



(12) 发明专利

(10) 授权公告号 CN 101977932 B

(45) 授权公告日 2014. 09. 03

(21) 申请号 200980110221. 1

113 页第 29-32 行.

(22) 申请日 2009. 01. 30

W0 2006072620 A1, 2006. 07. 13, 权利要求书, 实施例 1, 4-5, 7.

(30) 优先权数据

61/063, 245 2008. 01. 31 US

US 2007148170 A1, 2007. 06. 28, 序列 25-27.

(85) PCT 国际申请进入国家阶段日

2010. 09. 21

W0 2007076319 A2, 2007. 07. 05, 序列 43, 45.

(86) PCT 国际申请的申请数据

PCT/US2009/032692 2009. 01. 30

W0 2006036834 A2, 2006. 04. 06, 全文.

W0 2006114700 A2, 2006. 11. 02, 序列 1, 12 及实施例 1, 权利要求书.

(87) PCT 国际申请的公布数据

W02009/099961 EN 2009. 08. 13

Matthias J. Feige et al. Folding Mechanism of the CH2 Antibody Domain.

(73) 专利权人 美国政府健康及人类服务部

地址 美国马里兰州

《J. Mol. Biol. 》. 2004, 第 344 卷 (第 1 期), 第 107-118 页.

(72) 发明人 迪米特尔·S·迪米特罗夫

审查员 李谦

(74) 专利代理机构 中原信达知识产权代理有限

责任公司 11219

代理人 张颖 樊卫民

(51) Int. Cl.

C07K 16/00 (2006. 01)

C07K 16/10 (2006. 01)

C07K 19/00 (2006. 01)

(56) 对比文件

US 5965709 A, 1999. 10. 12, 序列 60.

US 20050136049 A1, 2005. 06. 23, 说明书第

权利要求书 2 页 说明书 48 页

序列表 49 页 附图 31 页

(54) 发明名称

工程化抗体恒定结构域分子

(57) 摘要

本申请描述的是工程化抗体恒定结构域分子, 例如 CH2 或 CH3 结构域分子, 所述工程化抗体恒定结构域分子包含至少一个突变, 或者包含在所述 CH2 结构域的环区域中移入的至少一个互补决定区 (CDR) 或其功能片段。本申请所述的 CH2 结构域分子是小的、稳定的、可溶的, 表现出很少毒性乃至没有毒性并且能够结合抗原。

1. 一种包含 IgG 的免疫球蛋白 CH2 结构域的多肽,其中所述 CH2 结构域包含在 N-端 A 链的第一氨基酸取代和在 C-端 G 链的第二氨基酸取代,其中所述第一氨基酸和第二氨基酸取代各自用半胱氨酸残基置换原始的残基,其中所述半胱氨酸残基形成二硫键,并且参考 SEQ ID NO:5 编号,所述第一氨基酸取代是 L12 至 C12,并且所述第二氨基酸取代是 K104 至 C104,或者所述第一氨基酸取代是 V10 至 C10,并且所述第二氨基酸取代是 K104 至 C104,并且其中所述多肽具有小于 15kD 的分子量。

2. 如权利要求 1 所述的多肽,其中所述 CH2 结构域包含 7 个氨基酸的 N 端平截。

3. 如权利要求 1 所述的多肽,其包括 m1a1、m1a2、m1a3、m1a1CH2 或 m1a2CH2,其中所述多肽特异性结合抗原。

4. 如权利要求 3 所述的多肽,其中抗原来自病原体。

5. 如权利要求 4 所述的多肽,其中病原体是病毒。

6. 如权利要求 5 所述的多肽,其中病毒是人免疫缺陷病毒(HIV)。

7. 如权利要求 1-6 任一项所述的多肽,其中所述 CH2 结构域是非糖基化的。

8. 如权利要求 1-6 任一项所述的多肽,其中所述 CH2 结构域是糖基化的。

9. 如权利要求 1-8 任一项所述的多肽,具有 12kD 至 14kD 的分子量。

10. 如权利要求 1-9 任一项所述的多肽,其中 CH2 结构域还包含 1 至 4 个氨基酸的 C 端平截。

11. 核酸分子,其编码权利要求 1-10 任一项的多肽。

12. 载体,其包含权利要求 11 的核酸分子。

13. 分离的宿主细胞,其包含权利要求 12 的载体。

14. 组合物,其包含权利要求 1-10 任一项的多肽和可药用载体。

15. 组合物,其包含权利要求 1-10 任一项的多肽,所述多肽与效应物分子或可检测标记结合。

16. 一种鉴定特异性结合靶抗原的重组 CH2 结构域的方法,所述方法包括:

