异源二聚化序列以对应于这两种序列中的异源二聚体沉降(参看 Kammerer 等人的表 1)。此外，单个异源二聚化序列可以在宿主细胞中表达，同时，它可以通过许多蛋白分析，包括但不限于 SDS-PAGE, Western 印迹和免疫沉淀法来证明宿主细胞中同源二聚体的缺乏。体外实验必须在生理缓冲液条件下进行，和/或优选在生理体温下进行。一般而言，生理缓冲液中含有生理浓度的盐，调到中性 pH 范围 6.5-7.8 左右，优选 pH 是 7.0-7.5 左右。许多生理缓冲液列在上面 Sambrook 等人的文章中，因此，本发明中没有详细介绍。优选的生理条件在上述 Kammerer 等人的文章中已有详细的描述。

异源二聚化序列的特异性结合典型地包括非共价相互作用。这种

相互作用包括每一种现有的稳定键，它不会形成共价键。非共价键相互作用的例子包括，但不限于静电键，氢键，范德华力，两亲性肽的空间交错。

5 在设计受试的 Abu 过程中，进一步考虑的是减少异源二聚化序列与获得的异源多聚体抗原结合位点之间的结构干扰。可以采用本领域的各种不同技术来设计具极小内部结构干扰的嵌合异源多聚体。例如，一种方法涉及应用极少的异源二聚化序列，其中只含异源二聚化必需的氨基酸残基。第二种方法是把异源二聚化序列连接到产生的异源多聚体的 N 端或 C 端。选择哪一端取决于异源多聚体生物活性结构域的定位。为了构建一种抗原结合位点位于轻链和重链可变区 N 端一半部分的嵌合 Abu，优选的是把异源二聚化序列连接到轻链或重链的 C 末端。另一种选择的设计是在抗原结合位点和异源多聚体的异源二聚化序列间整合了“flexon”。本发明中使用的 flexon，指的是可弯曲的多肽接头(或编码这种多肽的核酸序列)，典型的 flexon 包括具小侧链的氨基酸(如甘氨酸，丙氨酸，缬氨酸，亮氨酸，异亮氨酸和丝氨酸)。在一个或多个受试 Abu 的位点间整合入 flexon，使它们具互相相对独立的构型，从而启动功能性。这样的构建一般能给抗原结合区提供额外的柔韧性。合适的 flexons 优选包括 4-100 个左右氨基酸，更优选的是 4-50 个左右氨基酸，尤为优选的是 4-15 个左右氨基酸。

25 用于构建受试 Abus 的异源二聚化序列可以有不同的来源。一般而言，任何参与形成稳定异源二聚体的蛋白序列均可成为候选的异源二聚化序列。同样，这些序列可来自任何异源多聚体蛋白复合物。代表性的候选序列有病毒蛋白，如腺伴随病毒的衣壳蛋白，和含 SH2 结构域的蛋白相互作用的蛋白激酶磷酸化位点(Cantely 等，(1993)，细胞，72: 767-778; Cantely 等，(1995)J. Biol. Chem. 270(44):26029-26032)，转录因子和异源二聚受体的结构域，它能介导形成异源聚体。

30 优选的异源多聚体转录因子有 Pal/Max 复合物和 Hox/Pbx 复合

物。Hox 代表的是一个大的转录因子家族，参与了胚胎发育过程中前后轴的图案结构。Hox 蛋白用一个保守的三 α 螺旋同源域结合 DNA。为了结合特异性的 DNA，Hox 蛋白要求存在异配偶体如 Pbx 同源域。为了理解 Hox-Pbx 复合物是怎样形成的以及它是如何和 DNA 结合的，
5 Wolberge 等人解决了 HoxB1-Pbx1-DNA 三源复合物的 2.35Å 晶体结构。这个结构表明每种蛋白的同源域结合在 DNA 另一侧的相邻识别序列上。在 Pbx1 螺旋 3 与螺旋 1 和 2 之间形成了一个口袋，通过这个口袋与 HoxB1 同源域 N 末端的 6 氨基酸的肽之间形成接触，进行异源二聚化。Pbx1 同源域的 C 末端延伸形成一个 α 螺旋，这个螺旋挤
10 进螺旋 1，形成一个大的 4 螺旋的同源域(Wolberger 等(1999)细胞 96: 587-597; Wolberger 等, J. Mol. Biol, 291:521-530)。

数目众多的异源二聚受体已被鉴定出来了。它们包括但不限于那些能和生长因子(如 heregulin)，神经递质(如 γ -氨基丁酸)和别的有机或无机小分子(如盐皮质激素，糖皮质激素)结合的受体。优选的异源二聚受体是核激素受体(Belshaw 等(1996), Proc. Natl. Acad. Sci. U.S.A. 93(10):4604-4607)，erbB3 和 erbB2 受体复合物，G 蛋白偶联受体和 GABA_B 受体家族，而 G 蛋白偶联受体包括，但不限于阿片样物质(Gomes 等, (2000)J. Neuroscience 20(22): RC110); Jordan 等(1999)Nature
15 399:697-700)，毒绳碱，多巴胺，血清素，腺苷/多巴胺。对大多数已知的异源二聚受体而言，其 C 端序列被发现能介导异源二聚体的形成。
20

如果需要，来自新的异源二聚受体的序列能用来构建受试的 Abus。在这种情况下，不需要太多的试验，通过任何遗传或生化实验就可鉴定给定受体对中的候选异源二聚化序列。此外，基于在相关和不相关基因中出现的共同结构域的序列同源性，计算机建模和搜索技术进一步便利了异源二聚化序列的检测。能进行同源性搜索的程序
25 的例子包括，但不限于 Blast(<http://www.ncbi.nlm.nih.gov/BLAST>)，Fasta(遗传学计算机压缩包，Madison, Wisconsin)，DNA star, Clustlaw, TOFFEE, COBLATH, Genthreader 和 MegAlign。任何含有对应于靶受
30

体或其区段的 DNA 序列的序列数据库均可用来进行序列分析。共同采用的数据库包括但不限于 GenBank, EMBL, DDBJ, PDB, SWISS-PROT, EST, STS, GSS 和 HTGS。

5 优选的另一类异源二聚化序列由亲水性肽组成，这种肽采用的是卷曲螺旋结构，螺旋化的卷曲螺旋是蛋白中的主要亚基寡聚序列之一。一级序列分析表明大约 2-3%的所有蛋白残基形成卷曲螺旋(Wolf 等人(1997)Protein Sci. 6:1179-1189)。已经鉴定的含螺旋卷曲的蛋白包括细胞骨架蛋白家族成员(如 α 角蛋白，波形蛋白)，细胞骨架运动蛋白家族成员(例如，肌球蛋白，驱动蛋白，和动力蛋白)，病毒膜蛋白(如 Ebola 或 HIV 的膜蛋白)，DNA 结合蛋白，和细胞表面受体(如 GABA_B 受体 1 和 2)。本发明的卷曲螺旋异源二聚化序列在广义上可以分为两类：即左手和右手卷曲螺旋。左手卷曲螺旋的特征有一个七重复序列(heptad repeat)，叫“abcdefg”，其非极性残基则优选地定位在第一(a)和第四(d)位。这两个位置的残基典型地构成了“隆突和洞”的 Z 字型结构，它和别的链上的残基连锁形成紧密的疏水核。相反，覆盖在卷曲螺旋外周的第二，三和六位的残基优选的是带电荷的残基。带电荷的氨基酸的例子包括碱性氨基酸残基，如赖氨酸，精氨酸，组氨酸；还包括酸性氨基酸残基，如天冬氨酸，谷氨酸，天冬酰胺，谷氨酰胺。

10 不带电荷或非极性的适合设计异源二聚体卷曲螺旋的氨基酸包括，但不限于甘氨酸，丙氨酸，缬氨酸，亮氨酸，异亮氨酸，丝氨酸和苏氨酸。尽管不带电荷的氨基酸残基典型地形成疏水核，但即便在核心位置，包括带电残基的螺旋间和螺旋内部的盐桥可以用来稳定整个卷曲螺旋结构(Burkhard 等(2000), J. Biol. Chem. 275:11672-11677)。尽管

15 可以采用不同长度的卷曲螺旋，但优选的受试的异源二聚化序列包括 2 到 10 个七序列重复。更优选的异源二聚化序列包括 3 到 8 个七序列重复，尤为优选的是包括 4 到 5 个七序列重复。

20 25

30 在设计优化的卷曲螺旋状异源二聚化序列的过程中，可采用许多现有的能预测肽的二级结构的计算机软件程序。例如，一种计算机分

析便使用了 COILS 算法，它能把氨基酸序列和数据库中已知的双链卷曲螺旋序列进行比较，同时预测高概率的卷曲螺旋序列(Kammerer 等 (1999), Biochemistry, 38:13263-13269)。

5 尽管许多不同的参与异源寡聚化的卷曲螺旋可以应用到本发明，但优选的是衍生自异源二聚受体的卷曲螺旋。因此，本发明包括衍生自 GABA_B 受体 1 和 2 的卷曲螺旋二聚体序列。一个方面，受试的卷曲螺旋包括 GABA_B 受体 1 和 2 的 C 端序列。另一方面，受试的卷曲螺旋进一步与半胱氨酸残基相连。卷曲螺旋是至少 30 个氨基酸残基的 GABA_B 受体 1 和 2 多肽，其中一个和 SEQ ID NO.2 中所示可比长度的线性序列基本相同，而另一个和 SEQ ID NO.4 中所示可比长度的线性肽序列基本相同。

15 如果两条序列的氨基酸序列基本上具同源性，则一条肽线性序列就和另一条肽线性序列“基本相同”。一般而言，基本相同的序列在比对过同源区以后，至少有约 60%的同一性。优选的序列至少有约 70%的同一性，至少约 80%，约 90%和约 95%的同一性是更优选的，而序列有 100%的同一性则尤为优选。

20 在确定多肽序列是否基本相同时，和正被比较的多肽比较，能保留其功能性的序列是特别优选的。功能性可以根据不同的标准确定，例如和配对的卷曲螺旋序列形成异源二聚体的能力，以及在生理缓冲条件和/或生理体温下形成同源二聚体的不稳定性。

25 本发明包括修饰的 GABA_B 异源二聚化序列，其在功能上等同于本发明列举的序列。能给产生的 Abus 提供改进的稳定性的修饰多肽是优选的。修饰的多肽的例子包括那些具保守氨基酸残基取代的多肽，以及哪些缺失或添加一个或多个氨基酸，但并不显著恶性的改变异源二聚化特异性的多肽。取代可以改变或修饰一个或多个氨基酸残基，直到完全重新设计一个区域，只要成对亲和力得以保留。如果存

30

在氨基酸取代, 优选的是保守取代, 不会不利地影响肽的折叠或功能特性。能发生保守性取代的功能相关氨基酸组有甘氨酸/丙氨酸; 缬氨酸/异亮氨酸/亮氨酸; 天冬酰胺/谷氨酰胺; 天冬氨酸/谷氨酸; 丝氨酸/苏氨酸/蛋氨酸; 赖氨酸/精氨酸; 和苯丙氨酸/酪氨酸/色氨酸。本发明的多肽可以是糖基化或非糖基化形式, 可以是转录后修饰(如乙酰化和磷酸化), 或者通过合成进行修饰(如连接一个标记基团)。

抗原结合单位(Abus)的构型和修饰

本发明的 Abus 可以采用各种不同的构型。最小的非单链 Abu 是单价 ccFv 片段。ccFv 片段是一个由 VL 和 VH 区组成的二聚体蛋白, 这两个区通过分别与 VL 和 VH 区框内融合的第一条和第二异源二聚化序列间的成对亲和力进行二聚化。优选的 ccFv 包含了一条短的 flexon 序列, 它给 VL 和 VH 区提供额外的柔韧性(参看图 1 中的 ccFv 的例子)。Nsc Abu 是一个更复杂的多价分子, 它能和多于 1 个的同种抗原(即多价但单特异性)或异种抗原(即多价和多特异 Abus)结合。典型的多价 Abu 是异源多聚体, 它由多于 1 条的 L 和 H 链多肽组成, 其中在 L 或 H 多肽中, 或者这两条多肽均包含多于 1 个的 V 区。例如, 双价 Abus 的例子呈现出如图 12 所描述的(ccFv)₂ 构型。这个双价 Abus 例中的 H 链多肽包含两个 VH 区, 其中每个 VH 区和 VL 区二聚化形成两个抗原结合位点。另外, L 链多肽可提供两个 VL 区, 其中每个 VL 区与 VH 区二聚化重构两个结合位点。如图 12 所示, 多价 Abu 通过连接在 VL 和 VH 区的两条异源二聚化序列的成对亲和力得到稳定。Abu 被有效地装配, 因为至少一个, 优选的是 2 个异源二聚化序列不能形成同源二聚体。因此, 使得分子内二聚化最少以形成非功能性 VH/VH 或 VL/VL 二聚体。应用一般的抗体工程改造方案, 可以构建三价或四价的 Abus(例如参看图 12)。

构建多价 Abus 的变通方法采用了如图 13 所示的 scFv 或 dsFv 片段。除了构建单位 ccFv 能提供一个抗原结合位点外, 这种构型的 Abus 包括一个或多个连接在 ccFv 上的 scFv 或 dsFv 片段。连接的 scFv 或 dsFv

能提供额外的结合位点。例如，双价 Abus 就可以采用 ccFv-scFv 或 ccFv-scFv 构型(图 13)。然而，一个抗原结合位点通过与 VL 和 VH 区相连的异源二聚化序列的配对亲和力进行装配(如 ccFv)，而另一个则由 scFv 或 dsFv 片段提供，其和 VL 区进行框内融合。另外，scFv 或 dsFv 片段可以连接到 VH 区。

可以利用同样的方法来构建如图 14 所示的三价 ccFv-scFv 或 ccFv-dsFv Abus。一方面，三价 Abus 具有 ccFv-(scFv)₂ 构型，在这种构型中，两条多肽“VH-第一异源二聚化序列-scFv”和“VL-第二异源二聚化序列-scFv”通过这两段异源二聚化序列的配对亲和力二聚化，从而构成三个结合位点。一个结合位点由 ccFv 结果单位的 VL 和 VH 区组成，剩余的 2 个结合位点由 scFv 片段提供，所述 scFv 片段连接到各自的 VL 和 VH 多肽。另外，多价的 Abus 可以构型成 ccFv-scFv-dsFv。在这种构型中，一个抗原结合位点通过位于 dsFv 片段 VH 和 VL 区的一对半胱氨酸残基之间的分子间二硫键装配和稳定。这种构型的进一步变体是三价的 ccFv-(dsFv)₂，在这种构型中，两个结合位点呈现 dsFv 形式(例如参看图 14)。不管是单特异性或多特异性的，采用 ccFv 结果单位的任何别的多价 Abus 变体均包含在本发明中。

因此，本发明进一步提供多特异的 Abus。它们是能结合至少两个不同抗原的多价分子。优选的多特异性 Abus 是双特异和三特异的分子，它们分别对 2 和 3 种不同抗原表现出结合特异性。和以前鉴定的多特异抗体不同的是(例如，参看美国专利 No.5932448)，受试的多特异性的 Abus 包括一个或多个具不同结合特异性的 ccFv 结果单位。受试的多特异 Abus 也能整合一个或多个如上所详细描述 scFv 或 dsFv 片段。根据图 15-17 描述的一般结构，优选的是双特异和三特异性 Abus 构型。

除了非单链 Abus，本发明包括单链 Abus，它被受试的异源二聚化序列所稳定。典型的 Sc Abus 包括 VL 和 VH 区，它们通过连接到

这两个区域的异源二聚化序列的配对亲和力形成分子内二聚体。异源二聚化序列的构型可以是平行的或反平行的方式(例如参看图 18)。在平行构型中, 这两段异源二聚化序列按相同方向排列(从氨基端到羧基端)。在反平行构型中, 异源二聚化序列排列时, 一序列的氨基末端和另一序列的羧基末端比对, 反之亦然。一般来说, 异源二聚化序列通过一种 flexon 序列连接在一起。如本发明描述的那样, flexon 是一种可弯曲的多肽接头(或编码这种多肽的核酸序列), 典型的 flexon 包含具小侧链的氨基酸(如甘氨酸, 丙氨酸, 缬氨酸, 亮氨酸, 异亮氨酸和丝氨酸)。整合在这两段异源二聚化序列的 flexon 为它们提供了空间柔韧性, 以便形成一个分子内二聚体。反平行构型中的合适 flexon 优选包括 4 到 100 个左右氨基酸, 更优选的是 4-50 个左右氨基酸, 尤为优选的是 4-15 个左右氨基酸。平行构型的 flexon 一般更长, 优选 10-100 个左右的氨基酸, 更优选 50-30 个左右氨基酸残基。

如果需要的话, 可以在异源二聚化序列的 N 或 C 末端整合进一对或多对半胱氨酸残基, 以进一步稳定本发明的 Abus。

本发明的 Abus 可含有衍生自 L 链或 H 链恒定区的序列。这种衍生自恒定区的序列一般置于轻链或重链可变区与它所连接的异源二聚化序列之间。此外, 轻链和重链可包含部分或全部人序列。

将非人抗体人源化的方法在本领域众所周知。“人源化”抗体是至少部分序列被改变的抗体, 已从原始形式变得更像人的免疫球蛋白。在一种形式中, H 链和 L 链 C 区被人序列取代。这是一个融合多肽, 它包括 V 区和异源免疫球蛋白 C 区。在另一种形式中, 尽管 V 框架区也被转变成人序列, 但 CDR 区包括非人抗体序列。例如, 参考 EP0329400。在第三种形式中, 通过设计人和小鼠 V 区的共有序列, 以及将与共有序列不同的 CDR 外的残基转变, 构建人源化 V 区。

在制造人源化抗体中, 框架区残基的选择对于保持高结合亲和力

是关键的。原则上，来自任何 HuAb 的框架序列均可作为 CDR 嫁接的模板，然而，已经证明 CDR 直接取代这样的框架区将会导致对抗原的结合亲和力显著下降。Glaser 等(1992), J. Immunol. 149:2606; Tempest 等(1992)Biotechnology 9:266; 和 Shalaby 等(1992)J. Exp. Med. 17:217。HuAb 和原始的 muAb 越具同源性，人框架区将可降低亲和力的畸变导入鼠源 CDR 的可能性就越小。根据在抗体序列数据库中进行的序列同源性搜索，HuAb IC4 和 muM4TS.22 框架区具很好的同源性，尽管别的高度同源的 HuAbs 也适合，尤其是来自人亚基 I 的 κ L 链或来自人亚基 III 的 H 链。Kabat 等(1987)。各种不同的计算机程序如 ENCAD(Levitt 等(1983)J. Mol. Biol. 168:595)均可用来预测用于 V 区的理想序列。因此，本发明包括具不同 V 区的 HuAbs。确定合适的 V 区序列以及优化这些序列是本领域一名专业技术人员范围内事情。免疫原性降低的抗体的获得方法在美国专利 No.5270202 和 EP699755 中介绍了。

对抗体加以人源化同时保持对抗原高亲和力和别的有利的生物学特性是重要的。为了达到这个目的，根据一种优选的方法，利用亲本和人源化序列的三维模型，通过亲本序列和不同概念的人源化产物的分析方法，可以制备出人源化抗体。三维免疫球蛋白模型是本领域的技术人员所熟知的。计算机程序可用来例证和展示选择的候选免疫球蛋白可能的三维构象结构。对这些候选的免疫球蛋白序列进行分析时，检查这些展示能分析残基可能的作用，换句话说，对能影响候选免疫球蛋白结合其抗原能力的残基进行分析。根据这种方法，可以从共有和重要序列中选择和组合 FR 残基，这样就能获得所需要的抗体特征，如对靶抗原的亲和力增加。

本发明也包括交联在化学官能团上的 Abus。典型的基团是能产生可以检测信号的标记。例如，这些交联的 Abus 在检测系统如肿瘤负担的定量，以及转移病灶造影和肿瘤造影中是有用的。这样的标记是本领域众所周知的，包括但不限于放射性同位素，酶，荧光化合物，

免疫毒素的制备在本领域是众所周知的。可以在例如“Monoclonal antibody-toxin Conjugates: Aiming the Magic Bullet”, Thorpe 等 (1982) Monoclonal Antibodies in Clinical Medicine, Academic 出版社, 第 168-190 页; Vitatta(1987) Science 238:1098-1104; 和 Winter 和 Milstein(1991) Nature 349:293-299 中找到制备的方法。合适的毒素包括, 但不限于蓖麻毒蛋白, 放射性核素, 美洲商陆抗病毒蛋白, 假单胞菌外毒素 A, 白喉毒素, 蓖麻毒蛋白 A 链, 真菌毒素如局限曲菌素和磷脂酶。一般参看“嵌合毒素”(“Chimeric Toxins”), Olsnes 和 Pihl, Pharmac. Ther. 15:355-381(1981); 和用于癌症检测和治疗的单克隆抗体(Monoclonal Antibodies for Cancer Detection and Therapy), Baldwin 和 Byers 编辑, 第 159-179 和第 224-266 页, Academic 出版社(1985)。

化学官能团可以通过重组技术来制备, 例如构建一个编码 Abu 和官能团的融合基因。此外, Abu 也可通过任何已知的化学方法结合到基团上。例如, 当基团是蛋白时, 可以用异源双官能交联接头如 SPDP, 戊二醛碳二亚胺等进行连接。这些基团可以通过二级试剂如第二抗体, 蛋白 A, 或生物素-亲和素复合物进行共价连接, 或交联。顺磁基团和其抗体交联物在本领域众所周知。例如, 参看 Miltenyi 等 (1990) Cytometry 11: 231-238。

抗原结合单位(Abus)的制备

可以用重组 DNA 技术, 化学合成技术或它们的组合技术来制备受试的 Abus。例如, 典型组装编码包括 VL, VH 和异源二聚化序列在内的所需 Abus 成份的序列, 并将其片段连接到表达载体上。这些序列可以从编码所需蛋白序列的别的载体中组装, 可以利用各自模板核酸从 PCR 产生的片段中组装, 或者组装编码所需序列的合成寡核苷酸。然而, 所有编码 Abus 的核酸序列优选地通过框内融合编码序列来进行组装。为了增强每种成份呈现彼此相对独立构型的能力, 可以将上面介绍的 flexon 导入不同成份和结构域之间。为了制备 Nsc Abus, 可以先单独形成 L 和 H 链, 然后再组装, 或者通过两条链的表达系统

进行在体组装。用包含 L 和 H 链的各自可转录区的载体转染合适的细胞，或者用每条链的载体共转染同一细胞，可产生这样的表达系统。

5 组装的 Abus 可以用本领域已知的各种不同蛋白质纯化技术进行分离。一般来说，当想直接获得没有信号肽的 Abu 时，可以从培养基中分离分泌的 Abu，尽管它们可以从宿主细胞溶菌产物或细菌周质中回收。如果 Abus 是膜结合的，则可用本领域技术人员共用的合适的去污剂溶液来溶解。回收的 Abus 可以通过下列方法进一步加以纯化：
10 盐析(例如用硫酸铵)，离子交换层析(例如在中性 pH 下的阳离子或阴离子交换柱以及用离子强度增强的分级梯度洗脱)，凝胶过滤层析(包括凝胶过滤 HPLC)，以及在标记亲和柱或亲和树脂如蛋白 A，蛋白 G，羟磷灰石和抗免疫球蛋白上进行层析。

本发明的多核苷酸和载体

15 本发明提供了编码本发明 Abus 的各种不同多核苷酸。在一定程度上说，本发明的多核苷酸的特征在于含有上面详细介绍的独特异源二聚化序列。这种异源二聚化序列能有效地组装和筛选 Abus，例如那些特异性结合到所需抗原的异源二聚化序列。这种序列也能便利异源多聚体在活的生物实体包括噬菌体，细菌，别的原核或真核细胞中的
20 展示。优选的异源二聚化序列如 SEQ ID NOS.2 和 4 所示。

在一种实施方案中，本发明提供了编码受试 Nsc Abus 的分离的多核苷酸。在这个实施方案中的一方面，重组多核苷酸包括编码受试 Nsc Abu 轻链多肽的编码序列。另一方面，重组多核苷酸包括编码受
25 试 Nsc Abu 重链多肽的编码序列。还有另一方面，重组多核苷酸包括两条独立的编码序列，一条编码轻链多肽，另一条编码重链。

通过一些常规的技术，包括但不限于杂交，PCR 和 DNA 测序可以轻易获得对应于现有抗体 L 或 H 链不同区的核酸序列并测序。产生
30 单克隆抗体的杂交瘤细胞可以作为抗体核苷酸序列的优选来源。数目

众多的产生单克隆抗体阵列的杂交瘤细胞可以从公立或私有宝库中获得。最大的贮藏机构是美国典型培养物保藏中心(American Type Culture Collection)(<http://www.atcc.org>)，它能提供不同的鉴定好的杂交瘤细胞系。此外，可以从免疫或非免疫过的啮齿类动物或人上获得抗体核苷酸，也可从器官如脾和外周血淋巴细胞中获得抗体核苷酸。适用于抽提和纯化抗体核苷酸的具体技术在 Orlandi 等(1989)Proc. Natl. Acad. Sci. U.S.A 86:3833-3837; Larrick 等(1989)Biochem. Biophys. Res. Commun. 160:1250-1255; Sastry 等(1989)Proc. Natl. Acad. Sci. U.S.A. 86:5728-5732; 和美国专利 No.5,969,108 中。

抗体核苷酸序列也可以被修饰，例如，在同源的非人源序列上用 人重链和轻链恒定区的编码序列取代。通过这种方式，可以制备出仍保留原始抗体结合特异性的嵌合抗体。

本发明包括的多核苷酸包括那些例证多肽的功能性等同物及其片段的编码核苷酸也能理解。功能等同的多肽包括那些能增强，降低或不会显著地影响编码多肽的特性的多肽。功能等同物可以是具保守氨基酸取代的多肽，包括融合类似物以及突变体。

因为遗传密码的简并性，因此在 L 和 H 序列以及适合构建本发明的多核苷酸和载体的异源二聚化序列中有很大的变异。序列变体可以是修饰的 DNA 或氨基酸序列，一或多个取代，缺失，或添加，这种变异的净效果是保留了所需的抗原结合活性。例如，可以在编码区进行不同的取代，这种取代既不会改变编码的氨基酸，也不会导致保守改变。这些取代包括在本发明中。保守的氨基酸取代包括下列组的氨基酸取代：甘氨酸，丙氨酸；缬氨酸，异亮氨酸，亮氨酸；天冬氨酸，谷氨酸；天冬酰胺，谷氨酰胺；丝氨酸，苏氨酸；赖氨酸，精氨酸；和苯丙氨酸，酪氨酸。尽管保守性取代确实会有效地改变将要产生的多肽中所含的一个或多个氨基酸残基，但这种取代不会按预料的那样干扰产生的 Abus 的抗原结合活性。不改变编码氨基酸残基的核

核苷酸取代可以用于优化不同系统的基因表达。合适的取代是本领域技术人员众所周知的。例如，这反映了表达系统密码使用的偏爱性。

如果需要，重组的多核苷酸可以包含异源序列，其便利表达检测和基因产物的纯化。这种序列的例子在本领域众所周知的，并包括那些编码报告蛋白如 β -半乳糖苷酶， β -内酰胺酶，氯霉素乙酰转移酶 (CAT)，荧光素酶，绿色荧光蛋白(GFP)及它们的衍生物的序列。另一种有助于纯化的异源序列可以编码表位，如 Myc，HA(衍生自流感病毒血凝素)，His-6，FLAG，或免疫球蛋白 Fc 部分，谷胱甘肽-S-转移酶 (GST)，和麦芽糖结合蛋白(MBP)。

多核苷酸可以交联到上面描述的不同化学官能团。共用的基团包括能产生检测信号的标记，信号肽，能增强免疫活性的试剂，有助于偶联到固相支持物上的试剂，疫苗载体，生物反应效应物，顺磁标记和药物。基团可通过重组技术或本领域众所周知的其他技术共价连接到多核苷酸上。

本发明的多核苷酸包括附加序列，如在同一转录单位中的附加编码序列，控制元件，如启动子，核糖体结合位点和多聚腺苷化位点，在同一或不同启动子控制下的附加转录单位，能进行克隆，表达以及转化宿主细胞的序列，以及任何诸如可理想的提供本发明实施方案的构建体。

本发明中包含的多核苷酸可以通过化学合成，重组克隆方法，PCR 或它们的任何组合方法来获得。多核苷酸化学合成方法是本领域众所周知的，因此没必要再次详细叙述。本领域的技术人员通过 DNA 合成仪或从商业上订购可以从本发明提供的序列数据库获得所需要的多核苷酸。

含有所需序列的多核苷酸可以插入合适的载体，然后，导入合适

的宿主细胞进行复制和扩增。因此，本发明包括的许多载体含有本发明的一个或多个多核苷酸。此外，本发明也提供了表达载体的选择文库，其中至少一种载体编码受试 Abus。

5 本发明的载体一般可分为克隆载体和表达载体。克隆载体用于复制它们含有的多核苷酸的拷贝数，或者作为一种储存库中多核苷酸的工
15 具，以便以后回收。表达载体(和含有这些表达载体的宿主细胞)可以用来获得它们含有的多核苷酸所表达的多肽。合适的克隆和表达载体可以包括本发明的任何载体，例如用于细菌，哺乳动物，酵母，昆
20 虫和噬菌体展示表达系统的载体。

 合适的克隆载体根据基本的方法可以构建，也可从本领域数目众多的载体中筛选。尽管根据准备使用的宿主细胞，克隆载体有所变化，但有用的克隆载体一般都有自主复制的能力，都有特定的限制性内切
15 核酸酶作用的单一靶点，或携带标记基因。合适的例子包括质粒和细菌病毒，如 pBR322, pMB9, ColE1, pCR1, RP4, pUC18, mp18, mp19, 噬菌体 DNA(包括丝状和非丝状噬菌体 DNA)，以及穿梭载体如 pSA3 和 pAT28。这些载体和别的克隆载体可以从商业公司如 Clontech, BioRad, Stratagene 和 Invitrogen 获得。

20 含有这些核酸的表达载体可以用于获得宿主载体系统来制备蛋白和多肽。表达载体暗含它们在宿主生物体中或可作为附加体，或能作为染色体 DNA 的一部分被复制。合适的表达载体包括质粒，病毒载体，包括噬菌粒，腺病毒，腺伴随病毒，反转录病毒，粘粒等。许多
25 适合在真核细胞，包括酵母，禽类和哺乳动物细胞中表达的表达载体在本领域中众所周知。一个表达载体的例子是 pcDNA3(Invitrogen, San Diego, CA)，在这个载体中，转录是由巨细胞病毒(CMV)早期启动子/增强子所驱动。两类特定用于表达受试 Abus 的表达载体是噬菌体展示载体和细菌展示载体。

30

构建噬菌体载体的技术在本领域中已经建立好了(参看 Winter G 等(1994)的综述, *Ann. Rev. Immunol.* 12:433-55)。丝状噬菌体和非丝状噬菌体序列均可适用于构建展示载体。丝状噬菌体载体是优选的, 因为这类噬菌体中许多有代表性的噬菌体基因组已经被测序, 而且发现它们的基因组比非丝状噬菌体的基因组要小。这类噬菌体中的代表包括 M13, f1, fd, If1, Ike, Xf, Pf1 和 Pf3。噬菌体载体典型地构建来表达异源多聚体, 如抗体肽, 通过融合到部分或全部噬菌体外壳蛋白上。合适的外壳蛋白包括 M13 的 pIII, VIII, VI, VII 和 IX。采用不损害表达噬菌体外壳完整性的方式, 可以将异源多聚体序列插入噬菌体载体, 而优选的异源多聚体是具生物功能的。

为了构建 PIII融合载体, 共用的融合位点位于氨基端, 介于柔韧间隔序列和 PIII的两个结构域之间的位点(Smith 等, *Science* 288:1315-17), 或任何别的在美国专利 Nos.5969108 和 5837500 介绍的可替代融合位点。噬菌体的 PIII融合蛋白和别的蛋白可以完全由同一个噬菌体(page)复制子编码, 也可由不同复制子编码。当至少使用两个复制子时, PIII融合蛋白一般在噬菌粒中编码, 噬菌粒是含噬菌体复制起始位点的质粒。噬菌粒通过辅助噬菌体如 M13KO7 的“补救”, 包装进噬菌体颗粒中, 辅助噬菌体能提供所有的噬菌体蛋白, 包括 PIII, 但它由于缺乏起始位点, 在和噬菌粒竞争中, 本身很难包装进噬菌体。别的缺乏 PIII或包含能增强包装效率的修饰 PIII的多价辅助噬菌体(如 M13 Δ gIII)也可以采用(Rondot 等, *Nature Biotechnology*, 19:75-78)。

对别的丝状噬菌体也可进行同样的构建。Pf3 是一种众所周知的丝状噬菌体, 它能感染含 IncP1 质粒的绿脓假单胞菌(*Pseudomonas aeruginosa*)细胞。整个基因组已经测序, 而参与复制和组装的遗传信号已经鉴定出来。PF3 主要的外壳蛋白在没有信号肽指导分泌的情形下并不常见。序列有带电的残基 ASP₇, ARG₃₇, LYS₄₀ 和 PHE₄₄-COO-, 它们和被暴露的氨基端相一致。为了构建展示 Pf3 载体, 一般按需要, 通过基因工程把绿脓杆菌(*P. aeruginosa*)中已知能引起分泌的信号序列

框内融合到编码异源多肽的基因片段上，结果，编码异源多肽的基因片段又和编码成熟 Pf3 外壳蛋白的 DNA 框内融合在一起。

同样的一般构建方案用于产生含来自非丝状噬菌体的序列的展示载体，非丝状噬菌体包括细菌噬菌体 X174， λ ，T4 和 T7 噬菌体。这些非丝状噬菌体丰富的结构信息是本领域众所周知的。本领域的技术人员不用过多的实验就能轻易地获得相应的展示载体，这种载体利用唯一的异源二聚化序列来表达受试的异源多聚体。

除了噬菌体展示载体，另一类优选载体是细菌展示载体。上面概括的一般方案同样适用于构建这种载体。简而言之，这些载体使和细菌表面蛋白融合的异源多聚体的表达更容易。尤其是 Abus。以前的研究表明，数目众多的细菌表面蛋白适用于表达这种融合蛋白。非限制的细菌表面蛋白的例子有 LamB(Bremer 等，Proc. Natl. Acad. Sci. U.S.A(1984), 81:3830-34; Gene (1987)52:165-73; OmpA(Progs Biophys Molec Biol (1987)49:89-115); OmpC; OmpF(Pages 等，Biochimie (1990)72:169-76); PhoE(van der Ley 等，J. Biol. Chem. 261:12222-5); pilin(So 等，Curr. Top. in Microbiol & Immunol (1985)118:13-28); pldA(de Genus 等，EMBO. J., (1984)3(8):1799-1802)及它们的同源物。这些和别的表面蛋白的鉴定，以及使用这些蛋白用于展示异源多肽的方法在美国专利 No.5837500 以及它所引用的参考文献中已有详细描述。

本发明的载体一般包括一个 Abus 表达所需要的转录或翻译控制序列。合适的转录或翻译控制序列包括，但不限于复制起始位点，启动子，增强子，阻遏蛋白结合区，转录起始位点，核糖体结合位点，翻译起始位点，以及转录和翻译的终止位点。

如本发明使用的那样，“启动子”是一个 DNA 区，它在某些条件下能结合 RNA 聚合酶并起始启动子下游编码区的转录(按 3'方向)。

启动子可以是组成型或诱导型的。总之，启动子序列 3'端邻接的是转录起始位点，并向上游延伸(5'方向)，包括在高于背景的可检测水平上起始转录所必需的最小数目的碱基或元件。转录起始位点以及负责结合 RNA 聚合酶的蛋白结合区均在启动子序列中。真核启动子经常，
5 但并不总是包含“TATA”框和“CAT”框。

启动子的选择主要取决于载体导入的宿主细胞。对动物细胞而言，不同的强启动子，病毒或非病毒启动子均是本领域众所周知的。非限制的代表性病毒启动子包括 CMV，SV 病毒的早期或晚期启动子，
10 不同类型腺病毒(如腺病毒 2)以及腺伴随病毒的启动子。如果这种控制序列和宿主细胞系统相容的话，利用和所需的轻链或重链基因正常结合的启动子也是可能的，而且通常是理想的。

用于别的真核细胞的合适启动子序列包括 3-磷酸甘油酸激酶启动子，或别的糖酵解酶如烯醇化酶，甘油醛-3-磷酸脱氢酶，己糖激酶，
15 丙酮酸脱羧酶，磷酸果糖激酶，葡萄糖-6-磷酸异构酶，3-磷酸甘油酸变位酶，丙酮酸激酶，丙糖磷酸异构酶，磷酸葡萄糖异构酶，和葡萄糖糖激酶的启动子。其他由生长条件控制的附加转录优点的启动子有醇脱氢酶 2，isocytochrome C，酸性磷酸酶，和氮代谢相关的降解酶，
20 以及前述的甘油醛-3-磷酸脱氢酶和参与麦芽糖和半乳糖利用的酶的启动子区域。

在某个优选的实施方案中，本发明的载体使用的是强增强子和启动子表达盒。这种表达盒的例子包括人巨细胞立即早期(HCMV-IE)启动子(Boshart 等，细胞 41: 521，(1985))， β -肌动蛋白启动子(Gunning
25 等(1987)Proc. Natl. Acad. Sci(USA)84:5831)，组蛋白 H4 启动子(Guild 等(1988)，J. Viral. 62:3795)，小鼠金属硫蛋白启动子(McIvor 等(1987)，Mol. Cell. Biol. 7:838)，大鼠生长激素启动子(Millet 等(1985)，Mol. Cell. Biol. 5:431)，人腺苷脱氨酶启动子(Hantzopoulos 等(1989)，Proc. Natl
30 Acad. USA86:3519)，HSV tk 启动子 25(Tabin 等(1982)Mol. Cell. Biol.

2:426), α -1 抗胰蛋白酶增强子(Peng 等(1988)Proc. Natl. Acad. Sci. USA 85:8146)以及免疫球蛋白增强子/启动子(Blankenstein 等(1988)Nucleic Acid. Res. 16:10939), SV40 早期或晚期启动子, 腺病毒 2 主要晚期启动子, 或别的衍生自多瘤病毒, 牛乳头瘤病毒或别的反转录病毒或腺病毒的病毒启动子。免疫球蛋白(Ig)基因的启动子和增强子元件赋予 B 淋巴细胞有标记特异性(Banerji 等(1983)细胞 33: 729; Gillies 等(1983)细胞 33: 717; Mason 等(1985), 细胞 41: 479), 尽管控制 B 珠蛋白基因转录的元件仅在类红细胞中发挥功能(van Assendelft 等(1989)细胞 56: 969)。

细胞特异或组织特异的启动子也可以使用。数目众多的不同组织特异的启动子已经被本领域的技术人员所描述和采用。在选择性动物细胞中可操作地启动子例子包括肝细胞特异的启动子和心肌细胞特异的启动子。根据选择的受体细胞类型, 本领域的技术人员知道别的适合用来构建本发明表达载体的细胞特异或组织特异的启动子。

根据本发明, 利用众所周知的限制性酶切和连接技术, 可以从不同的 DNA 中切除合适的转录控制序列, 然后可操作地整合到要表达的完整选择性融合基因上。

在构建受试载体过程中, 和外源序列结合的终止序列插入到需要转录的序列的 3'端, 以提供 mRNA 聚腺苷酸信号和/或转录终止信号。优选的终止子序列含有一个或多个转录终止序列(如聚腺苷酸化序列), 同时也可通过添加 DNA 序列来加以延长, 以便进一步破坏转录通读。本发明优选的终止子序列(终止位点)有一个基因, 其后跟一个转录终止序列, 既可是基因自己的终止序列, 也可是异源的终止序列。这种终止序列的例子包括和不同聚腺苷酸化序列偶联的终止密码子, 这些聚腺苷酸化序列是本领域众所周知的, 广泛运用并在下面例证了的。如果终止子中含有基因, 利用编码检测或选择标记的基因是有益的, 因此, 提供了一种方法, 通过这种方法可以检测或选择终止子序

列的存在和/或缺失(相应于转录单位的失活和/或激活)。

除了上述的元件，载体可以包括选择标记(例如，一个蛋白的编码基因，该蛋白是转化有载体的宿主细胞存活或生长所必需的)，尽管
5 这些标记基因可以包含在另外的多核苷酸序列上，被共导入宿主细胞中。只有那些导入了选择基因的宿主细胞在选择条件下才可以存活和/或生长。典型的选择基因编码的蛋白(a)对抗生素或别的毒素如氨苄青霉素，新霉素，G418，氨甲蝶呤等具抗性；(b)补充营养缺陷；或(c)补充一些从复合培养基中得不到的关键营养物。合适标记基因的选择
10 取决于宿主细胞，而适合不同宿主的基因是本领域众所周知的。

在一个优选的实施方案中，载体是穿梭载体，它能在至少两种不相关的表达系统中复制。为了方便这种复制，这种载体一般含有至少两个复制起始位点，在各个表达系统中分别是有效的。典型的穿梭载体
15 能在真核表达系统中与原核表达系统中复制。这使得在真核宿主细胞(表达细胞类型)中检测蛋白表达以及在原核宿主细胞(扩增细胞类型)扩增载体成为可能。尽管任何本领域已知的复制起点均可使用，只要它能指导载体的复制，但优选的复制起点还是一个来自 SV40，一个衍生自 pBR322。如果这个载体是个穿梭载体，则这个载体优选含有
20 至少两个选择标记，一个用于表达细胞类型，另一个用于扩增细胞类型。任何本领域已知的或本发明介绍的均可使用，只要它们在被利用的表达系统中能发挥功能。

本发明包含的载体可以通过重组克隆技术和/或化学合成获得。
25 数目众多的重组克隆技术如 PCR，限制性内切核酸酶消化以及连接都是本领域众所周知的技术，没必要在这详细的介绍。本领域的技术人员也可利用本发明提供的序列数据或在公共或私有数据库中的数据通过本领域的任何合成方法去获得所需要的载体。

30 本发明的宿主细胞

本发明提供用上述载体或表达载体文库转染的宿主细胞。表达载体通过任何一种合适的方法，如电穿孔，微粒轰击，脂转染，侵染(在这种方法中，载体和侵染剂偶联在一起)，用 CaCl_2 ，氯化铷，磷酸钙，DEAE-葡聚糖或别的物质进行的转染，导入合适的原核或真核细胞。

5 载体导入方法的选择通常取决于宿主细胞的特征。

对于绝大部分动物细胞，上述任何方法均适合载体转移。优选的动物细胞是脊椎动物细胞，脊椎动物细胞中优选哺乳动物细胞，它能大量表达异源导入的基因产物，例如，在毫克水平上。非限制的优选细胞例子是 NIH3T3 细胞，COS, HeLa 和 CHO 细胞。

10

动物细胞可以在不同培养基中培养。商业上获得的培养基如 Ham 氏 F10(Sigma)，基本必需培养基(MEM, Sigma)，RPMI-1640(Sigma)和 Dulbecco 氏改良的伊格尔氏培养基(DMEM, Sigma)都适合培养宿主细胞。此外，动物细胞可以在合成培养基中生长，这种培养基缺乏血清，但补充有激素，生长因子或任何别的对特定细胞类型生存和/或生长所必需的因子。然而，支持细胞存活的合成培养基保持生存力，形态，代谢能力和潜能，细胞分化的能力，但促进细胞生长的合成培养基提供了细胞增生或增殖所必需的化学物质。控制哺乳动物细胞在体外存活和生长的一般参数在本领域中已经确定了。在不同的细胞培养系统中被控制的物化参数如 pH， pO_2 ，温度和摩尔渗透压浓度。细胞所必需的营养通常以标准培养基配方提供，开发这样的配方是为了优化环境。营养物可以分成若干类：氨基酸及衍生物，碳水化合物，糖，脂肪酸，复合脂，核酸衍生物和维生素。除了维持细胞代谢的营养之外，大部分细胞还需要一种或多种激素，它们选自下列组中的至少一种：类固醇，前列腺素，生长因子，垂体激素和用来在无血清培养基中扩增的肽激素(Sato, G.H.等“细胞在激素合成培养基中的生长”(Growth of cells in Hormonally Defined Media”，冷泉港出版社，N.Y.1982)。为了在体外存活和生长，除了激素，细胞还要求转运蛋白，如运铁蛋白(血浆铁转运蛋白)，血浆铜蓝蛋白(铜转运蛋白)，和高密度脂蛋白(脂载

15

20

25

30

体)。每一套优化的激素或转运蛋白将根据细胞类型有所变化。大部分激素和转运蛋白从外部加入，或者，在很少的情况下，已发现突变的细胞系不需要某个特定因子。不需太多的实验，本领域的技术人员将会知道保持细胞培养所必需的别的因子。

5

对于植物细胞，本领域有不同的载体转移技术可用。宿主细胞的形式可以是整株植物，分离的细胞或原生质体。举例的载体导入植物细胞的方法包括农杆菌介导的植物转化，原生质体转化，基因转入花粉，注射入生殖器官以及注射入未成熟胚。每一种方法有其长处与不足，这一点对本领域的一个技术人员而言很明显，因此，把载体导入具体植物物种的特定方法没必要对另一种植物物种最有效。

10

农杆菌(*Agrobacterium tumefaciens*)介导的转化是把载体导入植物细胞的一种广为采用的系统，因为载体能导入整个植物组织，没必要从原生质体生成完整植株。利用农杆菌介导的表达载体把载体导入植物细胞是本领域众所周知的。农杆菌通过把自己的一部分DNA(T-DNA)转入宿主细胞，这部分DNA整合进核DNA，在植物中定居下来，这个技术就是利用了农杆菌的这个共同特征。T-DNA的定义是长25个碱基对的边界序列，介于这个边界序列之间的任何DNA也能转入植物细胞。在T-DNA边界序列之间插入重组的植物病毒核酸会导致重组植物病毒核酸向植物细胞的转化，其中重组植物病毒核酸在植物细胞中复制，然后经系统扩散到整个植物。