(a)提供在它们的表面上展示重组 IgG CH2 结构域的噬菌体粒子的文库,其中所述 CH2 结构域包含 7 个氨基酸的 N 端缺失,并且其中 CH2 结构域包含在 N-端 A 链的第一氨基酸取代和在 C-端 G 链的第二氨基酸取代,其中所述第一和第二氨基酸取代各自用半胱氨酸残基置换原始的残基,并且参考 SEQ ID NO:5 编号,所述第一氨基酸取代是 L12 至 C12,并且所述第二氨基酸取代是 K104 至 C104,或者所述第一氨基酸取代是 V10 至 C10,并且所述第二氨基酸取代是 K104 至 C104,其中所述半胱氨酸残基形成二硫键,并且其中所述 CH2 结构域具有小于 15KD 的分子量,其中所述文库是通过以下产生:

(i) 提供编码 CH2 结构域的遗传多样性群体的核酸分子的文库,其中所述遗传多样性群体是通过将突变引入 CH2 结构域的一个或多个环区域中来提供,其中所述环区域选自环 1、环 2 和环 3,并且其中将突变引入一个或多个环区域中包括用 A, Y, D 或 S 残基随机置换所有的环残基,以及在每个环的 C-末端插入另外的 G 残基;并且

(ii) 在重组的宿主细胞中表达核酸分子的文库,从而 CH2 结构域被表达在所述粒子的表面上,并且所述 CH2 结构域的核酸分子被所述粒子的遗传材料编码;

(b) 将粒子的文库与所述靶抗原接触以选择特异性结合靶抗原的粒子;以及

(c) 从表达特异性结合所述靶抗原的 CH2 结构域的粒子中克隆所述 CH2 结构域的核酸

分子,由此鉴定特异性结合所述靶抗原的 CH2 结构域。

17. 一种制备 IgG 的重组 CH2 结构域的文库的方法,所述方法包括(i)将突变引入 CH2 结构域支架的一个或多个环区域中,其中所述环区域选自环 1、环 2 和环 3,并且其中将突变引入一个或多个环区域中包括用 A, Y, D 或 S 残基随机置换所有的环残基,以及在每个环的 C- 末端插入另外的 G 残基,或者(ii)用来自异源免疫球蛋白可变重链(VH)结构域的 CDR3 置换所述 CH2 结构域支架的环 A-B 或环 E-F 的一部分,其中 CH2 结构域支架还包含在 N- 端 A 链的第一氨基酸取代和在 C- 端 G 链的第二氨基酸取代,其中所述第一和第二氨基酸取代各自用半胱氨酸残基置换原始的残基,并且参考 SEQ ID NO:5 编号,所述第一氨基酸取代是 L12 至 C12,并且所述第二氨基酸取代是 K104 至 C104,或者所述第一氨基酸取代是 V10 至 C10,并且所述第二氨基酸取代是 K104 至 C104,其中所述半胱氨酸残基形成二硫键。

18. 如权利要求 17 所述的方法,其中所述 CH2 结构域支架还包含 7 个氨基酸的 N 端平截。

19. 如权利要求 17 或权利要求 18 所述的方法,其中所述 CH2 结构域支架还包含 1 个至 4 个氨基酸的 C 端平截。

20. 一种鉴定特异性结合靶抗原的重组 CH2 结构域的方法,所述方法包括使通过权利要求 17-19 任一项所述方法产生的文库与所述靶抗原接触以选择特异性结合所述靶抗原的重组 CH2 结构域。