15

20

因为不是所有的植物都是农杆菌的天然宿主，可以采用别的方法如原生质体转化将受试载体导入宿主细胞。对某些单子叶植物而言，可以利用基于磷酸钙沉淀，聚乙二醇处理，电穿孔及这些处理的组合的方法可以进行植物原生质体的转化。

25

除了原生质体转化外，粒子轰击是一种可供选择和方便的技术，它把本发明载体转移到植物宿主细胞。具体来说，植物细胞可以用包

30

被有多个受试载体的微粒子轰击。用包被 DNA 的微粒轰击已在植物和动物中成功地获得了稳定的转化子(例如, 参看 Sanford 等(1993)Methods in Enzymology, 217:483-509)。适合把载体导入植物细胞的微粒典型地是由金属制成, 优选的是钨或金。例如, 这些微粒可以从 Biolad 公司获得(如 Bio-Rad 的 PDS-1000/He)。本领域技术人员均知道, 对任何植物来说, 改变参数如 He 压力, 包被粒子的数量, 大载体和阻挡屏之间的距离以及从阻挡屏到靶点的飞行距离, 均可以优化粒子轰击方案。

如 Zhou 等人描述的那样, 通过直接把 DNA 转化花粉, 也可以把载体导入植物, Methods in Enzymology, 101:433(1983)D. Hess. Intem. Rev. Cytol., 107:367(1987); Lou 等, Plant Mol. Biol, Reporter, 6:165(1988)。另外, 也可如 Pena 等人(Nature, 325:274(1987))介绍的那样, 把载体注射入植物的繁殖器官。

别的把核酸导入植物细胞的技术包括:

(a)手工接种。手工接种是在中性 pH, 低摩尔浓度磷酸缓冲液中进行的, 还添加了硅藻土或金刚沙(通常大约是 1%)。在叶子上表面滴加 1-4 滴制剂, 然后轻轻涂开。

(b)植物床 (plant bed) 的机械接种。当拖拉机牵引的割草机切割叶子时, 把载体溶液喷洒入(气体驱动的)割草机, 完成植物床的接种。或者, 也可以在植物床被切割后, 立即把载体溶液喷洒入切割的叶子上。

(c)对单叶进行高压喷洒。单叶接种的做法是用含大约 1%金刚砂的载体缓冲液狭窄直接喷洒(50psi, 离叶面 6-12 英尺)叶子。

(d)真空渗入。为了便于感染, 把宿主生物体置于基本上是真空压力的环境下, 完成接种。

别的适合克隆和表达受试载体的宿主细胞有原核和真核微生物如真菌或酵母细胞。适合这种用途的原核生物包括细菌, 细菌中包含革

兰氏阴性和革兰氏阳性细菌。这类微生物的代表有肠内菌科(如大肠杆菌), 肠杆菌属, 欧氏杆菌属, 克雷伯氏杆菌属, 变形杆菌属, 沙门菌属(如鼠伤寒杆菌(*Salmonella typhimurium*)), 沙雷氏菌属(例如, 粘质沙雷菌(*Serratiamarcescans*)), 志贺菌属, 奈瑟球菌属(如脑膜炎奈瑟球菌(*Neisseria meningitidis*))以及杆菌属(如枯草杆菌(*Bacilli subtilis*)和地衣芽孢杆菌(*Bacilli licheniformis*))。优选的宿主细胞分泌极少量的表达 Abus 的蛋白水解片段。公用的真菌(包括酵母)宿主细胞有酿酒酵母(*S. cerevisiae*), 乳酸克鲁维酵母(*Kluyveromyces lactis*)(*K. lactis*), 念珠菌种属包括白色念珠菌(*C. albicans*)和光滑念珠菌(*C. glabrata*), 麦芽糖念珠菌(*C. Maltosa*), 产朊假丝酵母(*C. utilis*), 类星型念珠菌(*C. stellatoidea*), 近平滑念珠菌(*C. parapsilosis*), 热带假丝酵母(*C. tropicalus*), 链孢菌(*Neurospora crassas*), 构巢曲霉(*Aspergillus nidulans*), 粟酒裂殖酵母(*Schizosaccharomyces pombe* (*S. Pombe*)), 巴氏毕赤酵母(*Pichia pastoris*)和 *Yarrowia lipolytica*。

一旦导入合适的宿主细胞, Abus 的表达可以用本领域已知的任何核酸与蛋白实验来进行检验。例如, 可以用常规的杂交实验(如 Northern 印迹分析), 扩增方法(如 RT-PCR), SAGE(美国专利 No5695937)以及以阵列为基础的技术(例如参看美国专利 Nos5405783, 5412087 和 5445934), 利用和 Abu 多核苷酸任何区域互补的探针, 对 L 或 H 链, 以及 Sc Abu 的转录 mRNA 的存在加以检测和/或定量分析。

载体的表达也可通过检查 Abu 的表达来确定。本领域的许多技术可用来进行蛋白分析。它们包括, 但不限于放射性免疫测定, ELISA(酶联免疫吸附测定), “三明治”免疫实验, 免疫放射分析, 原位免疫实验(例如使用胶体金, 酶或放射性同位素标记), Western 印迹分析, 免疫沉淀实验, 免疫荧光实验和 SDS-PAGE。

本发明的多核苷酸, 载体和宿主细胞的使用

本发明的多核苷酸和载体有若干特殊的用法。例如, 它们用于生

产 Sc 和 Nsc Abus 的表达系统。多核苷酸也能用作引物去影响所需多核苷酸的扩增。此外，本发明的多核苷酸还能用于药物组合物，包括疫苗，诊断剂和药物。

5 本发明的宿主细胞基于它们的抗原结合特异性，尤其能用作受试多核苷酸，载体的宿主，或作为产生和筛选所需 Abus 的媒介物。

10 因此，本发明提供一种能鉴定和所需抗原进行免疫反应的 Nsc Abu 的方法。所述方法包括下列几步：(a)制备遗传多样性的 Abus 库，其中所述库包括至少一个受试 Abus；(b)将抗原结合单位库和所需抗原接触；(c)检测 Abus 和抗原之间的特异性结合，从而确定和所需抗原发生免疫反应的 Abu。

15 Abu 和所需抗原特异结合的能力可以用本领域建立好的各种不同方法来进行实验。参考 Harlow 和 Lane(1988)抗体：实验手册，冷泉港实验室，纽约，Gherardi 等(1990)J. Immunol. Meth. 126:66-68。典型地，表现出所需结合特异性的 Abu 可以用免疫实验直接检测出来，例如，使标记的 Abus 和固定在固相支持物或基质上的抗原反应。总之，抗原粘附的基质由一些物质组成，这些物质在免疫实验中表现出低水平非特异结合。优选的固相支持物由下列一种或多种类型的材料组成：塑料多聚体，玻璃，纤维素，硝化纤维素，半导体材料和金属。优选的基质是培养皿，层析珠，磁珠等。

25 对于这种固相实验，未反应的 Abus 可以通过冲洗去掉。然而，在液相实验中，未反应的 Abus 是通过别的分离技术，如过滤或层析去掉的。抗原结合到标记的 Abus 上以后，结合标记的数量就被确定了。这种技术的变体是竞争实验，在竞争实验中，抗原和原始结合分子结合到饱和状态。当一群受试的 Abus 被导入这个复合物时，只有那些表示出更高亲和力的 Abus 能竞争，因此仍然结合在抗原上。

30

此外，与给定抗原的特异结合可通过细胞分选来测定，细胞分选包括把所需的抗原呈递到要分选的细胞上，然后用偶联有检测剂的 Abus 标记靶细胞，紧接着在细胞分选器上把标记的细胞从未标记的细胞中分离出来。一个复杂的细胞分离方法是荧光激活细胞分类术 (FACS)。一系列流行走的细胞通过激光束时，然后检测每个被荧光标记的 Abus 所结合的细胞的荧光。

随后对洗脱的 Abus 的分析涉及描述 L 和 H 链的氨基酸序列的蛋白质测序。根据推导的氨基酸序列，然后通过重组克隆技术包括 PCR，文库筛选，对现有核酸数据库的同源型搜索，或它们的组合方法可以获得编码抗体多肽的 cDNA。共用的数据库包括但不限于 GenBank, EMBJ, DDBJ, PDB, SWISS-PPOT, EST, STS, GSS 和 HTGS。

当全部 Abus 在噬菌体或细菌粒子中展示时，选择方法优选亲和层析。这种方法典型的是把全部噬菌体 Abus 结合剂抗原包被平板，柱基体，细胞或溶液中的生物素化抗原，然后捕获 Abus。冲洗结合到固相的噬菌体或细菌，然后用可溶性半抗原，酸或碱洗脱。此外，增加抗原的浓度能把 Abus 从亲和基体中解离出来。对某种与抗原具极高亲和性或亲和力的 Abus 而言，有效的洗脱要求如 WO92/01047 所介绍的高 pH 或温和的还原溶液。

为了避免回收具所需结合特异性的结合 Abus 的潜在困难，在异源二聚化序列和用来展示 Abus 的噬菌体外壳蛋白之间导入蛋白酶剪切位点。适用于这个目的的蛋白酶剪切位点包括但不限于 X 因子，胰蛋白酶和凝血酶识别位点。把全部噬菌体结合到亲和基体并洗脱掉非特异的噬菌体后，展示具所需亲和力的 Abus 的剩余噬菌体。在适合于剪切位点消化的条件下通过用蛋白酶冲洗抗原亲和基体可得以收集。这种消化将会把 Abus 从噬菌体粒子中释放出来。

上述方法的另一种选择是取出亲和基体，其已保留具强烈结合的

噬菌体或细菌粒子，并提取这些粒子的核酸，例如在 SDS 溶液中煮沸。提取出来的核酸可用来直接转化大肠杆菌宿主细胞或另一种选择是用合适的引物，通过 PCR 扩增抗体编码序列。

5 选择的效率可能取决于若干因子的共同作用，包括冲洗过程中解离动力学以及单个噬菌体或细菌中的多个 Abus 是否能同时结合到固相支持物上的抗原上。例如，具快速解离动力学的抗体(和弱的结合亲和力和力)通过短时间冲洗，多价展示以及固相支持物上抗原的高包被密度应被保留。相反，长时间的冲洗，单价噬菌体以及低包被密度的抗原
10 应有利于低解离动力学(和强的结合亲和力)Abus 的选择。

 如果需要，为了反选择不需要的 Abus，可以将与不相关抗原作用的全部 Abus 预选择。例如，为了分离抗独特性 Abus，全部 Abus 也得通过与相关抗原作用来进行预选择。

15 受试 Abu 库能快速分离具所需特异性的 Abus。正如所料，通过常规杂交瘤或转基因动物技术很难或不可能获得许多分离的 Abus。

包含本发明载体的试剂盒

20 本发明也包括以合适包装含有本发明载体的试剂盒。本发明的试剂盒包括那些能产生重构 Abus 的试剂，Abus 是通过本发明介绍的唯一异源二聚化序列对的配对亲和力重构的。

 每个试剂盒必须包含有可能使载体导入宿主细胞的试剂。有助于
25 载体转移的试剂的选择可以根据所用的特定的转染或感染方法而有所变化。试剂盒也包含用于产生标记多核苷酸探针或蛋白样探针的试剂，这些探针可检测 Abus。每种试剂可以是固体形式，也可以溶解/悬浮在液体缓冲液中，适合编目保存，以后如果进行实验的话，可以交换或加入反应介质中。提供了合适的包装。试剂盒能选择性地提供
30 别的在实验中用得着的成分。这些任选的成分包括但不限于缓冲液，

捕捉试剂，显影试剂，标记，反应表面，检测手段，对照样品，说明书以及解释信息。

5 根据本发明，Abus，多核苷酸，载体和宿主细胞的开发和使用在下面的实施例中进一步例证。提供的实施例只是作为本领域普通从业人员的指南，并不意味着只限于这种方法。

实施例

非单链抗原结合单位：卷曲螺旋 Fv(ccFv) 的构建

10 如上所介绍，Fv 片段是含完整抗原结合单位的最小抗体片段。Fv 片段包括重链和轻链的两个可变区(VH 和 VL)，它位于 Y 型免疫球蛋白分子的“顶部尖端”。Fv 片段的 VH 和 VL 片段之间的相互作用能量很低，而且在生理条件下应用时经常太不稳定。在天然存在的免疫球蛋白(如 Ig)中，位于恒定区 CH1 和 CL 的链间二硫键经常用来连接
15 VH 和 VL。这种连接使得稳定的抗原结合单位 Fab 的分子量能达到 50kDa。VH 和 VL 片段也可通过一条短的连接肽人为连接成单链 Fv 抗体片段(scFv)，连接肽位于一个片段的羧基端与另一个片段的氨基端之间。ScFv 抗原结合单位仅是 Fab 大小的一半。然而，一些 scFv 蛋白也是不稳定的。scFv 中的连接肽在某些情况下不会干扰结合。链
20 间二硫键也可导入 VH 和 VL 的框架区形成二硫键稳定的 Fv(dsFv)。dsFv 构型也有许多深奥的限制。两个 Cys 残基导入抗原结合可变区可能改变 VH 或 VL 的链内二硫键，从而干扰抗原结合。

25 我们设计了一种新的方法去稳定 VH 和 VL 异源二聚体。我们设计和使用了唯一的异源二聚化序列对来创造一种类 Fab 的，功能人为的 Fv 片段，即卷曲螺旋 Fv 片段(ccFv)。这一对异源二聚化序列衍生自异源二聚受体 GABA_B 受体 1 和 2。这对序列形成一种卷曲螺旋的结构并介导 GABA_B-R1 和 GABA_B-R2 受体的功能性异源二聚化。

30 和以前鉴定的来自 Fos 和 Jun 蛋白的卷曲螺旋亮氨酸拉链不同的

是, GABA_B-R1 和 GABA_B-R2 受体的 C 末端卷曲螺旋在生理条件下(如在体内)不形成可检测到的同源二聚体, 在生理体温下也不能形成同源二聚体。Kuner 等人和 White 等人(科学(1999), 283: 74-77; 自然(1998)396: 679-682)证明了 GABA_B-R1 和 GABA_B-R2 在体内具异源二聚化特异性。事实上, White 等人根据这对异源二聚受体的独有的特异性, 能够从酵母细胞中克隆 GABA_B-R2。Kammerer 等人, 同上的体外研究证明, GABA_B-R1 和 GABA_B-R2 C 末端序列当在生理体温下试验时, 在生理缓冲条件下均不能形成同源二聚体(参看 Kammerer 的表 1)。然而, 参与初始分离 GABA_B-R2 基因和鉴定卷曲螺旋的这些研究者没有一个介绍或甚至于暗示使用这种唯一的异源二聚化序列去构建异源多聚体如抗原结合单位。

我们修饰 GR1 和 GR2 结构域的羧基端, 是通过在 GR1 和 GR2 结构域的氨末端添加一个 flexon “SerArgGlyGlyGlyGly”, 从而使 V 区有额外的柔韧性。为了进一步稳定 ccFv, 我们在卷曲螺旋的 C 末端添加了 ValGlyGlyGly 间隔序列, 从而导入一对半胱氨酸残基。GR1 和 GR2 结构域分别融合到 VH 和 VL 片段的羧基端。VH-GR1 和 VL-GR2 融合序列在大肠杆菌中表达, 并由噬菌体展示。如图 10-11 所示, 获得了由平行卷曲螺旋 α 螺旋稳定的功能性异源二聚体 ccFv Abus。由于卷曲螺旋异源二聚化序列是 CH1 和 CL 结构域大小的一半左右, 因此, ccFv(大约 35kDa)比常规的 Fab 片段(大约是 50kDa)要小。ccFv 及其衍生物因为尺寸小, 因此在临床应用如肿瘤和组织侵入上更有潜力。ccFv 更有效的表达和展示是预料之中的。此外, 因为唯一异源二聚化序列的配对亲和力, VH 和 VL 区的特异装配使得强大的, 数目众多的 Abus 库的构建更可行。

材料和方法

细菌和噬菌体菌株: 用大肠杆菌 TG1(supE Δ (hsdM-mcrB) 5 (r_k⁻mk⁻McrB⁻)thi Δ (lac-proAB/F' traD36, lacI^q Δ (lacZ)M15))来获得质粒 DNA 和噬菌体; KO7 辅助噬菌体和 HRP 交联抗 M13 抗体来自 Phamersham

Pharmacia Biotech; pbluescript SK(+)来自 Stratagene; 抗 HA 抗体来自 Santa Cruz Biotechnology。

实施例 1: 载体构建

5 pABMX1 和 pABMX2 载体:

噬菌粒展示载体 pABMX1 和 pABMX2 衍生自 pbluescriptSK(+).
用 一 对 引 物 (pBS -Ska : 5'GGAATTGTGAGCGGATAACAAT
TTACCGGTCACACAGGAAACAGCTATGACCATG-3' 和 pBS-SKb :
5'CATGGTCATAGCTGTTTCCTGTGTGACCGGTAAATTGTTATCCG
10 CTCACAATTCC-3'), 通过基于 PCR 定点突变, 紧接在 Lac 启动子之
后导入一个唯一的 Age I 限制性酶切位点, 并通过剪切和钝末段连接
去掉 Xho I 和 Kpn I 位点。其后, 合成的 DNA 片段的 5'端侧翼是 Age
I 位点, 而 3'端侧翼是 Bgl II/EcoR I 位点, 含有来自 T7 噬菌体基因
10(TTAAC TT TA) 的 翻 译 增 强 序 列 , 核 糖 体 结 合 序 列
S/D(TAAGGAGG), 带有 HindIII位点的噬菌体基因 8 前导序列(对
15 pABMX2 而言是 ATGAAAAAGTCTTTAGTCCTCAAAGCCTCCGT
AGCCGTTGCTCCCTCGTTCCGATGCTAAGCTTCGCT)或 pelB 前导
序 列 (对 pABMX2 而 言 是 ATGAAATACCTATTGCCTACG
GCAGCCGCTGGATTGTTATTACTCGCGGCCAGCCGGCCATGGCG)
20 以 及 HA- (His)₆ 标 志 (DH 标 志)(TATCCATACGACGT
ACCAGACTACGCAGGAGGTCATCACCATCATCACCATTAG), 将这
段 DNA 克隆进修饰的 pbluescriptSK(+). 产生的载体命名为 pABMX1
和 pABMX2(参看图 3A-B 限制性图谱和序列)。把编码异源多聚体如
Nsc Abus 的异源序列进一步亚克隆进这些载体, 进行外周表达。

25

PABMD1 和 pABMD2 载体:

侧翼是 Bgl II 和 Sal I 位点的 PCR 扩增的 fd 基因III(或基因 3)被
插入 pABMX1 和 pABMX2 载体(参看图 4)。要展示的异源序列可以插
入在前导序列之后。Lac 启动子驱动 PIII衣壳融合序列的表达, PIII衣
壳融合序列又通过辅助噬菌体如 KO7 的超感染后展示在噬菌体表面。
30

PABMX5 和 pABMX6 载体:

这两个载体衍生自 pABMX1 和 pABMX2。5'端侧翼是 Xba I /Asc I 位点, 而 3'端侧翼是 Mlu I /Xho I /Not I 位点的合成 DNA 片段含有核糖体结合序列 S/D(TAAGGAGG) 和基因 3 前导序列 (ATGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTTCCTTTCTATTCTCACTCCGCT), 它被插入 pABMX1 和 pABMX2 的 Xba I /Not I 位点之间。然后, 将 GR1 区编码序列(图 2)亚克隆入 Xba I /Asc I 位点, 而 GR2 区编码序列(图 2)亚克隆入 Xho I /Not I 位点。接着 VH 和 VL 区分别插入 GR1 和 GR2 序列之前。载体 pABMX5 和 pABMX6 的示意图如图 5A 所示。这些载体在一个 Lac 启动子下表达两种蛋白: VH-GR1 和 VL-GR2。

PABMD5 和 pABMD6 载体:

来自载体 pABMX5 和 pABMX6 的 ccFv 片段亚克隆到 pABMD1 和 pABMD2, 得到载体 pABMD5 和 pABMD6(参看图 6A-B 的限制性图谱和序列)。这些载体表达两种蛋白: VH-GR1 和 VL-GR2-pIII 融合蛋白。表达的 VH-GR1 和 VL-GR2-pIII 融合蛋白被分泌到周质空间, 在周质空间通过卷曲螺旋区的异源二聚化发生二聚化。通过辅助噬菌体如 KO7 的超感染将装配的 Abu 展示在噬菌体表面。

实施例 2: 功能性 ccFv 的表达

为了表达 ccFv 片段, 将抗体 AM1 的可变区亚克隆到载体 pABMX6 中。然后把载体导入 TG1 细胞或 BL21 细胞。转化的细菌在 37°C, 500mL 含大约 100µg/mL 羧苄青霉素和 0.1%葡萄糖的 2XYT 中从一个单菌落长到 OD₆₀₀=0.7(左右)。加入 1mM IPTG, 在 30°C 诱导 4 小时。收集细菌沉淀, 准备外周休克和渗透休克。沉淀重悬在 12.5mL 加有 1.25mL 蛋白酶抑制剂混合物(sigma)的 PPB 缓冲液中(200mg/mL 蔗糖, 1mM EDTA, 30mM Tris-HCl, pH8.0), 接着把它们置冰上 20 分钟。通过离心收集上清。沉淀重悬在 5mM MgSO₄ 中, 并在冰上培养 20 分钟。将 MgSO₄

和 PPB 上清混合,并用 PBS 透析。上样到 1mL Ni-NTA 柱后,用 350mM 咪唑洗脱,纯化 His 标志的蛋白。图 10 所示的是纯化的 ccFv 在非还原性凝胶上有 35kDa 的电泳迁移率。在还原条件下分析时,观察到了两个对应于 VL 和 VH 的亚基。通过 Western 印迹证实,上面那条带是 VL-His 标志融合蛋白。

为了检测可溶性 AM1-ccFv 的结合特异性,进行了 ELISA 试验。AM1 抗原(0.2-1 μ g/孔)包被在 ELISA 平板上 4 $^{\circ}$ C 过夜。用 5%牛奶/PBS 封闭后,5%牛奶/PBS 的抗体溶液加入 ELISA 平板,在室温下培养 1-2 小时。冲洗掉没有结合的 Abus。图 10B 所示的是 AM1-ccFv 和它的抗原的特异性结合。对照含有 5%牛奶的 PBS。这个结果证实了功能性 ccFv 的装配是通过卷曲螺旋 GABA_BR1/R2 异源二聚化序列。

实施例 3: 功能性 ccFv 的展示

利用遗传包装展示抗体是从大文库中富集和分离特异的 Abu 的一种有力工具。为了分析 ccFv 在噬菌体展示系统中是否能被利用,我们通过把 AM1 抗体的 ccFv 基因亚克隆到 pABMD6 载体中,构建了一个噬菌粒载体。用辅助噬菌体 KO7 超感染携带有噬菌粒载体的 TG1 细胞。感染的 TG1 细胞在 2XYT/Amp/Kan 中 30 $^{\circ}$ C 过夜生长。用 PEG/NaCl 从培养上清中两次沉淀噬菌粒粒子,并重悬在 PBS 中。通过噬菌体 ELISA 试验,根据抗原结合活性检测展示在噬菌体上的抗体。简而言之,首先抗原包被在 ELISA 平板上。用 5%牛奶/PBS 封闭后,将噬菌体溶液加入 ELISA 平板中。通过和 HRP 交联的抗 M13 抗体培养,检测结合到抗原上的噬菌体。底物 ABTS[2,2'-连氮-双-(3-乙基苯并噻唑啉-6-磺酸)]被用来检测 HRP 活性。抗 HA 标志的抗体也用来检测展示在噬菌体粒子上的蛋白。抗 HA 的抗体也被包被在 96 孔的平板上(2 μ g/孔)。与包被在 ELISA 平板上的抗 HA 抗体结合的噬菌体可以用 HRP 交联的抗 M13 来检测。

为了比较 ccFv 和 scFv 噬菌体展示,也制备了单链抗体噬菌体。

如图 11A-B 所示, ccFv 噬菌体的结合能力和常规的 scFv 噬菌体具有可比性。对于某些 ccFv 表达噬菌体, 它们的结合能力比表达常规 scFv 的噬菌体几乎要高一个等级(图 11B)。因此, ccFv 即便展示在噬菌体粒子上也是一个功能性 Abu。

5

单链抗原结合单位的表达:

实施例 4: 常规 scFv 的表达

AM2-scFv 被亚克隆入可溶性表达载体 pABMX1 的 HindIII/Not I 位点。如上概括的那样进行外周准备。一个从 NI-NTA 柱上纯化出来的 30kDa 抗体蛋白通过 SDS-PAGE 获得证实。同时用 ELISA 对它的抗原结合特异性进行了检测。AM2 抗原首先以 0.2 μ g/孔的浓度包被在 ELISA 平板上。不同数量的 AM2-scFv 片段和抗原培养。用抗 HA 标志抗体检测结合的 AM2-scFv 片段。这个实验表明 AM2-scFv 和它的 AM2 抗原的结合是剂量依赖性的(图 8)。

15

实施例 5: 常规 scFv 在噬菌体上的展示

AM2-scFv 片段首先被亚克隆进噬菌粒载体 pABMD1 的 HindIII/Not I 位点。携带该噬菌粒载体的 TG1 细胞被辅助噬菌体 KO7 超感染。从上清中纯化噬菌体。随后进行噬菌体 ELISA 实验去检测展示在噬菌体粒子上的 AM2-scFv。因为外壳 PIII基因用 HA-标志标记上了, 因此可以用抗 HA 抗体来检测融合蛋白。使用 AM2 抗原和抗 HA 抗体的 ELISA 实验证实, 展示的 scFv 能特异性地和相应抗原结合(图 9)。对照中噬菌体展示的是没有 HA 标记的无关抗体。

20

受试抗原结合单位在真核细胞中的表达:

实施例 6: ccFv 在酵母中的表达

构建了酵母载体 pABMEX7, 它携带有连接在受试异源二聚化序列的 VL 和 VH 序列。根据本领域已知的任何方法制备感受态的酵母细胞如 AH109 细胞, 并用 pABMEX7 载体转化。转化的酵母细胞在适合蛋白表达的情况下培养。这些条件是本领域的技术人员众所周知

30

的，因此没必要在这详细介绍。应用本领域常规方法和/或本发明中介绍的方法收获表达的 ccFv Abus。根据上面介绍的方法用 ELISA 检测收获的 ccFv 的抗原结合力。

5 具所需结合特异性的 ccFv 噬菌体的富集和鉴定：

实施例 7：来自模式文库的 AM2-ccFv 噬菌体的富集

为了证明具所需结合特异性的 ccFv 展示噬菌体能从背景噬菌体中选择出来并最终被富集，我们从模式文库中淘选 AM2-ccFv 噬菌体。把 AM2-ccFv 噬菌体和不相关的 AM1-ccFv 展示噬菌体以 $1:10^6$ 或 $1:10^7$ 10 比率混合，制备模式文库。对包被在 Nunc Maxisorb 96 孔平板上的特异蛋白抗原的文库进行了两轮淘选，用含 5%牛奶的磷酸缓冲液封闭后，把含 1×10^{12} 文库的 2%牛奶/PBS 加入孔内，在室温下培养 2 小时。然后去掉噬菌体溶液，并用 PBST(0.05%TWEEN-20 的 PBS)和 PBS 分别洗 5 次。用 100mM 三乙胺洗脱结合的噬菌体，并将它加入 TG1 培养物中进行感染。从感染的 TG1 细胞中制备的噬菌体用于下一轮淘选 15 以及针对 AM2-ccFv 的固定化蛋白抗原进行 ELISA。每轮淘选后，通过对随机挑取的克隆进行 PCR 分析来检测 AM2-ccFv 噬菌体和背景噬菌体 AM1-ccFv 噬菌体的比率，PCR 的一对引物是只为 AM2-ccFv 特异性设计的(一条~1kb PCR 产物在 1%琼脂糖凝胶上可见)，而不是为 20 AM1-ccFv 设计的(没有 PCR 产物)。如图 20 所示，对 AM2-ccFv 固定化蛋白抗原进行了 ELISA 实验。和第一轮相比，第 2 轮淘选中的噬菌体产生的 OD405 读数更高，表明 AM2-ccFv 噬菌体在淘选过程中被成功地挑选和富集。图 21 PCR 分析所示的是随机挑取克隆的 PCR 产物的 1%琼脂糖胶的影像。结果表明经过第一轮淘选后 AM2-ccFv 的出现 25 率在 $1:10^7$ 文库中是 4.4%，而经过第二轮淘选后出现率则达到 100%。

实施例 8：用展示功能性 ccFv 的噬菌体文库进行文库筛选

噬菌体文库中有数目众多的 Abus，所有的 Abus 都能结合所需的抗原，但每个 Abu 在关键结合区的某一部位有差异，对文库进行筛选 30 使我们能鉴定和进一步设计具更高亲和力的抗体。为了证明 ccFv 能

被有效地用作筛选抗体文库的媒介物，我们以 ccFv 的形式构建了一个 AM2 文库。AM2 的框架结构由 VH 和 VL 区组成，每一个区均被亚克隆入上述的载体 pABMD6 的相应位置上。简而言之，VH 连接到 GR1 的 N 末端，而 VL 连接到 GR2 的 N 末端。GR2 又和 PIII 融合是为了在噬菌体粒子表面展示 ccFv 抗体片段(参看实施例 1, 3 和 7)。为了通过 ccFv 噬菌体文库淘选筛选噬菌体高亲和力结合剂，把 VH 的 CDR3 构建成一个文库，它在多个位点含有多种选择的残基。通过对变性的寡聚 DNA 进行标准合成，制备文库，寡聚 DNA 侧翼两端有限制性酶切位点，使得它在 PCR 扩增和限制性消化后能亚克隆入 AM2 位置。图 22 所示的是变性寡聚 DNA 的设计，扩增引物和方向。不同的文库被设计成总共大约 10^6 种。连接以后，文库被电穿孔入 TG1 感受态细胞。然后在如上文实施例所述的方法收集文库的噬菌体粒子之前，收获转化子，并用辅助噬菌体 KO7 补救。为了淘选文库，将重组抗原固定在 Nunc Maxisorb 96 孔平板上，在 0.05M, pH9.6 的 NaHCO_3 缓冲液中 4℃ 过夜。含 10^{12} 个噬菌体粒子的噬菌体文库等份试样在含 2% 牛奶的 PBS 缓冲液中稀释，然后加入包被有抗原的孔中。在 37℃ 培养 2 小时后，冲洗孔，并用实施例 3 和 7 所介绍的方法冲洗噬菌体。洗脱的噬菌体被用来感染 TG1 细胞，然后用 KO7 辅助噬菌体补救。在 30℃ 过夜生长后，噬菌体在 TG1 细胞中扩增并准备收获以用于第二轮淘选。以前的研究表明，解离速率 K_{off} 最显著地有助于这种特定抗体的亲和力，而 K_{on} 相对恒定。因此，为了筛选具高结合亲和力的噬菌体，设计了一种延长冲洗时间来筛选存活噬菌体的方法，慢速淘选(off-rate panning)，从而特异地筛选出那些具更慢 K_{off} 的噬菌体。对这个文库进行了总共连续 7 轮淘选。

在噬菌体淘选过程中，经过第 5 和第 7 轮淘选后，随机挑取单克隆，利用同一蛋白抗原对制备的噬菌体进行 ELISA 试验。用 96 孔微量滴定板根据上文实施例 7 所介绍的方法进行 ELISA。如图 23 所示，在第 5 和第 7 轮淘选中选择的所有克隆均能和抗原进行阳性反应。记住阴性对照放在第 5 轮的位置 H10, 11, 12 和第 7 轮的 H11, 12, 所有

的阴性对照低于 0.05 OD₄₀₅。

- 5 在 ELISA 第 7 轮淘选中表示出阳性反应活性的克隆被随机挑取测序，并对 VH 的 CD3 上的残基的图谱见图 24。靶位点的共有序列通过文库进行鉴定。对代表共有序列和野生型(原始的 AM2)的克隆进行蛋白表达，随后用 BiaCore(表面胞质基因)确定这些蛋白和固定化蛋白抗原结合的 K_{off}。如图 25 所示，选自文库的 AM2 变体有着显著低的 K_{off}。

序列表

<110> 埃比玛克斯公司 (ABMAXIS, INC.)

<120> 产生嵌合异源多聚体的组合物和方法

(COMPOSITIONS AND METHODS FOR GENERATING CHIMERIC HETEROMULTIMERS)

<130> SCT040348-47

<140>

<141>

<160> 30

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 146

<212> DNA

<213> 智人 (Homo sapiens)

<220>

<221> CDS

<222> (1)...(132)

<400> 1

tct	aga	ggt	gga	gga	ggt	gag	gag	aag	tcc	cgg	ctg	ttg	gag	aag	gag	48
Ser	Arg	Gly	Gly	Gly	Gly	Glu	Glu	Lys	Ser	Arg	Leu	Leu	Glu	Lys	Glu	
1				5				10					15			

aac	cgt	gaa	ctg	gaa	aag	atc	att	gct	gag	aaa	gag	gag	cgt	gtc	tct	96
Asn	Arg	Glu	Leu	Glu	Lys	Ile	Ile	Ala	Glu	Lys	Glu	Glu	Arg	Val	Ser	
			20					25					30			

gaa	ctg	cgc	cat	caa	ctc	cag	tct	gta	gga	ggt	tgt	taatagggcg	142
Glu	Leu	Arg	His	Gln	Leu	Gln	Ser	Val	Gly	Gly	Cys		
			35					40					

cgcc	146
------	-----

<210> 2

<211> 44

<212> PRT

<213> 智人 (Homo sapiens)

<400> 2

Ser	Arg	Gly	Gly	Gly	Gly	Glu	Glu	Lys	Ser	Arg	Leu	Leu	Glu	Lys	Glu
1				5				10					15		
Asn	Arg	Glu	Leu	Glu	Lys	Ile	Ile	Ala	Glu	Lys	Glu	Glu	Arg	Val	Ser
			20					25					30		
Glu	Leu	Arg	His	Gln	Leu	Gln	Ser	Val	Gly	Gly	Cys				
			35					40							

<210> 3

<211> 140

<212> DNA

<213> 智人 (Homo sapiens)

<220>

<221> CDS

<222> (1)...(140)

<400> 3

tct	cga	gga	ggt	ggt	gga	aca	tcc	cgc	ctg	gag	ggc	cta	cag	tca	gaa	48
Ser	Arg	Gly	Gly	Gly	Gly	Thr	Ser	Arg	Leu	Glu	Gly	Leu	Gln	Ser	Glu	
1				5					10					15		

aac	cat	cgc	ctg	cga	atg	aag	atc	aca	gag	ctg	gat	aaa	gac	ttg	gaa	96
Asn	His	Arg	Leu	Arg	Met	Lys	Ile	Thr	Glu	Leu	Asp	Lys	Asp	Leu	Glu	
			20					25					30			

gag	gtc	acc	atg	cag	ctg	cag	gac	gtc	gga	ggt	tgc	gcg	gcc	gc		140
Glu	Val	Thr	Met	Gln	Leu	Gln	Asp	Val	Gly	Gly	Cys	Ala	Ala			
		35					40					45				

<210> 4

<211> 46

<212> PRT

<213> 智人 (Homo sapiens)

<400> 4

Ser	Arg	Gly	Gly	Gly	Gly	Thr	Ser	Arg	Leu	Glu	Gly	Leu	Gln	Ser	Glu	
1				5					10					15		
His	Arg	Leu	Arg	Met	Lys	Ile	Thr	Glu	Leu	Asp	Lys	Asp	Leu	Glu	Glu	
			20					25					30			
Val	Thr	Met	Gln	Leu	Gln	Asp	Val	Gly	Gly	Cys	Ala	Ala	Ala			
		35					40					45				

<210> 5

<211> 203

<212> DNA

<213> 人工序列

<220>

<223> Bluescript 载体

<400> 5

aat	tgt	gag	cgg	ata	aca	att	tac	cgg	ttc	ttt	taa	ctt	tag	taa	gga	48
gga	att	aaa	aaa	tga	aaa	agt	ctt	tag	tcc	tca	aag	cct	ccg	tag	ccg	96
ttg	cta	ccc	tcg	ttc	cga	tgc	taa	gct	tcg	ctt	cta	gag	cgg	ccg	ctt	144
atc	cat	acg	acg	tac	cag	act	acg	cag	gag	gtc	atc	acc	atc	atc	acc	192
att	aga	gat	ct													203

<210> 6

<211> 45

<212> PRT

<213> 人工序列

<220>

<223> Bluescript

<400> 6

Met	Lys	Lys	Ser	Leu	Val	Leu	Lys	Ala	Ser	Val	Ala	Val	Ala	Thr	Leu
1				5				10					15		
Val	Pro	Met	Leu	Ser	Phe	Ala	Ser	Arg	Ala	Ala	Ala	Tyr	Pro	Tyr	Asp
			20					25					30		
Val	Pro	Asp	Tyr	Ala	Gly	Gly	His	His	His	His	His	His			
		35					40					45			

<210> 7

<211> 212

<212> DNA

<213> 人工序列

<220>

<223> Bluescript 载体

<400> 7

aat	tgt	gag	cgg	ata	aca	att	tac	cgg	ttc	ttt	taa	ctt	tag	taa	gga	48
gga	att	aaa	aaa	tga	aat	acc	tat	tgc	cta	cgg	cag	ccg	ctg	gat	tgt	96
tat	tac	tcg	cgg	ccc	agc	cgg	cca	tgg	cgg	ccc	tgc	agg	cct	cta	gag	144
cgg	ccg	ctt	atc	cat	acg	acg	tac	cag	act	acg	cag	gag	gtc	atc	acc	192
atc	atc	acc	att	aga	gat	ct										212

<210> 8

<211> 48

<212> PRT

<213> 人工序列

<220>

<223> Bluescript

<400> 8

Met	Lys	Tyr	Leu	Leu	Pro	Thr	Ala	Ala	Ala	Gly	Leu	Leu	Leu	Leu	Ala
1				5				10					15		
Ala	Gln	Pro	Ala	Met	Ala	Ala	Leu	Gln	Ala	Ser	Arg	Ala	Ala	Ala	Tyr
			20					25					30		
Pro	Tyr	Asp	Val	Pro	Asp	Tyr	Ala	Gly	Gly	His	His	His	His	His	His
		35					40					45			

<210> 9

<211> 272

<212> DNA

<213> 人工序列

<220>

<223> Bluescript 载体

<400> 9

aat	tgt	gag	cgg	ata	aca	att	tac	cgg	ttc	ttt	taa	ctt	tag	taa	gga	48
gga	att	aaa	aaa	tga	aaa	agt	ctt	tag	tcc	tca	aag	cct	ccg	tag	ccg	96
ttg	cta	ccc	tcg	ttc	cga	tgc	taa	gct	tcg	ctt	cta	gag	cgg	ccg	ctt	144
atc	cat	acg	acg	tac	cag	act	acg	cag	gag	gtc	atc	acc	atc	atc	acc	192
att	aga	gat	ctg	gag	gcg	gta	ctg	ttg	aaa	gtt	gtt	tag	caa	aa		236
g	cta	aca	tac	tgc	gta	ata	agg	agt	ctt	aag	tcg	ac				272

<210> 10
 <211> 67
 <212> PRT
 <213> 人工序列

<220>
 <223> BlueScript 载体

<400> 10
 Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
 1 5 10 15
 Val Pro Met Leu Ser Phe Ala Ser Arg Ala Ala Ala Tyr Pro Tyr Asp
 20 25 30
 Val Pro Asp Tyr Ala Gly Gly His His His His His His Arg Ser Gly
 35 40 45
 Gly Gly Thr Val Glu Ser Cys Leu Ala Lys Ala Asn Ile Leu Arg Asn
 50 55 60
 Lys Glu Ser
 65

<210> 11
 <211> 281
 <212> DNA
 <213> 人工序列

<220>
 <223> Bluescript 载体

<400> 11
 aat tgt gag cgg ata aca att tac cgg ttc ttt taa ctt tag taa gga 48
 gga att aaa aaa tga aat acc tat tgc cta cgg cag ccg ctg gat tgt 96
 tat tac tcg cgg ccc agc cgg cca tgg cgg ccc tgc agg cct cta gag 144
 cgg ccg ctt atc cat acg acg tac cag act acg cag gag gtc atc acc 192
 atc atc acc att aga gat ctg gag gcg gta ctg ttg aaa gtt gtt tag 240
 caa a ag cta aca tac tgc gta ata agg agt ctt aag tcg ac 281

<210> 12
 <211> 70
 <212> PRT
 <213> 人工序列

<220>
 <223> Bluescript 载体

<400> 12
 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1 5 10 15
 Ala Gln Pro Ala Met Ala Ala Leu Gln Ala Ser Arg Ala Ala Tyr
 20 25 30
 Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly His His His His His
 35 40 45
 Arg Ser Gly Gly Gly Thr Val Glu Ser Cys Leu Ala Lys Ala Asn Ile
 50 55 60
 Leu Arg Asn Lys Glu Ser
 65 70

```

<210> 13
<211> 501
<212> DNA
<213> 人工序列

<220>
<221> CDS
<222> (1)...(501)
<223> Bluescript 载体

<400> 13
atg aaa aag tct tta gtc ctc aaa gcc tcc gta gcc gtt gct acc ctc      48
Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
  1             5             10             15

gtt ccg atg cta agc ttc gct tct aga ggt gga gga ggt gag gag aag      96
Val Pro Met Leu Ser Phe Ala Ser Arg Gly Gly Gly Gly Glu Glu Lys
             20             25             30

tcc cgg ctg ttg gag aag gag aac cgt gaa ctg gaa aag atc att gct      144
Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala
             35             40             45

gag aaa gag gag cgt gtc tct gaa ctg cgc cat caa ctc cag tct gta      192
Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val
             50             55             60

gga ggt tgt taa tag ggc gcg cca caa ttt cac agt aag gag gtt taa      240
Gly Gly Cys * * Gly Ala Pro Gln Phe His Ser Lys Glu Val *
             65             70             75

ctt atg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat      288
Leu Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr
             80             85             90

tct cac tcc gct acg cgt tct cga gga ggt ggt gga aca tcc cgc ctg      336
Ser His Ser Ala Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu
             95             100             105

gag ggc cta cag tca gaa aac cat cgc ctg cga atg aag atc aca gag      384
Glu Gly Leu Gln Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu
             110             115             120             125

ctg gat aaa gac ttg gaa gag gtc acc atg cag ctg cag gac gtc gga      432
Leu Asp Lys Asp Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly
             130             135             140

ggt tgc gcg gcc gct tat cca tac gac gta cca gac tac gca gga ggt      480
Gly Cys Ala Ala Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly
             145             150             155

cat cac cat cat cac cat tag      501
His His His His His His *
             160

```

<210> 14
 <211> 163
 <212> PRT
 <213> 人工序列

<220>
 <223> Bluescript 载体

<400> 14
 Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
 1 5 10 15
 Val Pro Met Leu Ser Phe Ala Ser Arg Gly Gly Gly Gly Glu Glu Lys
 20 25 30
 Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala
 35 40 45
 Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val
 50 55 60
 Gly Gly Cys Gly Ala Pro Gln Phe His Ser Lys Glu Val Leu Met Lys
 65 70 75 80
 Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser His Ser
 85 90 95
 Ala Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu Glu Gly Leu
 100 105 110
 Gln Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu Leu Asp Lys
 115 120 125
 Asp Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly Gly Cys Ala
 130 135 140
 Ala Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly His His His
 145 150 155 160
 His His His

<210> 15
 <211> 498
 <212> DNA
 <213> 人工序列

<220>
 <221> CDS
 <222> (1)...(498)
 <223> Bluescript 载体

<400> 15
 atg aaa tac cta ttg cct acg gca gcc gct gga ttg tta tta ctc gcg 48
 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Ala
 1 5 10 15
 gcc cag ccg gcc atg gcg tct aga ggt gga gga ggt gag gag aag tcc 96
 Ala Gln Pro Ala Met Ala Ser Arg Gly Gly Gly Glu Glu Lys Ser
 20 25 30
 cgg ctg ttg gag aag gag aac cgt gaa ctg gaa aag atc att gct gag 144
 Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu
 35 40 45

```

aaa gag gag cgt gtc tct gaa ctg cgc cat caa ctc cag tct gta gga      192
Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val Gly
   50                      55                      60

ggt tgt taa tag ggc gcg cca caa ttt cac agt aag gag gtt taa ctt      240
Gly Cys * * Gly Ala Pro Gln Phe His Ser Lys Glu Val * Leu
   65                      70                      75

atg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat tct      288
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
   80                      85                      90

cac tcc gct acg cgt tct cga gga ggt ggt gga aca tcc cgc ctg gag      336
His Ser Ala Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu Glu
   95                      100                     105

ggc cta cag tca gaa aac cat cgc ctg cga atg aag atc aca gag ctg      384
Gly Leu Gln Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu Leu
  110                      115                      120                     125

gat aaa gac ttg gaa gag gtc acc atg cag ctg cag gac gtc gga ggt      432
Asp Lys Asp Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly Gly
                130                      135                      140

tgc gcg gcc gct tat cca tac gac gta cca gac tac gca gga ggt cat      480
Cys Ala Ala Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly His
                145                      150                     155

cac cat cat cac cat tag      498
His His His His His *
                160

```

<210> 16
 <211> 162
 <212> PRT
 <213> 人工序列

<220>
 <223> Bluescript 载体

```

<400> 16
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
 1          5          10          15
Ala Gln Pro Ala Met Ala Ser Arg Gly Gly Gly Gly Glu Glu Lys Ser
 20          25          30
Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu
 35          40          45
Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val Gly
 50          55          60
Gly Cys Gly Ala Pro Gln Phe His Ser Lys Glu Val Leu Met Lys Lys
 65          70          75          80
Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser His Ser Ala
 85          90          95
Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu Glu Gly Leu Gln
100          105          110