工程化抗体恒定结构域分子

[0001] 相关申请的交叉引用

[0002] 本申请要求享有 2008 年 1 月 31 日提交的美国临时申请第 61/063,245 号的优先权,通过引用将其整体并入本文。

[0003] 领域

[0004] 本发明涉及抗体,特别地涉及在特定位置突变的抗体恒定结构域和 / 或从特异性结合目的抗原的异源抗体移入一个或多个可变链环的抗体恒定结构域。

[0005] 背景

[0006] 常规抗体是包含至少四条多肽链的大型多亚基蛋白复合物,所述四条多肽链包括两条轻链和两条重链。抗体的重链和轻链含有结合抗原的可变(V)区和提供结构支撑和效应物功能的恒定(C)区。抗原结合区包含两个分离的结构域,重链可变结构域(V_H)和轻链可变结构域(V_L)。互补决定区(CDR)为抗体可变结构域中的短氨基酸序列,它提供抗原特异性。抗体分子的重链和轻链各提供三个 CDR(CDR1、CDR2 和 CDR3),因此能够与抗原接触的每个抗体具有六个 CDR,产生抗原特异性。

[0007] 典型的抗体如 IgG 分子具有大约 150kD 的分子量。治疗应用可能受限,因为抗体相对较大的尺寸能够限制组织渗透或表位进入。

[0008] 已在鉴定了蛋白酶消化后天然存在的抗体的许多较小的抗原结合片段(例如, Fab、Fab' 和 $F(ab')_2$)。这些抗体片段具有范围在大约 50kD 至 100kD 的分子量。已使用重组的方法来产生替代的抗原结合片段,称为单链可变片段(scFv),它是由通过合成的肽接头相接合的 V_L 和 V_H 构成。scFv 分子具有大约 25-30kD 的分子量。

[0009] 尽管在人类和大多数其他的哺乳动物中的天然存在的抗体的抗原结合单元一般已知是由一对可变区组成,但是骆驼科(camelid)物种表达大比例的缺乏轻链序列的全功能、高特异性抗体。骆驼科重链抗体作为通过其恒定区二聚的单一重链的同源二聚体而存在(美国专利第 5,840,526 号和第 6,838,254 号;以及美国专利申请公开号 2003-0088074)。这些骆驼科重链抗体的可变结构域称为 V_HH 结构域,它们在被分离为 V_H 链的片段时保留了高特异性结合抗原的能力(Hamers-Casterman 等人, Nature 363: 446-448, 1993; Gahroudi 等人, FEBS Lett. 414: 521-526, 1997)。

[0010] 结合抗原的单 V_H 结构域称作结构域抗体(dAb),也已将其从由免疫小鼠的基因组 DNA 扩增的小鼠 V_H 基因文库中鉴定出来(Ward 等, Nature 341: 544-546, 1989)。还已描述了能够高亲和力结合抗原的人类单免疫球蛋白可变结构域多肽(参见例如, PCT 公开号 WO 2005/035572 和 WO 2003/002609)。

[0011] 然而,对能够特异性结合抗原的非常小的抗体的需求仍存在。此类小分子可能提供提高的表位进入、更好的组织渗透,并且可能被用于利用抗体或其片段的任何诊断性或治疗性的应用。

[0012] 概述

[0013] 本公开涉及工程化抗体恒定结构域分子。在一个实施方案中,抗体恒定结构域是来自 IgG、IgA 或 IgD 的 CH2 结构域。在另一个实施方案中,抗体恒定结构域是来自 IgE 或

IgM 的 CH3 结构域。如本文所述, CH2 或 CH3 结构域分子是小的、稳定的、可溶的, 具有最小毒性乃至没有毒性并且有效地结合抗原。因此, 本文提供了包含免疫球蛋白 CH2 或 CH3 结构域的多肽, 其中所述 CH2 或 CH3 结构域的环的至少一个被突变, 或者所述 CH2 或 CH3 结构域的环区域的至少一部分被来自异源免疫球蛋白可变结构域的互补决定区 (CDR) 或其功能片段 (如含有特异性决定残基 (SDR) 的功能片段) 取代, 或者二者兼有。本文所述的 CH2 和 CH3 结构域分子具有小于约 15kD 的分子量。本文还提供了包含 CH2 或 CH3 结构域分子的组合物、文库和试剂盒以及使用方法。还提供了表现出提高的稳定性的重组恒定结构域, 其可被用作构建抗原结合的 CH2 或 CH3 结构域的支架。还提供了鉴定特异性结合抗原的重组 CH2 或 CH3 结构域的方法和产生包含重组 CH2 或 CH3 结构域的文库的方法。