```

```

Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu Leu Asp Lys Asp
      115      120      125
Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly Gly Cys Ala Ala
      130      135      140
Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly His His His His
      145      150      155      160
His His

```

<210> 17

<211> 552

<212> DNA

<213> 人工序列

<220>

<221> CDS

<222> (1)...(543)

<223> Bluescript 载体

<400> 17

```

atg aaa aag tct tta gtc ctc aaa gcc tcc gta gcc gtt gct acc ctc      48
Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
  1              5              10              15

```

```

gtt ccg atg cta agc ttc gct tct aga ggt gga gga ggt gag gag aag      96
Val Pro Met Leu Ser Phe Ala Ser Arg Gly Gly Gly Gly Glu Glu Lys
              20              25              30

```

```

tcc cgg ctg ttg gag aag gag aac cgt gaa ctg gaa aag atc att gct      144
Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala
              35              40              45

```

```

gag aaa gag gag cgt gtc tct gaa ctg cgc cat caa ctc cag tct gta      192
Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val
              50              55              60

```

```

gga ggt tgt taa tag ggc gcg cca caa ttt cac agt aag gag gtt taa      240
Gly Gly Cys * * Gly Ala Pro Gln Phe His Ser Lys Glu Val *
              65              70              75

```

```

ctt atg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat      288
Leu Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr
              80              85              90

```

```

tct cac tcc gct acg cgt tct cga gga ggt ggt gga aca tcc cgc ctg      336
Ser His Ser Ala Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu
              95              100              105

```

```

gag ggc cta cag tca gaa aac cat cgc ctg cga atg aag atc aca gag      384
Glu Gly Leu Gln Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu
              110              115              120              125

```

```

ctg gat aaa gac ttg gaa gag gtc acc atg cag ctg cag gac gtc gga      432
Leu Asp Lys Asp Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly
              130              135              140

```

```

ggt tgc gcg gcc gct tat cca tac gac gta cca gac tac gca gga ggt      480
Gly Cys Ala Ala Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly
                145                150                155

cat cac cat cat cac cat tag gga ggc ggt act gtt gaa agt tgt ctg      528
His His His His His His * Gly Gly Gly Thr Val Glu Ser Cys Leu
                160                165                170

cgt aat aag gag tct taagtcgac      552
Arg Asn Lys Glu Ser
                175

```

<210> 18
 <211> 185
 <212> PRT
 <213> 人工序列

<220>
 <223> Bluescript 载体

```

<400> 18
Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
 1          5          10          15
Val Pro Met Leu Ser Phe Ala Ser Arg Gly Gly Gly Gly Glu Lys
 20          25          30
Ser Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala
 35          40          45
Glu Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val
 50          55          60
Gly Gly Cys Gly Ala Pro Gln Phe His Ser Lys Glu Val Leu Met Lys
 65          70          75          80
Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser His Ser
 85          90          95
Ala Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu Glu Gly Leu
100          105          110
Gln Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu Leu Asp Lys
115          120          125
Asp Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly Gly Cys Ala
130          135          140
Ala Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly His His His
145          150          155          160
His His His Arg Ser Gly Gly Gly Thr Val Glu Ser Cys Leu Ala Lys
165          170          175
Ala Asn Ile Leu Arg Asn Lys Glu Ser
180          185

```

<210> 19
 <211> 549
 <212> DNA
 <213> 人工序列

<220>
 <221> CDS
 <222> (1)...(540)
 <223> Bluescript 载体

```

<400> 19
atg aaa tac cta ttg cct acg gca gcc gct gga ttg tta tta ctc gcg      48
Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
  1             5             10             15

gcc cag ccg gcc atg gcg tct aga ggt gga gga ggt gag gag aag tcc      96
Ala Gln Pro Ala Met Ala Ser Arg Gly Gly Gly Gly Glu Glu Lys Ser
             20             25             30

cgg ctg ttg gag aag gag aac cgt gaa ctg gaa aag atc att gct gag      144
Arg Leu Leu Glu Lys Glu Asn Arg Glu Leu Glu Lys Ile Ile Ala Glu
             35             40             45

aaa gag gag cgt gtc tct gaa ctg cgc cat caa ctc cag tct gta gga      192
Lys Glu Glu Arg Val Ser Glu Leu Arg His Gln Leu Gln Ser Val Gly
             50             55             60

ggt tgt taa tag ggc gcg cca caa ttt cac agt aag gag gtt taa ctt      240
Gly Cys * * Gly Ala Pro Gln Phe His Ser Lys Glu Val * Leu
             65             70             75

atg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat tct      288
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
             80             85             90

cac tcc gct acg cgt tct cga gga ggt ggt gga aca tcc cgc ctg gag      336
His Ser Ala Thr Arg Ser Arg Gly Gly Gly Gly Thr Ser Arg Leu Glu
             95             100             105

ggc cta cag tca gaa aac cat cgc ctg cga atg aag atc aca gag ctg      384
Gly Leu Gln Ser Glu Asn His Arg Leu Arg Met Lys Ile Thr Glu Leu
             110             115             120             125

gat aaa gac ttg gaa gag gtc acc atg cag ctg cag gac gtc gga ggt      432
Asp Lys Asp Leu Glu Glu Val Thr Met Gln Leu Gln Asp Val Gly Gly
             130             135             140

tgc gcg gcc gct tat cca tac gac gta cca gac tac gca gga ggt cat      480
Cys Ala Ala Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Gly His
             145             150             155

cac cat cat cac cat tag gga ggc ggt act gtt gaa agt tgt ctg cgt      528
His His His His His * Gly Gly Gly Thr Val Glu Ser Cys Leu Arg
             160             165             170

aat aag gag tct taagtcgac      549
Asn Lys Glu Ser
             175

```

```

<210> 20
<211> 184
<212> PRT
<213> 人工序列

```

```

<220>

```

<223> Bluescript 载体

<400> 20

```

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

```

<210> 21

<211> 55

<212> DNA

<213> 人工序列

<220>

<223> 引物

<400> 21

ggaattgtga gcggataaca atttaccggt cacacaggaa acagctatga ccatg

55

<210> 22

<211> 55

<212> DNA

<213> 人工序列

<220>

<223> 引物

<400> 22

catggtcata gctgtttcct gtgtgaccgg taaattgtta tccgctcaca attcc

55

<210> 23

<211> 9

<212> DNA

<213> 病毒

<400> 23

ttaacttta	9
<210> 24	
<211> 8	
<212> DNA	
<213> 病毒	
<400> 24	
taaggagg	8
<210> 25	
<211> 68	
<212> DNA	
<213> 病毒	
<400> 25	
atgaaaaagt ctttagtcct caaagcctcc gtagccgttg ctccctcggt ccgatgctaa	60
gcttcgct	68
<210> 26	
<211> 66	
<212> DNA	
<213> 病毒	
<400> 26	
atgaaatacc tattgcctac ggcagccgct ggattgttat tactcgcggc ccagccggcc	60
atggcg	66
<210> 27	
<211> 54	
<212> DNA	
<213> 人工序列	
<220>	
<223> 噬菌体	
<400> 27	
tatccatacg acgtaccaga ctacgcagga ggatcatcacc atcatcacca ttag	54
<210> 28	
<211> 57	
<212> DNA	
<213> 人工序列	
<220>	
<223> 噬菌体	
<400> 28	
atgaaaaaat tattattcgc aattccttta gttgttcctt tctattctca ctccgct	57
<210> 29	
<211> 4	
<212> PRT	
<213> 人工序列	

<220>
<223> Bluescript

<400> 29
Val Gly Gly Cys
1

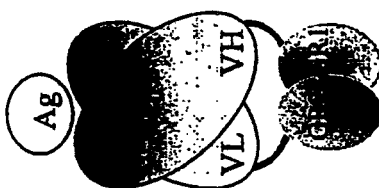
<210> 30
<211> 4
<212> PRT
<213> 人工序列

<220>
<223> Bluescript

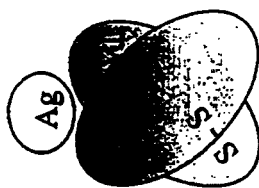
<400> 30
Gly Gly Gly Gly
1

图1

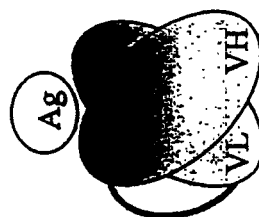
抗原结合单位



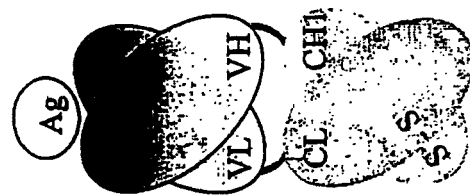
cFv
异源二聚体卷曲
螺旋稳定化的Fv



dsFv
二硫键稳定化
的Fv



scFv
单链Fv



Fab

Ag 抗原

图2

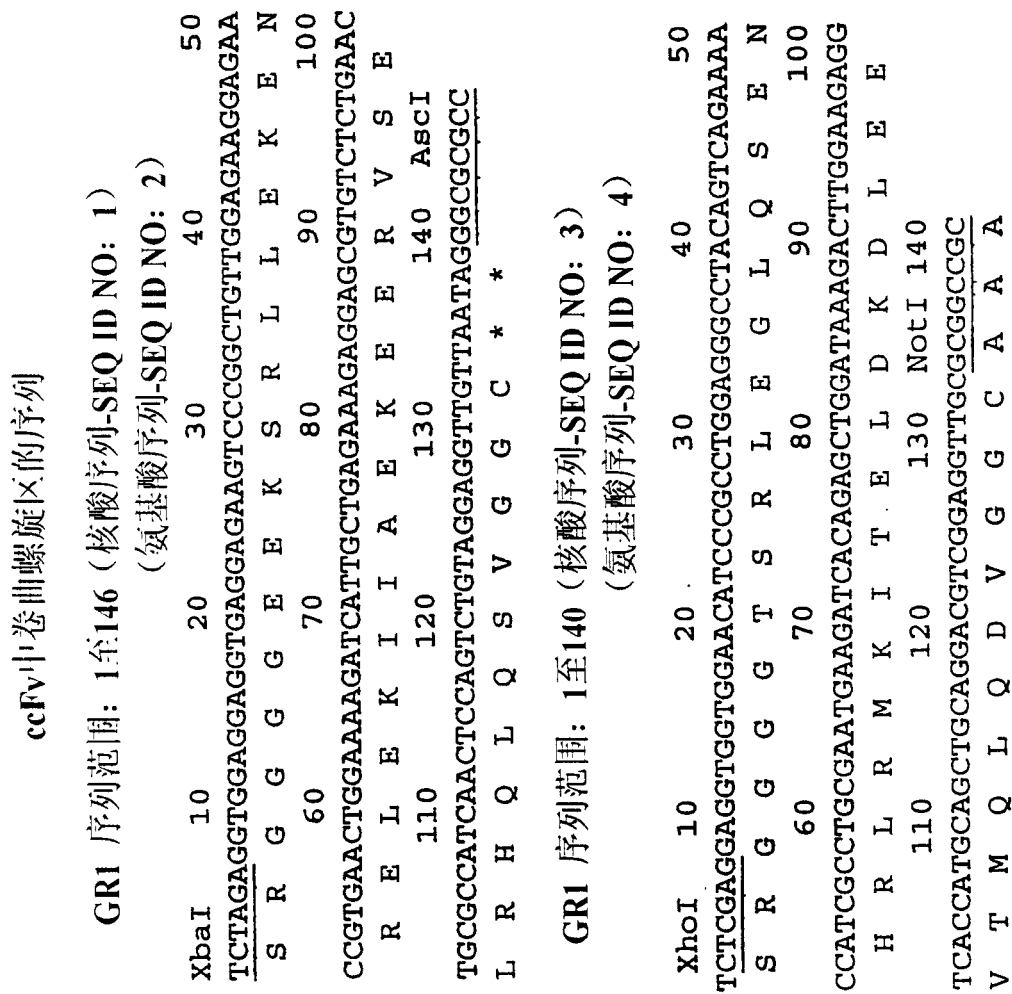


图3A

抗体表达的载体

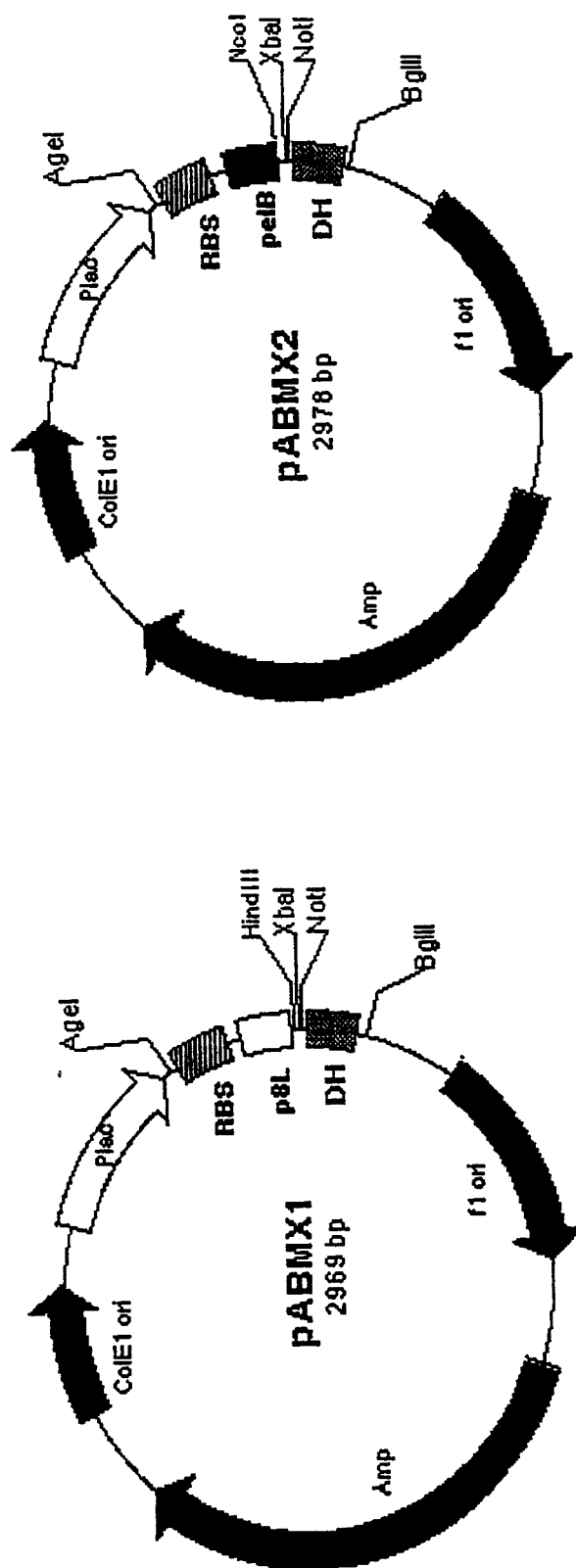


图3B

PABMX1 载体: AgeI至BglIII的序列 (核酸序列-SEQ ID NO: 5)

(氨基酸序列-SEQ ID NO: 6)

lac 启动子 /lac O1 AgeI EP S/D
AATTGTGAGCGGATAACAATT TACCGGT TCTT TTAACCTTAG TAAGGAGG AATTAAAAA
P8 前导序列 HindIII XbaI
ATGAAAAAGTCTTTAGTCCTCAAAAGCCTCCGTAGCCGTTGCTACCCCTCGTCCGATGCTAAGCTTCGCT TCTAGA
M K K S L V L K A S V A V A T L V P M L S F A S R
NotI HA-标志 His-标志 BglIII
GCGGCCGCT TATCCATACGACGTACCACTACGCA GGAGT CATCACCATCATCACCAT TAG AGATCT
A A A Y P Y D V P D Y A G G H H H H H * R S

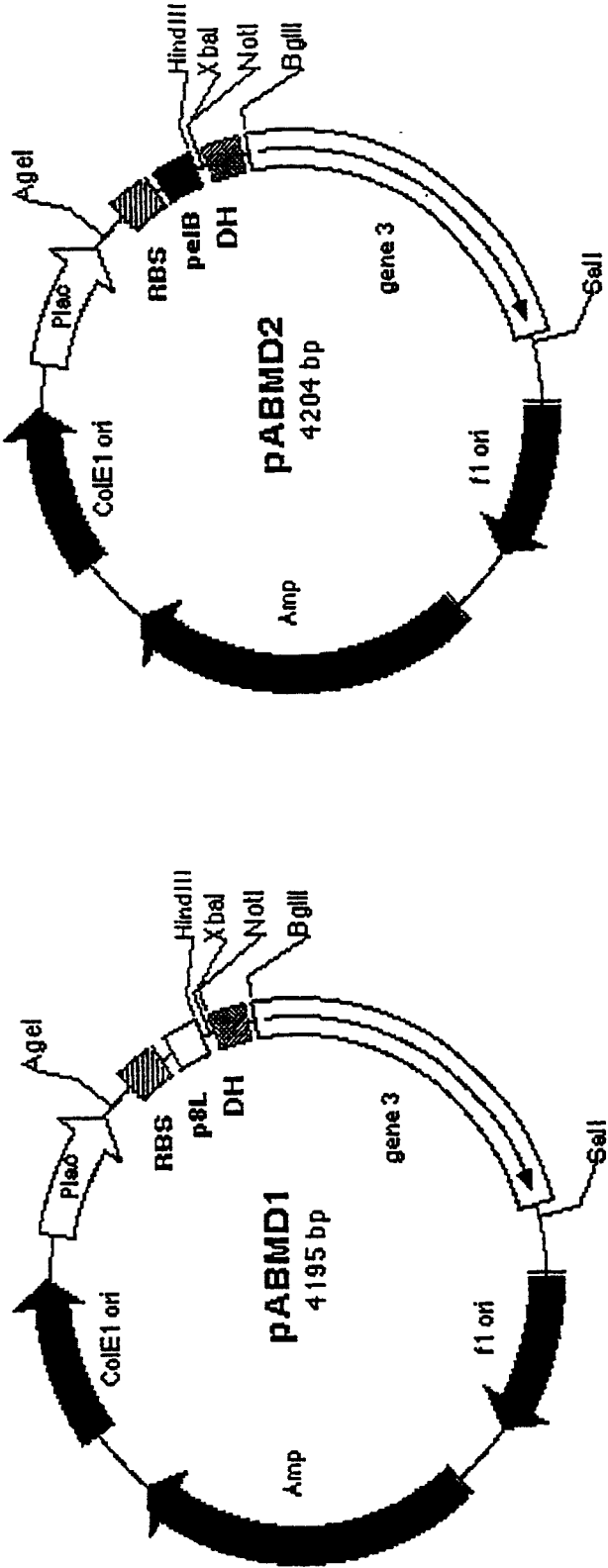
PABMX2 载体: AgeI至BglIII的序列 (核酸序列-SEQ ID NO: 7)

(氨基酸序列-SEQ ID NO: 8)

lac 启动子 /lac O1 AgeI EP S/D
AATTGTGAGCGGATAACAATT TACCGGT TCTT TTAACCTTAG TAAGGAGG AATTAAAAA
pelB 前导序列 Nco I PstI XbaI
ATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCAGCCGCCCTGCAGGGCCTCTAGA
M K Y L L P T A A A G L L L A A Q P A M A A L Q A S R
NotI HA-标志 His-标志 BglIII
GCGGCCGCT TATCCATACGACGTACCACTACGCA GGAGT CATCACCATCATCACCAT TAG AGATCT
A A A Y P Y D V P D Y A G G H H H H H H * R S

图4A

抗体展示的载体

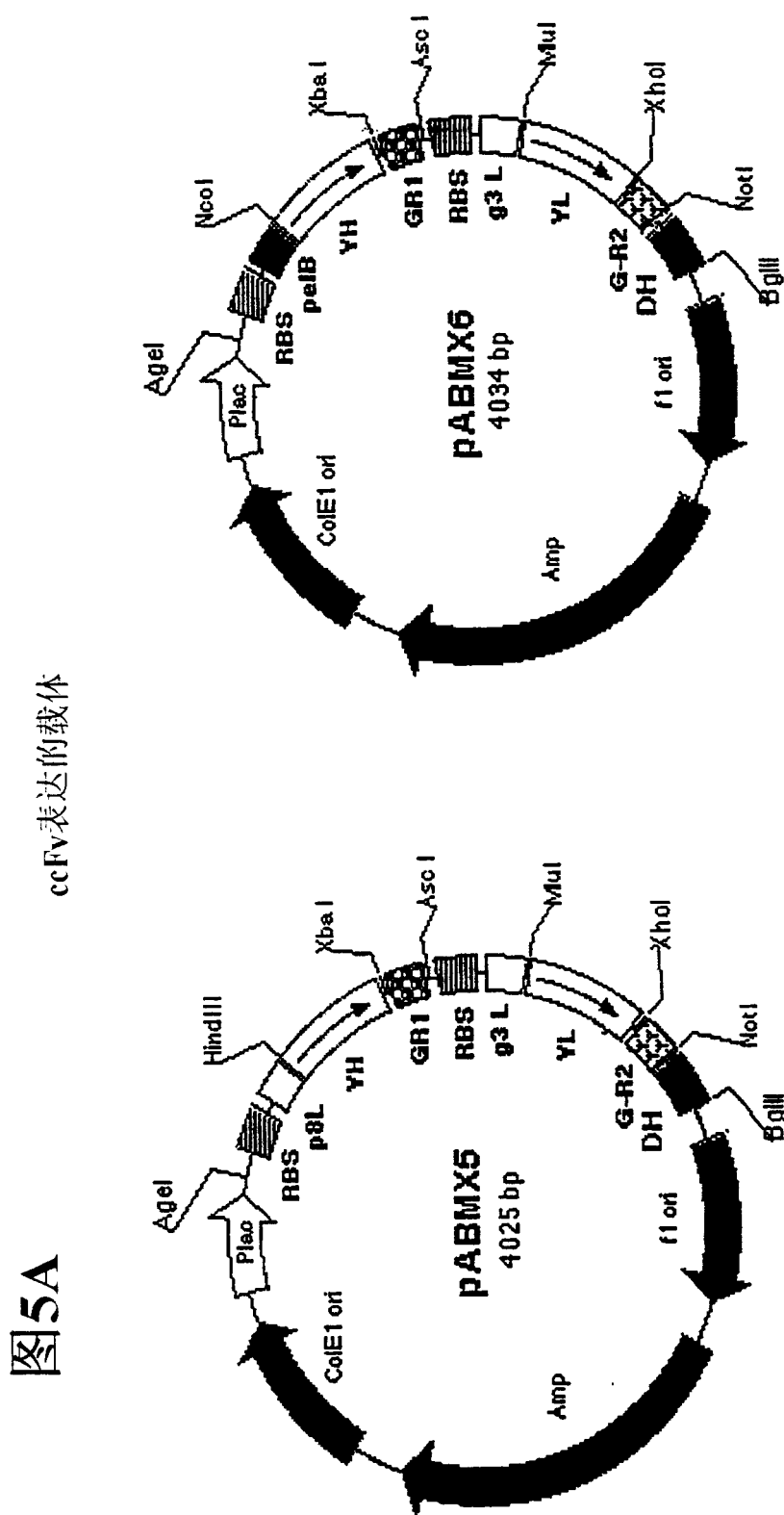


载体: Agel至SalI的序列 (核酸序列-SEQ ID NO: 9)

(氨基酸序列-SEQ ID NO: 10)

PABMX2 载体: AgeI至SalI的序列 (核酸序列-SEQ ID NO: 11)
(氨基酸序列-SEQ ID NO: 12)

94



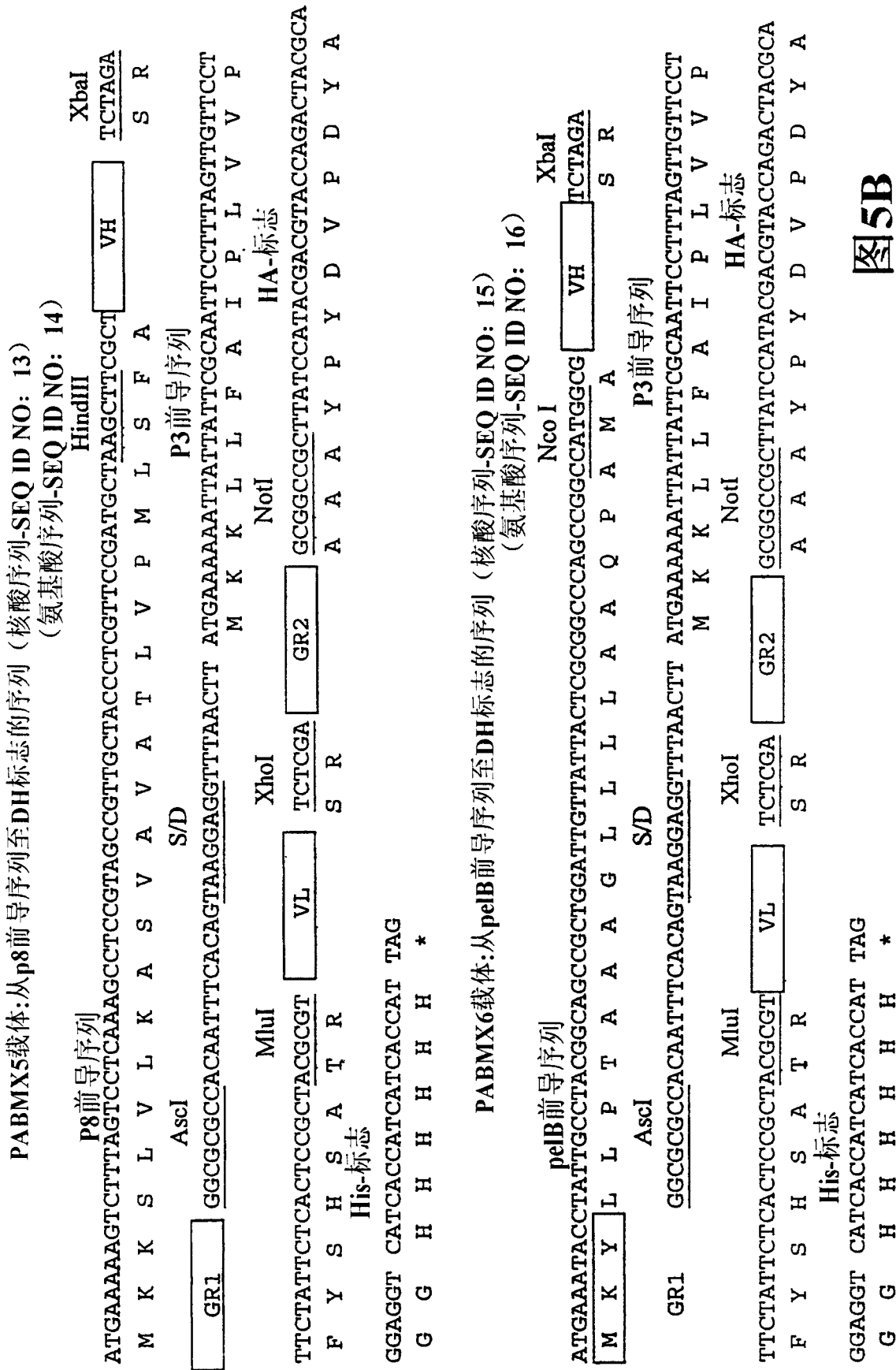
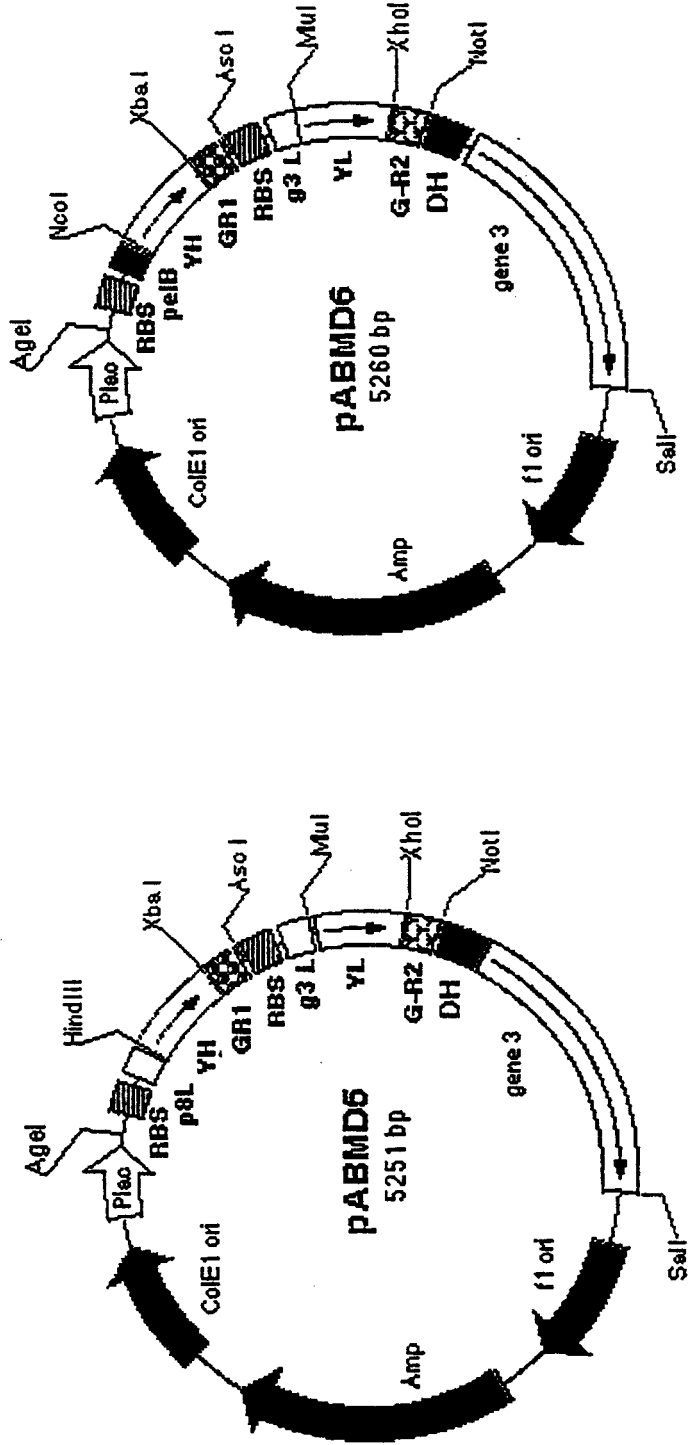


图5B

图6A

ccRv展示的载体



PABMD5 载体: 从HindIII到SalI的序列

(核酸序列-SEQ ID NO: 17)
(氨基酸序列-SEQ ID NO: 18)

ATGAAAAAGTCTTTAGTCTCAAGCCTCCGTAGCCGTTGCTACCCCTCGTTCCGATGCTAAGCTTCGCT VH XbaI
M K K S L V L K A S V A V A T L V P M L S F A S R

GR1 S/D

GGCGCGCCACAAATTTACAGTAAGGAGGTTTAACTT ATGAAAAAATTATATTTCGCAATTCCTTTAGTTGTTCCCT
M K K L L F A I P L V P P

His-标志

GGAGGT CATCACCATCATCACCAT TAG GGAGGCGGT ACTGTTGAAAGTTGT---CTGCGTAATAAGGAGTCTTAA GTCGAC
G G H H H H H * G G G T V E S C --- L R N K E S *

基因3

His-标志

GGAGGT CATCACCATCATCACCAT TAG GGAGGCGGT ACTGTTGAAAGTTGT---CTGCGTAATAAGGAGTCTTAA GTCGAC
G G H H H H H * G G G T V E S C --- L R N K E S *

PABMX6 载体: 从pelB前导序列至DH-标志的序列 (核酸序列-SEQ ID NO: 19)
(氨基酸序列-SEQ ID NO: 20)

pelB前导序列

ATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCCGCCATGGCG VH XbaI
M K Y L L P T A A A G L L L L A A Q P A M A S R

His-标志

GGAGGT CATCACCATCATCACCAT TAG GGAGGCGGT ACTGTTGAAAGTTGT---CTGCGTAATAAGGAGTCTTAA GTCGAC
G G H H H H H * G G G T V E S C --- L R N K E S *

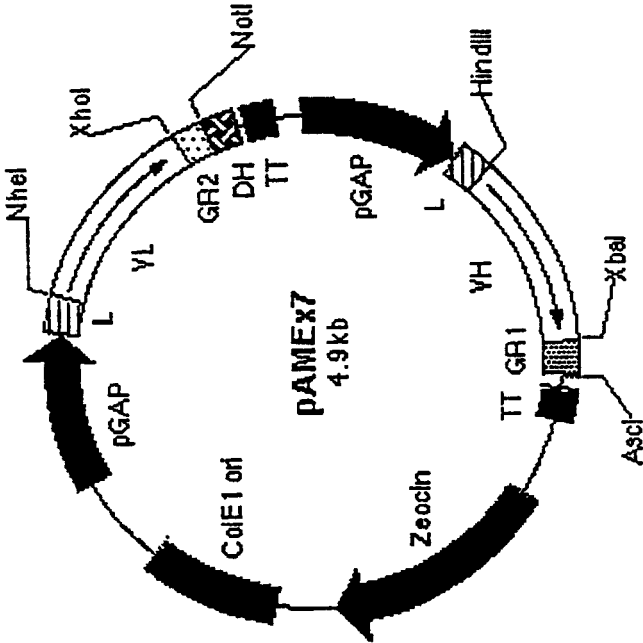
基因3

GGAGGT CATCACCATCATCACCAT TAG GGAGGCGGT ACTGTTGAAAGTTGT---CTGCGTAATAAGGAGTCTTAA GTCGAC
G G H H H H H * G G G T V E S C --- L R N K E S *

图6B

图7

在酵母中用于ccFv表达的载体



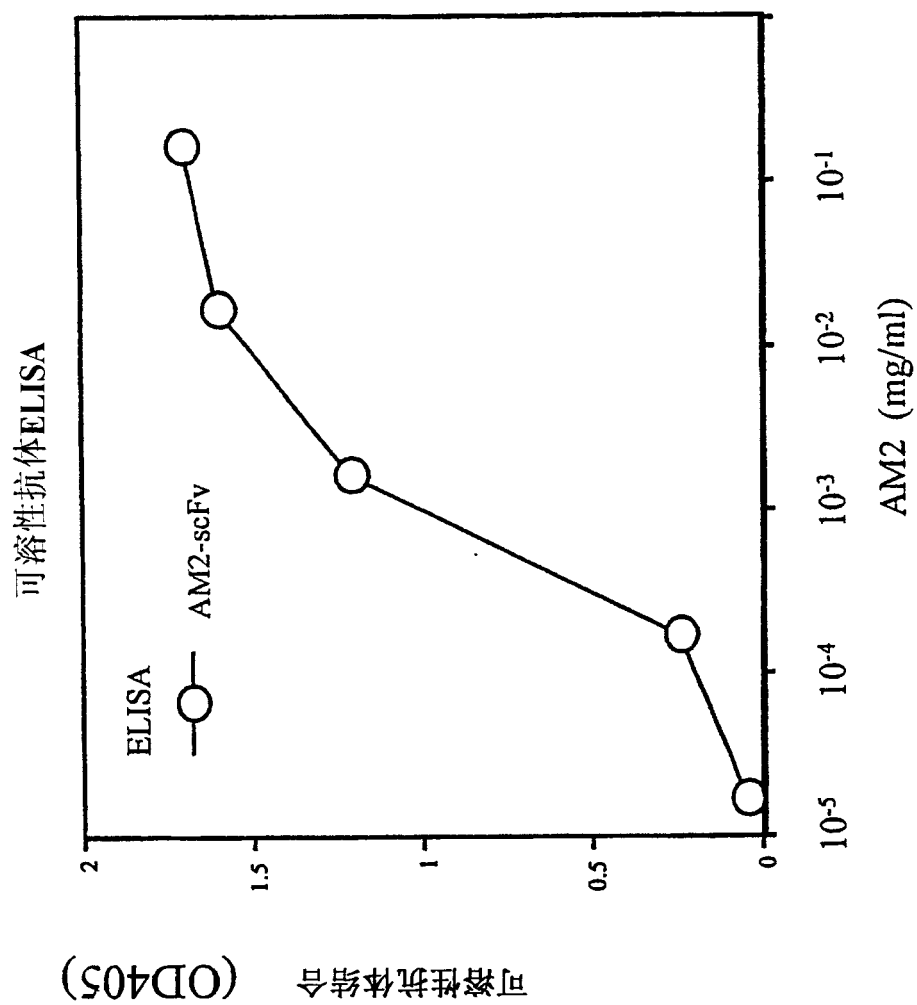
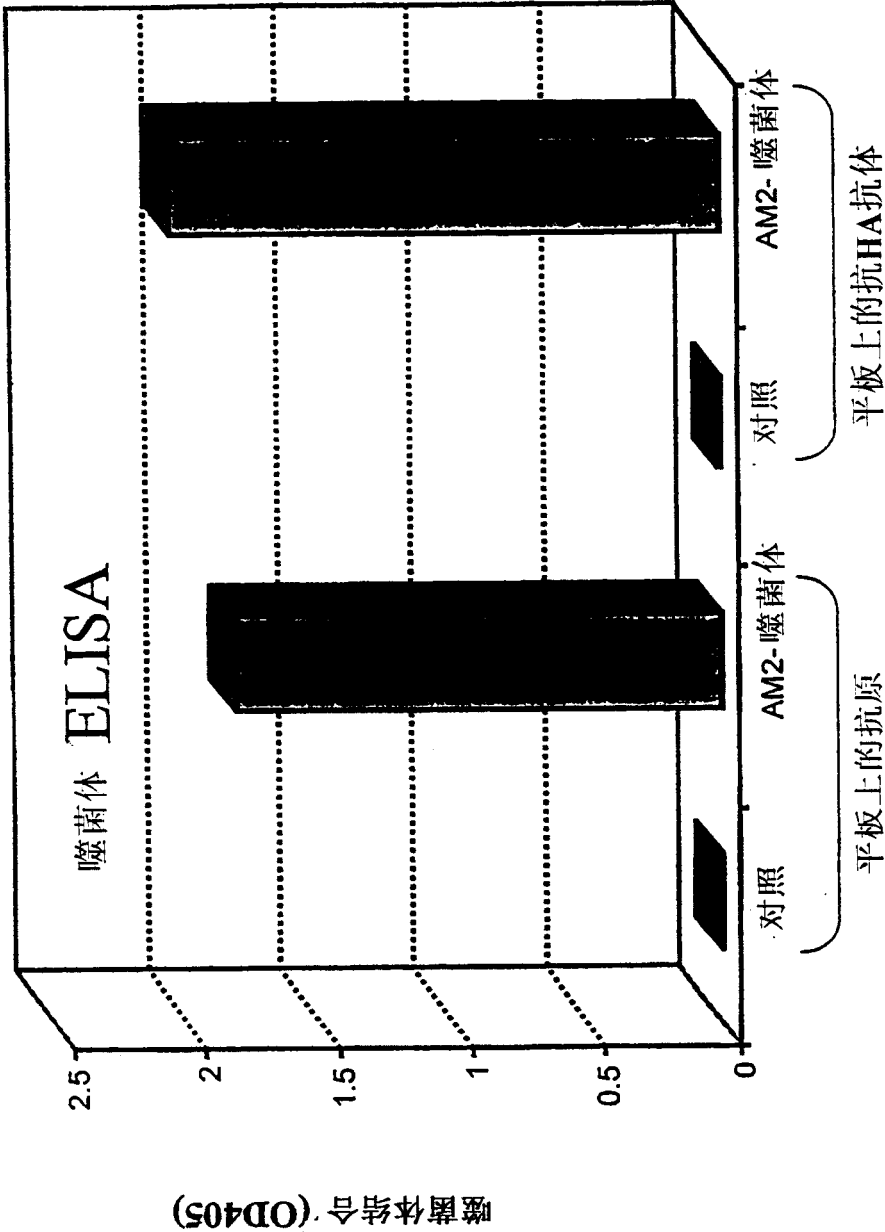