[0014] 上述的和其他的特征和优点将从以下关于附图进行的若干实施方案的详细描述中变得更明显。

[0015] 附图简述

[0016] 图 1A 是免疫球蛋白分子的示意图。常规抗体是包含至少四条多肽链的大型多亚基蛋白复合物, 所述四条多肽链包括两条轻 (L) 链和两条重 (H) 链。抗体的重链和轻链含有结合抗原的可变 (V) 区和提供结构支撑和效应物功能的恒定 (C) 区 (如 CH1、CH2 和 CH3 结构域)。抗原结合区包含两个分离的结构域, 重链可变结构域 (V_H) 和轻链可变结构域 (V_L)。

[0017] 图 1B 显示人类重链可变结构域的共有氨基酸序列 (SEQ ID NO :1)。指出了 CDR1、CDR2、CDR3 (表示为 H1、H2 和 H3) 的位置。还显示了三种不同的抗原特异性人类抗体的重链的氨基酸序列 (SEQ ID NO :2-4)。显示的数字是基于 Kabat 编号系统 (Wu 和 Kabat, J. Exp. Med. 132 (2) :211-250, 1970)。

[0018] 图 2 显示人类 $\gamma 1$ CH2 结构域的氨基酸序列 (SEQ ID NO :5)。指出了在 β -折叠 (...) 和 α -螺旋 (***) 区域中的残基。还显示了环 B-C (这里表示为环 1)、环 D-E (这里表示为环 2)、环 F-G (这里表示为环 3)、环 A-B、环 C-D 和环 E-F 的位置。每个环中的残基以粗体显示。

[0019] 图 3A-3C 是图解了在 CH2 结构域上移植 CDR (或高变环) 的可能策略的示意图。

[0020] 图 4 显示了展现工程化 CH2 结构域的蛋白表达的凝胶图像, 所述工程化 CH2 结构域由箭头指出。

[0021] 图 5A 显示人类 CH2 (NCB 登录号 J00228 ;SEQ ID NO :5) 与小鼠 CH2 (NCB 登录号 J00453 ;SEQ ID NO :92) 的氨基酸序列比对。相同和相似的残基分别是 67% 和 92%。图 5B 是显示人类 CH2 的尺寸排阻色谱的图。插图显示标准曲线。图 5C 是显示 CH2 结构域分子 (以每泳道 1-10 μ g 或 2-5 μ g 的浓度)、单链可变片段 (scFv)、抗体片段 (Fab) 和完整的抗体分子 (IgG) 的分子量的 SDS-PAGE 凝胶的图像。

[0022] 图 6A-6B 是显示通过圆二色性 (CD) 和差示扫描热量法 (DSC) 测量的人类 CH2 的稳定性的图。(A) 通过 CD 测量, 在 25°C 的折叠曲线 (-), 在 90°C 的解折叠 (□□□) 和在 25°C 的重折叠 (---)。将蛋白的折叠级分 (ff) 计算为 $ff = ([\theta] - [\theta_m]) / ([\theta_T] - [\theta_m])$ 。 $[\theta_T]$ 和 $[\theta_m]$ 其中在 25°C 的折叠态和 90°C 的解折叠态的平均残基椭圆率是 216。从对温度 (T) 的一阶导数 $[d(\text{折叠级分})/dT]$ 确定了来自 CD 的精确 T_m 值 ($54.1 \pm 1.2^\circ\text{C}$)。(B) 来自 DSC 的热诱导的解折叠曲线。 $T_m = 55.4^\circ\text{C}$, 其与来自 CD 的 T_m 相似。

[0023] 图 7 是显示基于 CH2 结构的 m01 和 m02 的设计的示意图。在两个天然 Cys 中的两

个 C^as 之间的距离是 **6.53 Å**。这两个 Cys 残基形成了天然二硫键（由黑色箭头指出）。在被半胱氨酸取代的 V10 与 K104(m01) 之间或 L12 与 K104(m02) 之间引入工程化二硫键。