图8

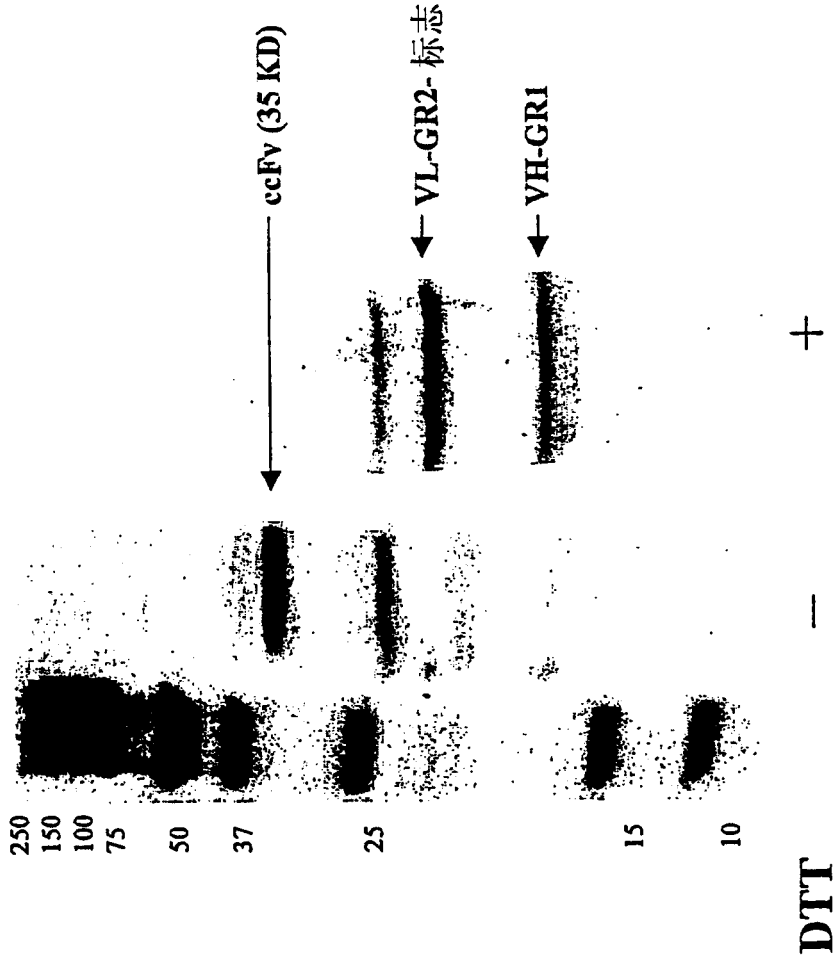
图9

通过pABMD1载体在噬菌体上展示AM2-scFv



可溶性AM1-ccFv抗体在大肠杆菌中的表达

图10A



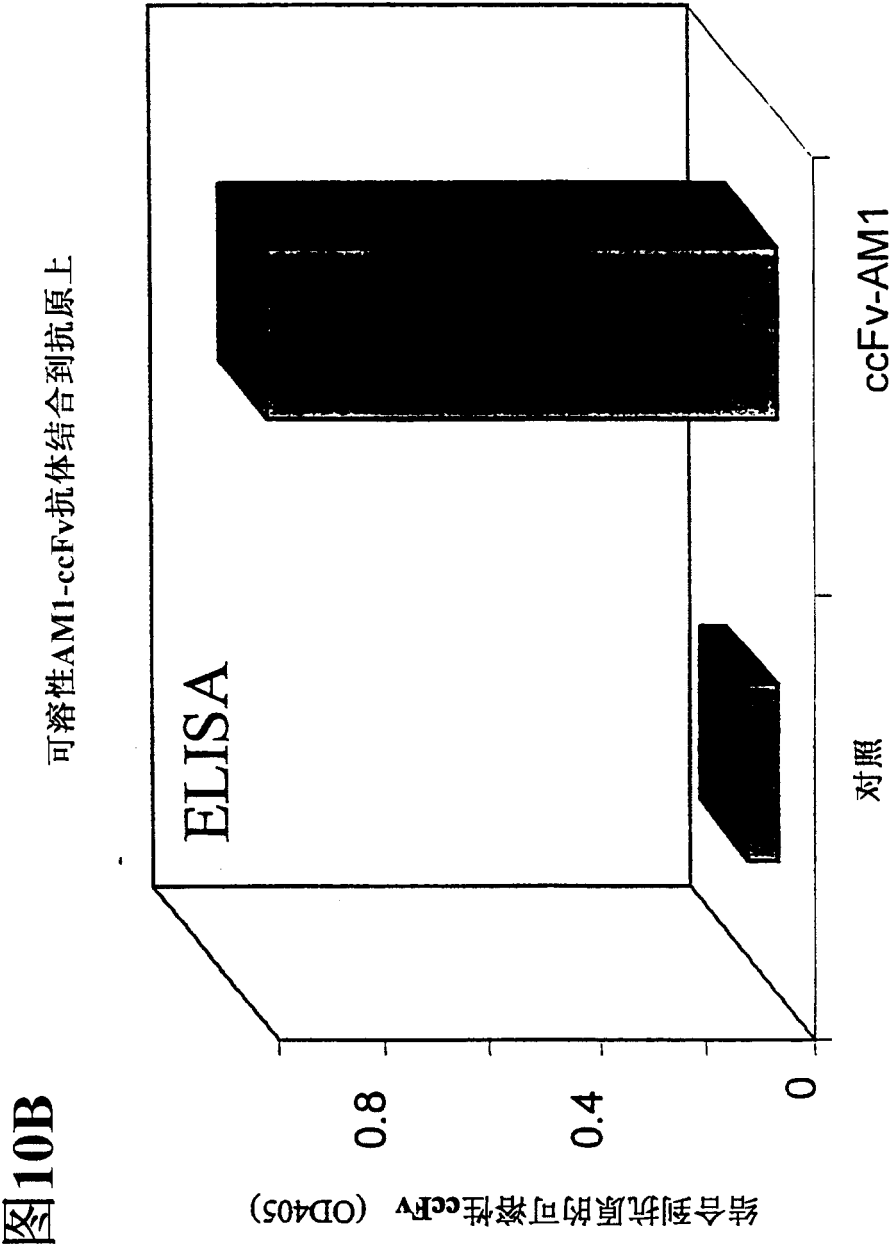


图11A AMI-ccFv抗体展示在噬菌体上

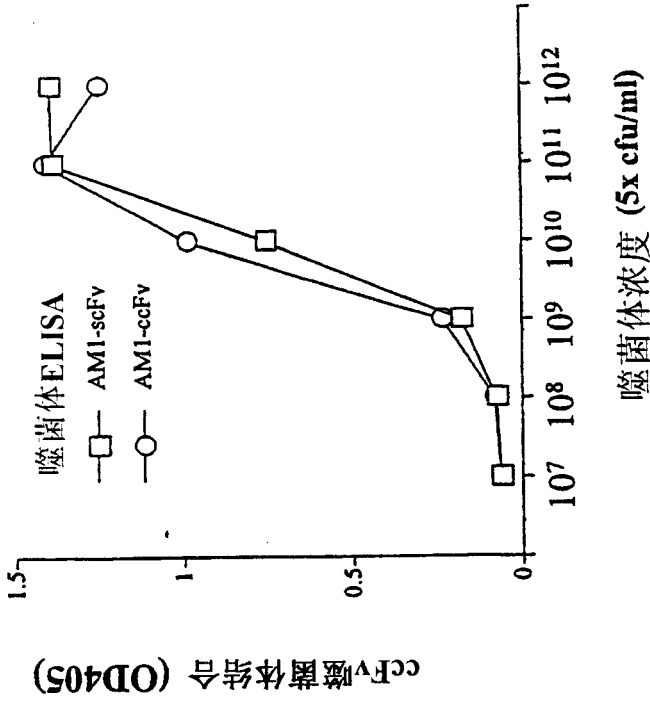
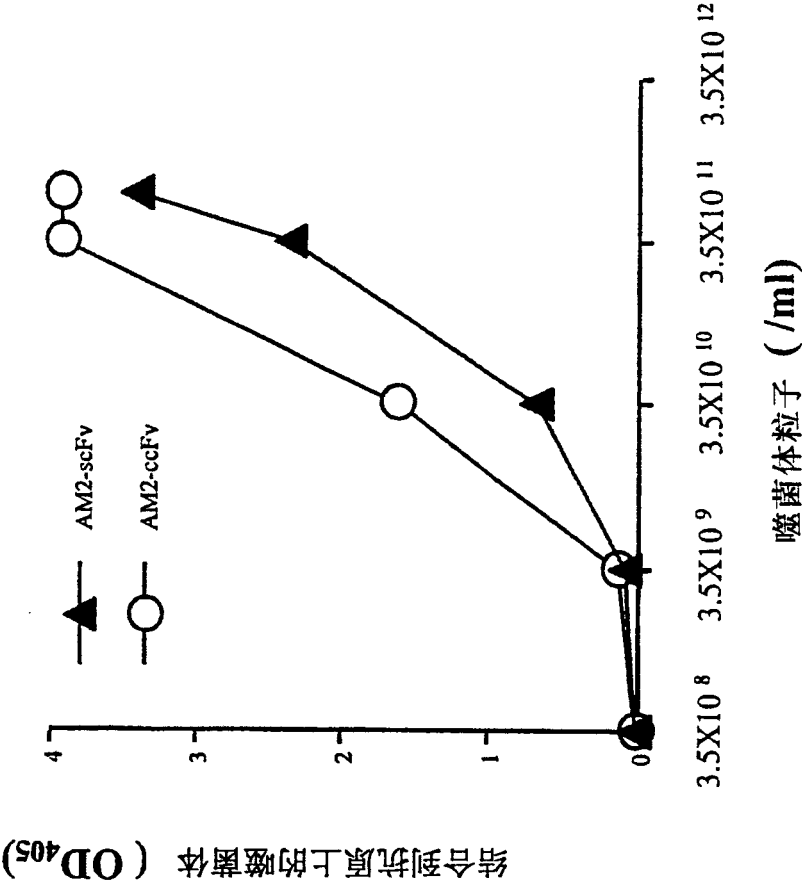
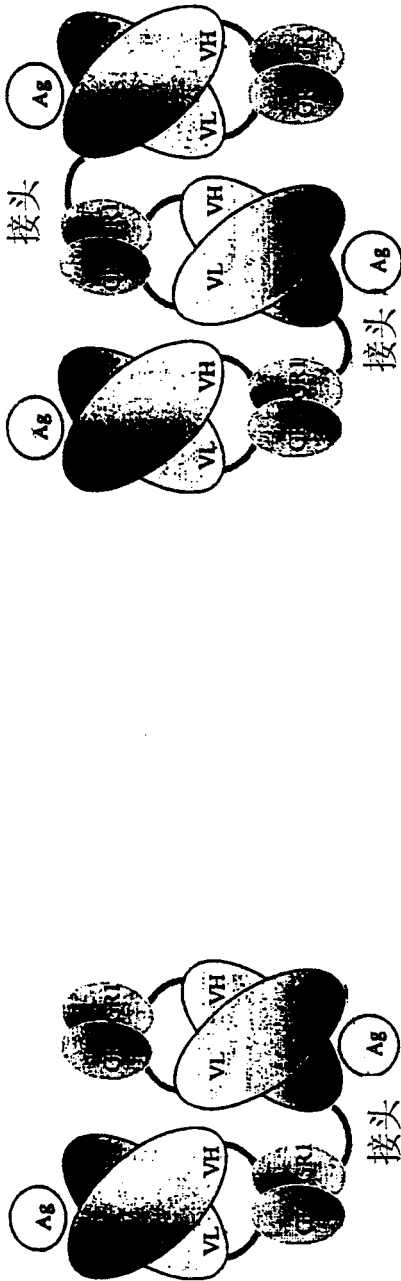


图11B
展示在噬菌体粒子上的AM2-ccFv和AM2-scFv的抗原结合能力的比较

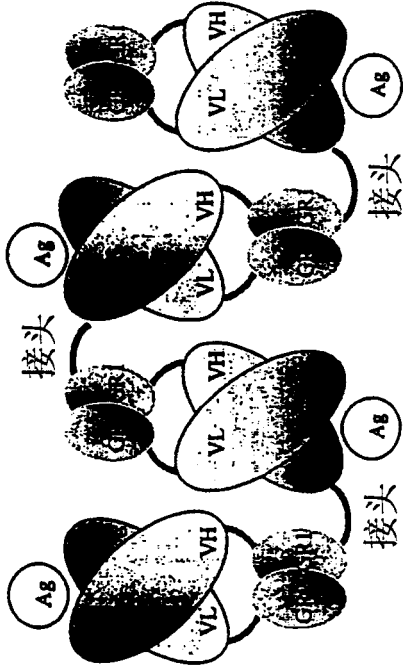


多价ccFv抗体

图12



二价(ccFv)2 肽 1: VH-GR1-VH-GR1
肽 2: VL-GR2



三价(ccFv)2 肽 1: VH-GR1-VH-GR1-VH-GR1
肽 2: VL-GR2

四价 (ccFv)4
肽1: VH-GR1-VH-GR1-VH-GR1-VH-GR1
肽2: VL-GR2

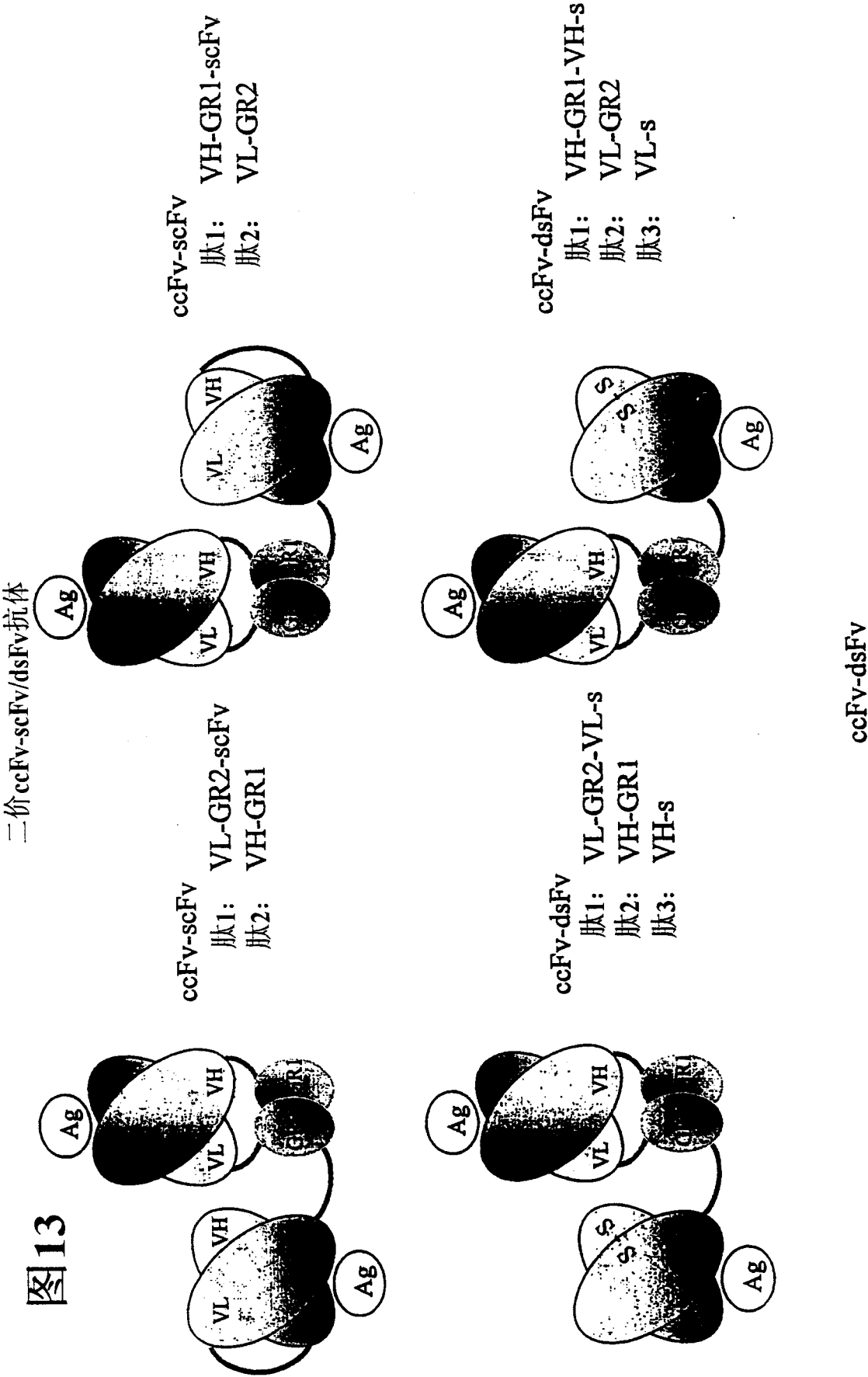


图14

三价ccFv-scFv-scFv/dsFv抗体

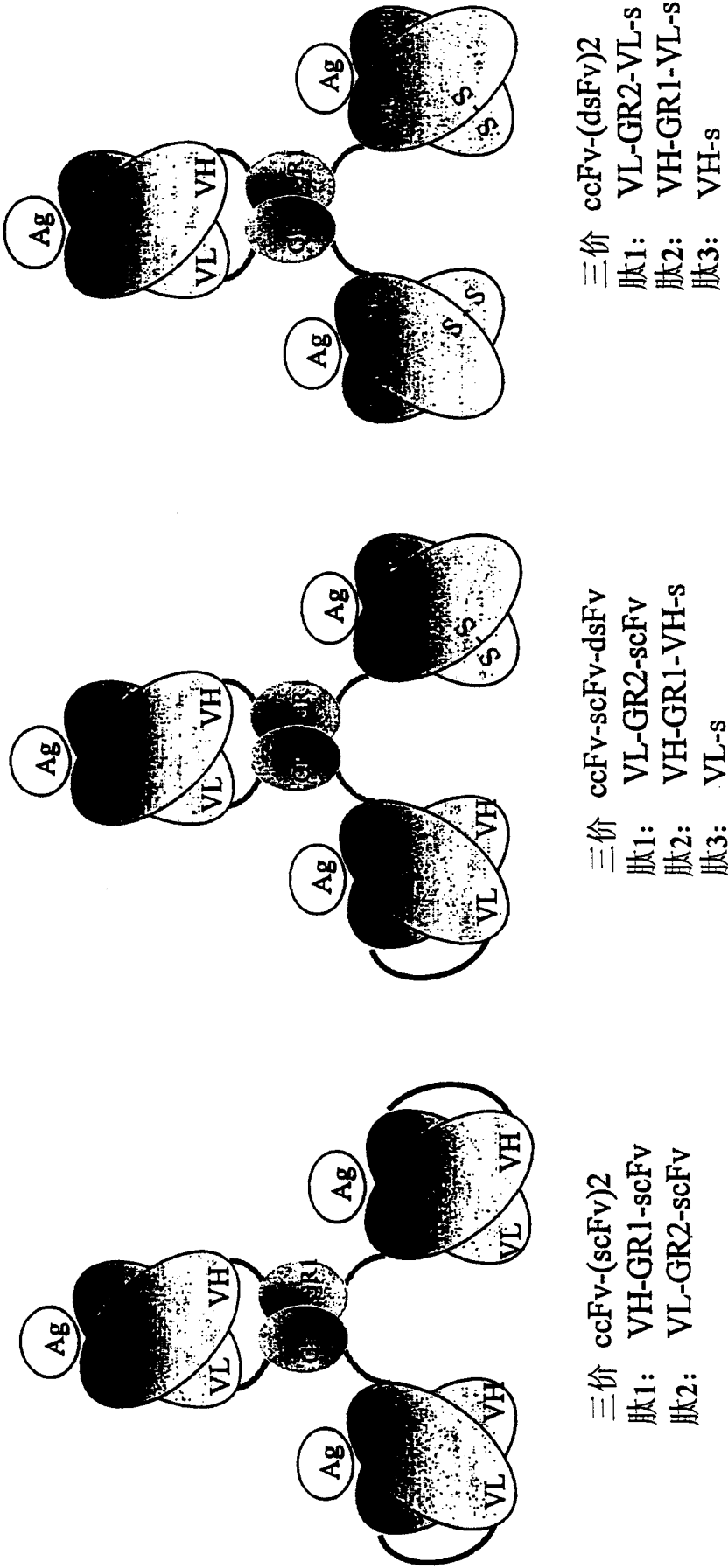


图15

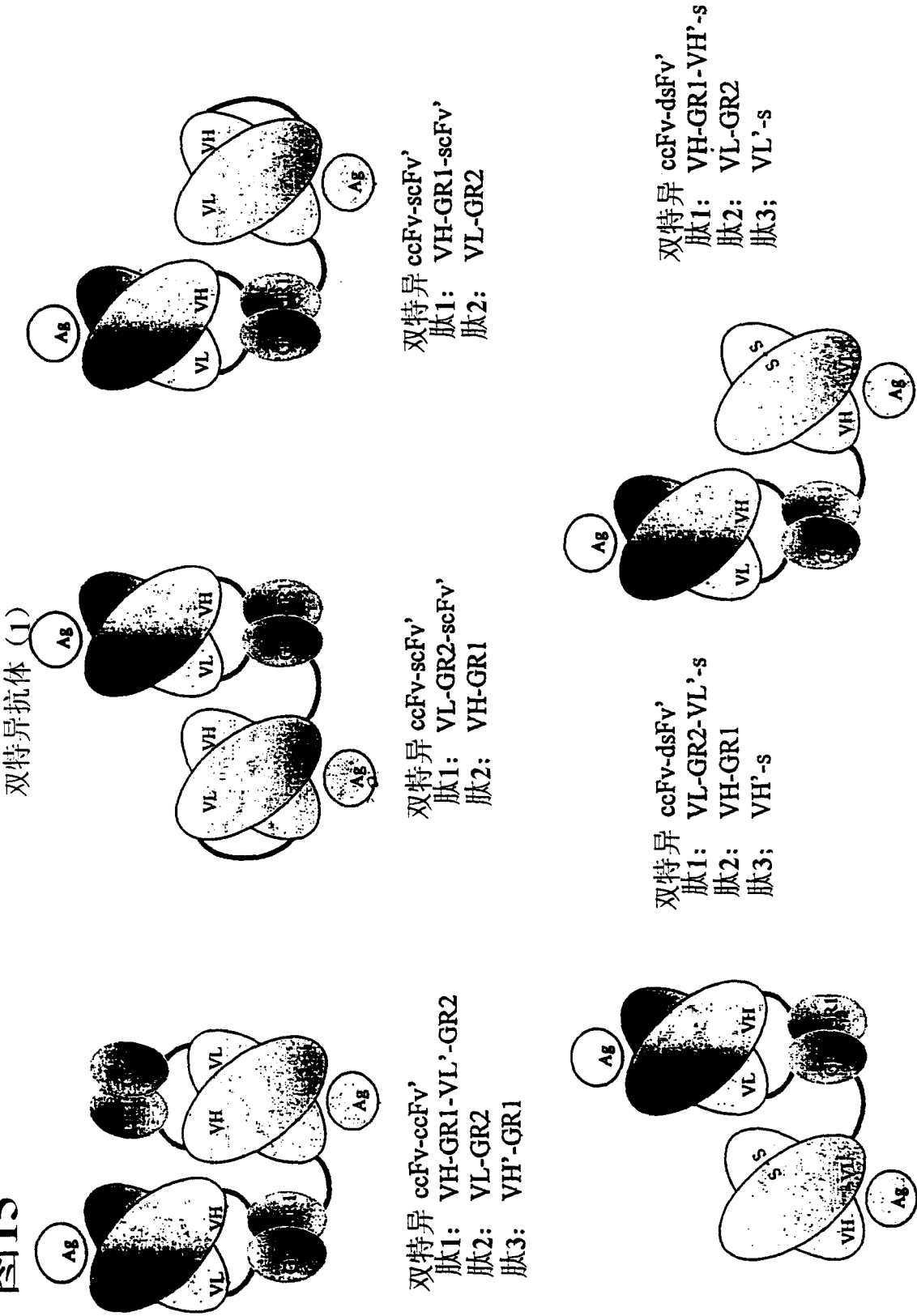


图16

双特异抗体 (2)

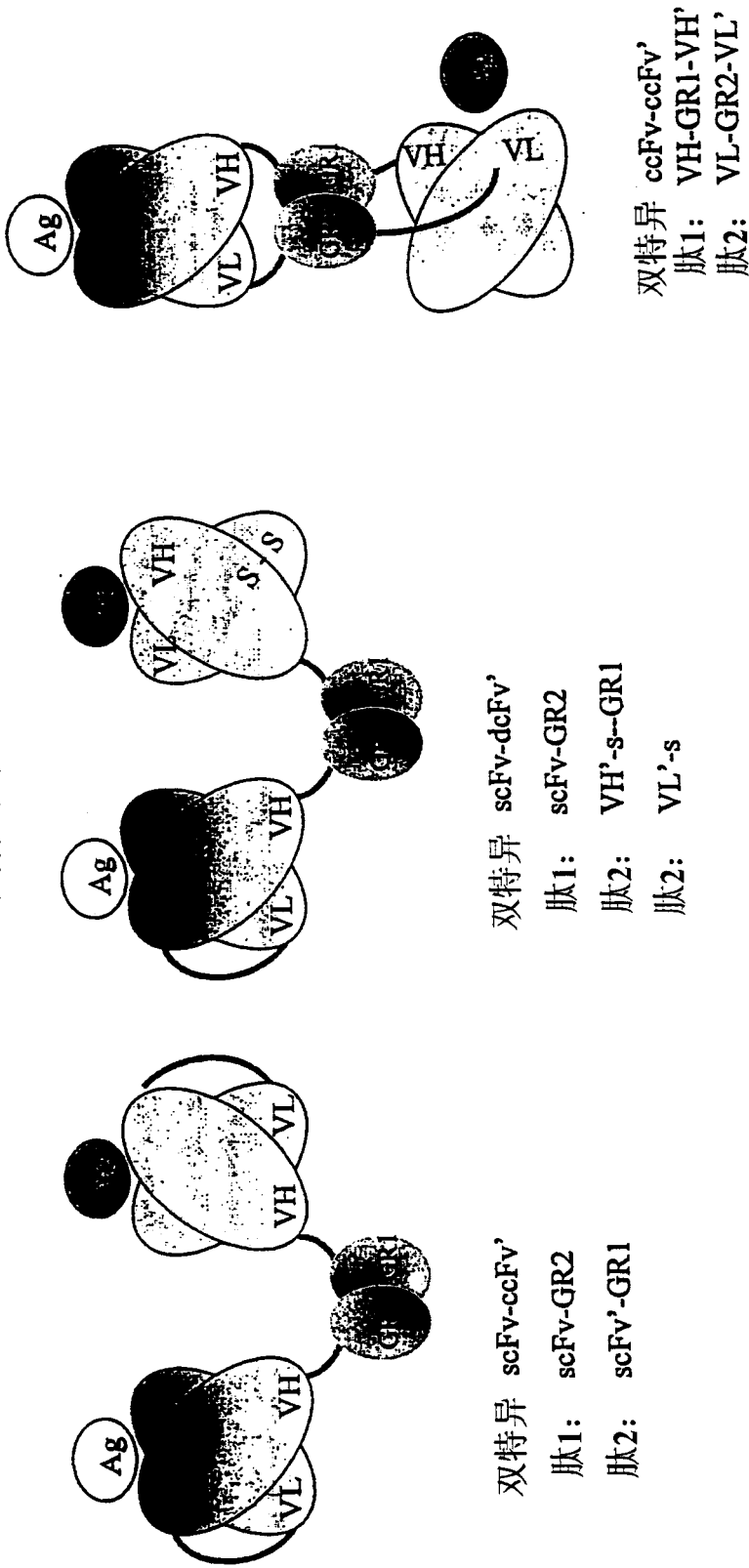


图17

三特异抗体

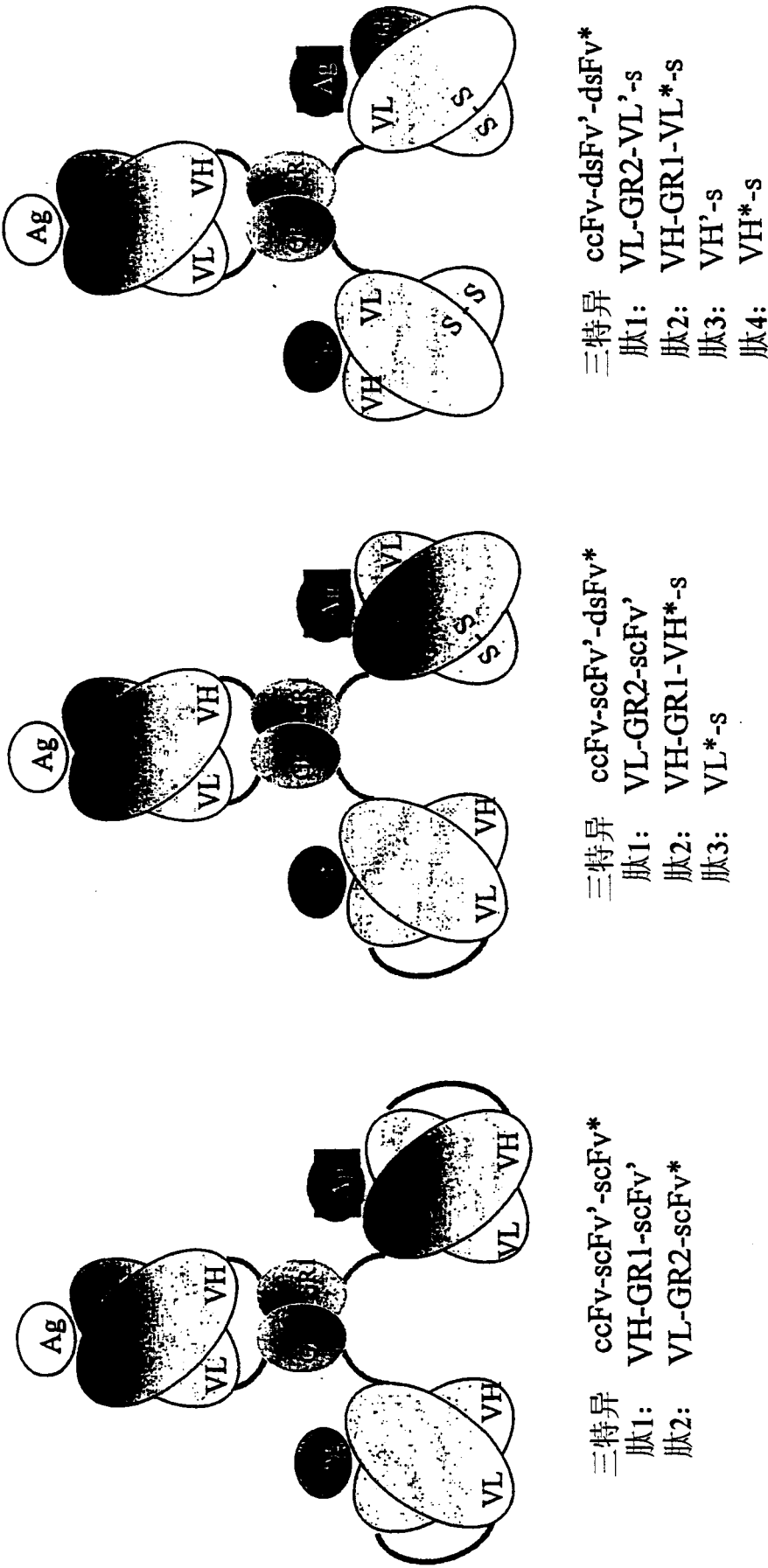


图18

单链ecFv抗体

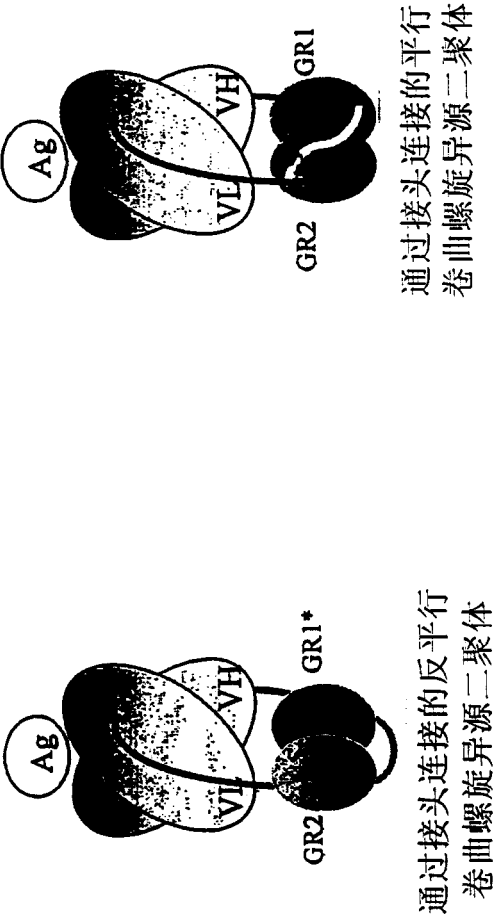


图19

ccFv展示及其在抗体文库构建中的用途

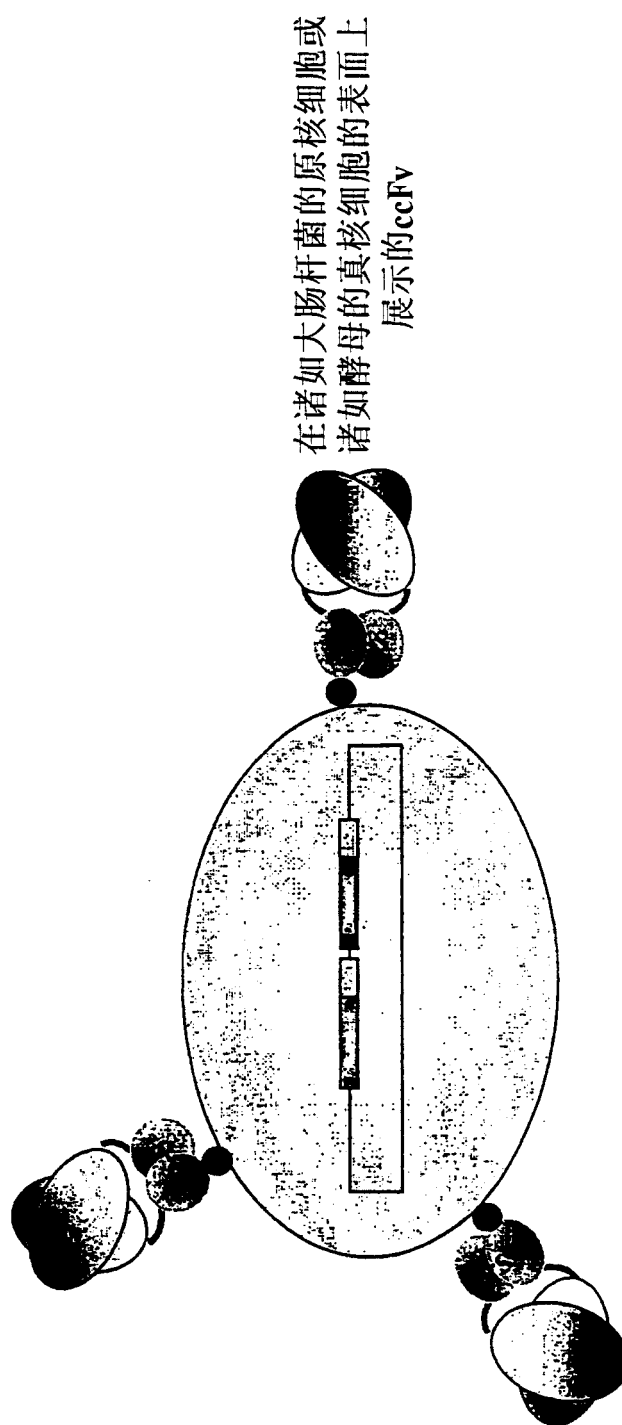
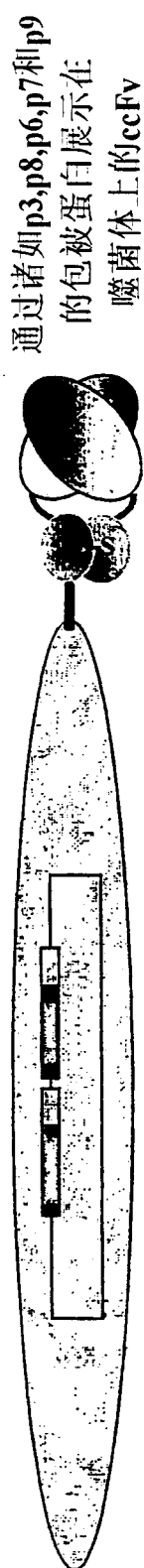
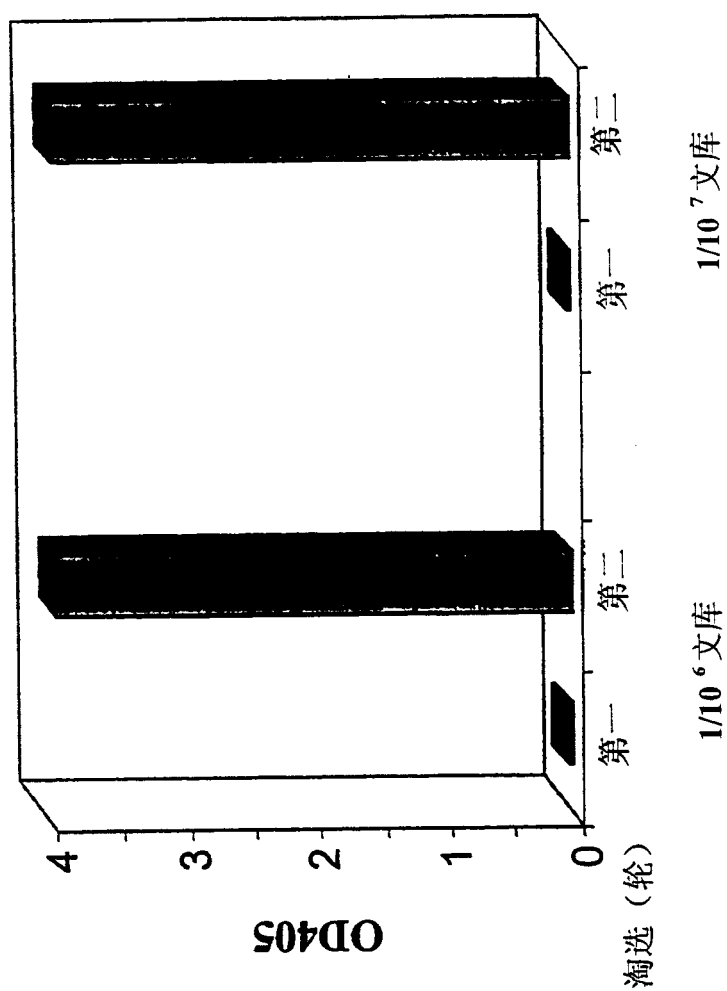


图20

在"模式文库"淘选中展示AM2-ccFv噬菌体的富集



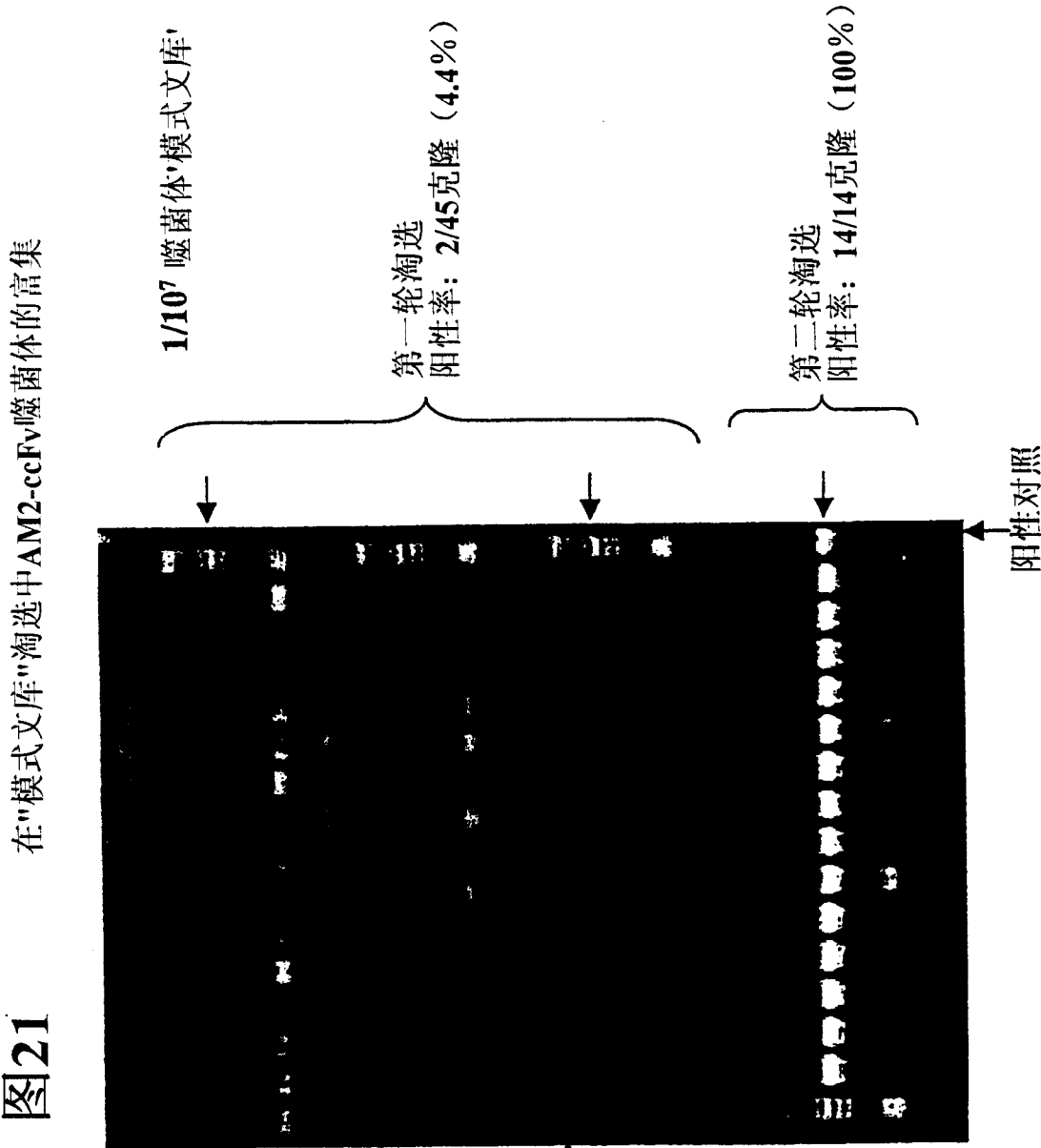


图22
抗体AM2-ccFv的CDR3的噬菌体文库设计



从淘选AM2的CDR3 VH文库中选择的ccFv变体的ELISA

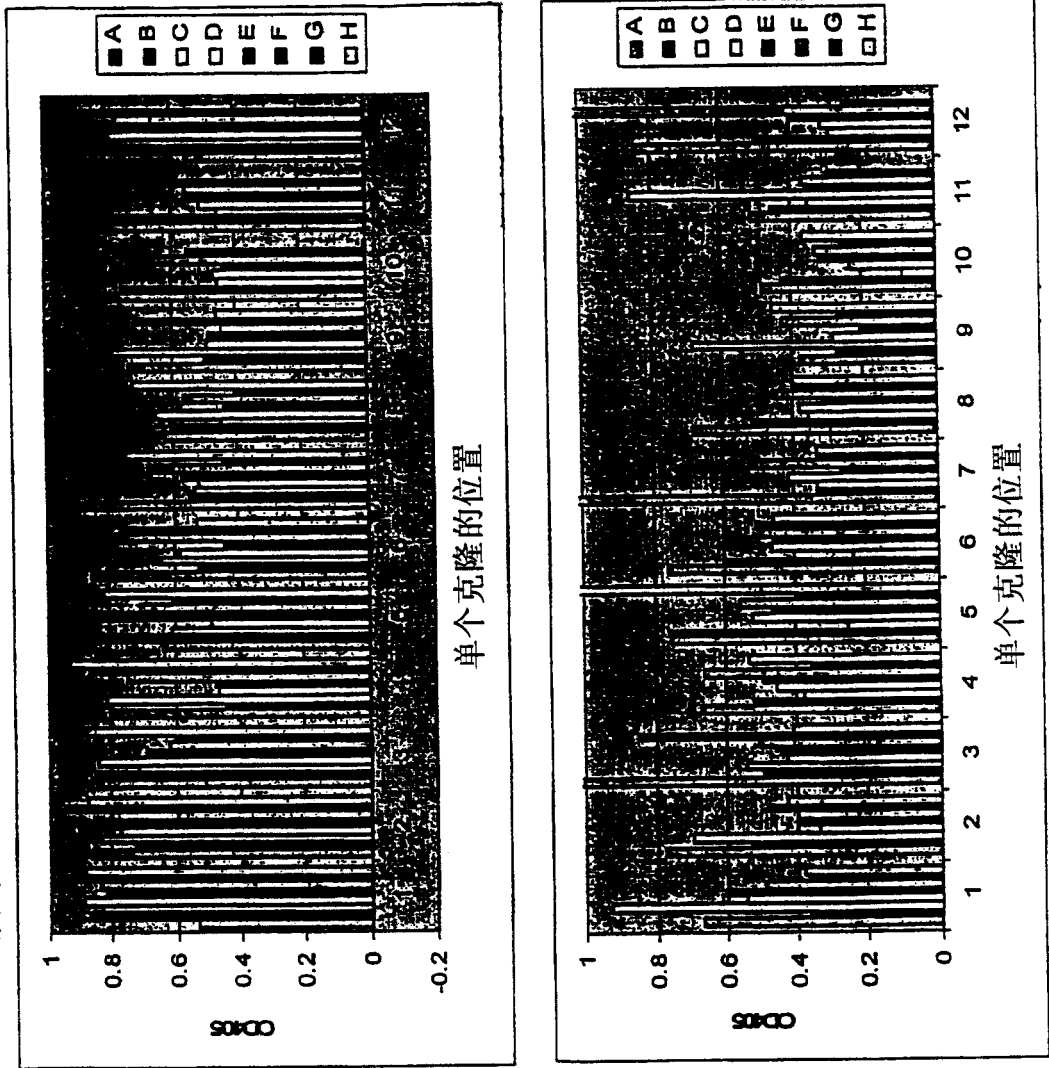
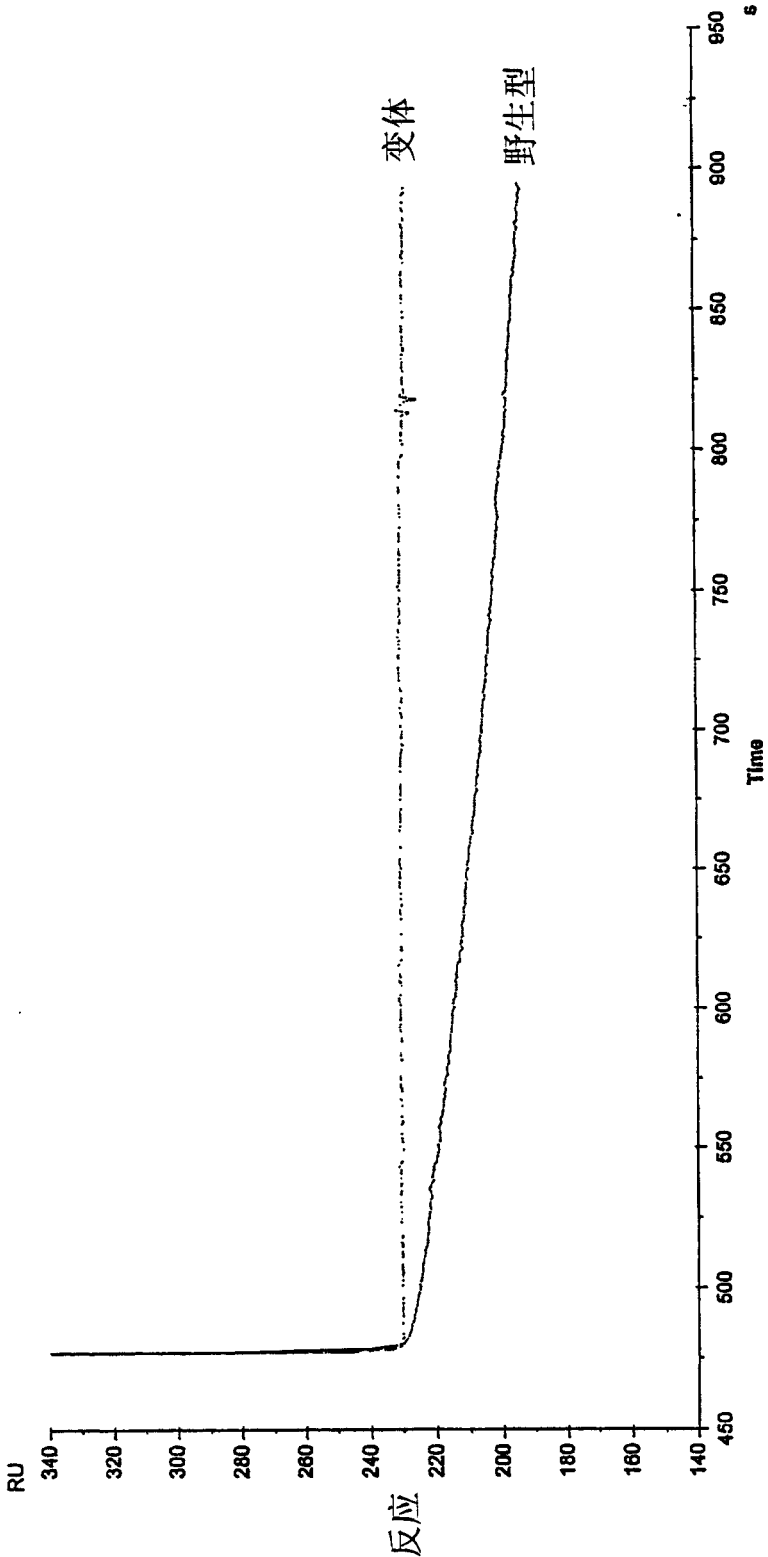


图24 从淘选AM2的CDR3 VH文库中选择的ccFv变体

氨基酸序列	命中率%
CAKXPYXXXXXXWYFDYWGQ	100
Y	100
Y	60
Y	20
S	10
H	10
D	70
G	30
S	80
T	20
S	100
S	100
H	100
V	100
CAKYPYYGTSHWYFDYWGQ	共有序列

图25
与野生型AM2相比AM2变体具有较低的 K_{on}



专利名称(译)	产生嵌合异源多聚体的组合物和方法		
公开(公告)号	CN1558916A	公开(公告)日	2004-12-29
申请号	CN02818596.X	申请日	2002-07-31
[标]发明人	王材力 钟平宇 刘胜江 罗培志 李胜峰 王欣慰		
发明人	王材力 钟平宇 刘胜江 罗培志 李胜峰 王欣慰		
IPC分类号	G01N33/53 C07K14/705 C07K16/00 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12P21/08 G01N33/566 C12N15/00 C12N15/63		
CPC分类号	C07K2319/00 C07K16/00 C07K2317/56		
代理人(译)	丁业平 王维玉		
优先权	09/921144 2001-08-01 US		
外部链接	Espacenet SIPO		

摘要(译)

本发明提供一种把单体多肽特异性组装成异源多聚体的方法。尤其可用本技术来获得具遗传多样性的异源多聚体如抗原结合单位库。本发明也提供由本发明所述的技术组装的非单链和单链抗原结合单位。本发明也提供重组多核苷酸，载体，宿主细胞以及制造受试抗原结合单位的试剂盒。本发明进一步提供使用受试抗原结合单位的方法。

图1