[0024] 图 8 是显示 m01 和 m02 高水平表达的 SDS-PAGE 凝胶的图像。将 m01 和 m02 可溶性表达与 CH2 的可溶性表达对比。表达是由箭头指出。

[0025] 图 9A-9E 是显示通过 CD(A-C)、DSC(D) 和荧光分光光度法 (E) 测量的两种突变体的提高的稳定性的图。显示了 m01(A) 和 m02(B) 的在 25℃ 的折叠曲线 (—)，在 90℃ 的解折叠 (□□□) 和在 25℃ 的重折叠 (---)。(C) 通过与 CH2 相同的方法计算 m01 和 m02 的折叠级分。m01 的 $T_m = 77.4 \pm 1.7^\circ\text{C}$, m02 的 $T_m = 68.6 \pm 0.6^\circ\text{C}$ 。(D) m01 和 m02 的热诱导的解折叠曲线也由 DSC 记录。与 CH2 相比, m01 的 T_m 和 m02 的 T_m 分别提高约 20℃ 和 10℃。(E) 通过荧光分光光度法对比在 CH2、m01 和 m02 之间的尿素诱导的解折叠。CH2、m01 和 m02 的解折叠的中点分别是 4.2M、6.8M 和 5.8M。

[0026] 图 10A 显示 m01 和 m02 的尺寸排阻色谱。如 CH2, m01 只形成单体, 而 m02 主要形成单体并且在较小程度上形成二聚体。图 10B 是显示 N 端平截的 CH2 (CH2) 和平截的 m01 (m01) 的高稳定性的图。前七个 N 端残基被缺失 (SEQ ID NO :5 的残基 1-7)。通过 CD 测量的 50% 解折叠温度 (T_m) (分别为 62℃ 和 79℃) 显著高 (分别高 8℃ 和 5℃) 于相应的 CH2 和 m01 的解折叠温度 (分别为 54℃ 和 74℃)。

[0027] 图 11 是显示 CH2 文库的设计的示意图。显示了 CH2 片段的图示, 其中实心矩形代表环 (L1-L3)。阴影矩形代表面向从环 1 到环 3 的相反方向环 (L) 和螺旋 (H1、H2)。以字母 A-G 标记的空矩形代表形成 β 夹层结构的七条 β 链。数字 231 和 341 代表在 IgG1 的情况下 CH2 片段的起始残基和末端残基。CH2 环 1 (SEQ ID NO :93) 和环 3 (SEQ ID NO :95) 的序列在下面示出, 并加注下划线。引入的突变显示在括号中 (SEQ ID NO :94 和 SEQ ID NO :96)。

[0028] 图 12A-12B 显示 CH2 结合物 (binder) 的表征。(A) 四个 Balgp120-CD4 特异性 CH2 克隆如下实施例所述地被表达并纯化。纯化的产物通过蛋白质印迹分析。样品 1-4 代表来自可溶性分克隆 m1a1 到 m1a3', 而 5-8 从包涵体中复性。(B) CH2 克隆与 Bal gp120-CD4 的结合的 ELISA 分析。

[0029] 图 13A-13B 是显示 CH2 特异性结合的决定簇的电泳凝胶的图和图像。(A) 环 1 决定了结合能力。两个优势克隆 m1a1 和 m1a2 以及含有来自 m1a1 和 m1a3 的环 1 序列而非原始的 CH2 环 3 序列的两个杂合子从包涵体中被表达并纯化, 并且被重折叠 (左图)。然后将这些蛋白用在 ELISA 分析中 (右图)。(B) CH2 为环 1 提供关键性结构支持。携带额外的二硫键的优势克隆 m1a1 及其突变体被表达、纯化以及重折叠 (左图)。然后将它们用在 ELISA 分析中 (右图)。

[0030] 图 14 是显示由 CH2 结合物对 HIV Env 假型病毒感染进行的广泛中和的图。两个 CH2 克隆 m1a1 和 m1a2 以 100 $\mu\text{g/ml}$ 的固定浓度用来测试其抗一组九种 HIV 假病毒的中和能力。浓度为 4 $\mu\text{g/ml}$ 或 6 $\mu\text{g/ml}$ 的 C34 肽用作阳性对照。

[0031] 图 15 显示基于 m1a1 的第二 CH2 文库的设计。来自 CH2 克隆 m1a1 的环 2 (SEQ ID NO :97) 和环 3 (SEQ ID NO :99) 序列 (以下划线标出) 被括号中所示的序列 (SEQ ID NO :98 和 SEQ ID NO :100) 取代。

[0032] 图 16A-16D 显示了选自于第二 CH2 文库的 CH2 克隆的表征。(A) 选自于第二文库

的 CH2 克隆的表达和纯化。(B)m1b3 的凝胶过滤分析。(C)CH2 克隆的 ELISA 分析。(D) 克隆 m1b3 的环 2 和环 3 序列,其与原始 CH2 序列 (SEQ ID NO :97-100) 相比主要具有单体的形式。

[0033] 图 17 是显示由来自第二 CH2 文库的克隆的假病毒中和的图。分析分离自第二文库的三个克隆在 100 μ g/ml 的浓度针对相同组的 HIV 假病毒的中和能力。20 μ g/ml 的浓度的平行纯化的 ScFv X5 用作对照。

[0034] 图 18 是显示识别保守表位的 CH2 结合物的图。主要为单体的 CH2 克隆 m1b3 是生物素标记并且将其在竞争 ELISA 测定中使用。将 ELISA 抗原 Ba1 gp120-CD4 包被在 ELISA 板的底部。将固定量的 1.7 μ M 的生物素酰化的 m1b3 与指示量的未标记的 m1b3、scFv X5 或 m36-Fc 混合并添加到每个孔。用链亲和素 -HRP 检测结合的 m1b3。

[0035] 序列表

[0036] 在附随的序列表中列出的核酸序列和氨基酸序列是使用如在 37C. F. R. 1. 822 中所定义的核苷酸碱基的标准字母缩写和氨基酸的三个字母密码示出。只显示每个核酸序列的一条链,但是将互补链理解为通过任何提及的显示的链而被包括在内。在随附的序列表中:

[0037] SEQ ID NO :1 是人类 V_H 结构域的氨基酸序列。

[0038] SEQ ID NO :2-4 是三种人类抗体的 V_H 结构域的氨基酸序列。

[0039] SEQ ID NO :5 是人类 γ 1CH2 结构域的氨基酸序列。

[0040] SEQ ID NO :6-10 是用于产生突变 CH2 结构域的文库的 PCR 引物的核苷酸序列。

[0041] SEQ ID NO :11-30 是具有随机的环 1 的突变 CH2 结构域的片段的氨基酸序列。

[0042] SEQ ID NO :31-50 是具有随机的环 3 的突变 CH2 结构域的片段的氨基酸序列。

[0043] SEQ ID NO :51-68 是用于将来自人抗体的 CDR3 移入到 CH2 支架中的 PCR 引物的核苷酸序列。

[0044] SEQ ID NO :69-87 是具有移植的 H3 的工程化 CH2 结构域的片段的氨基酸序列。

[0045] SEQ ID NO :88 和 89 是 CH2 结构域突变体 m01 的片段的氨基酸序列。

[0046] SEQ ID NO :90 和 91 是 CH2 结构域突变体 m02 的片段的氨基酸序列。

[0047] SEQ ID NO :92 是鼠 CH2 的氨基酸序列。

[0048] SEQ ID NO :93 是 CH2 环 1 的氨基酸序列。

[0049] SEQ ID NO :94 是突变体 CH2 环 1 的共有氨基酸序列。

[0050] SEQ ID NO :95 是 CH2 环 3 的氨基酸序列。

[0051] SEQ ID NO :96 是突变体 CH2 环 3 的共有氨基酸序列。

[0052] SEQ ID NO :97 是来自克隆 m1a1 的 CH2 环 2 的氨基酸序列。

[0053] SEQ ID NO :98 是来自克隆 m1a1 的突变体 CH2 环 2 的共有氨基酸序列。

[0054] SEQ ID NO :99 是来自克隆 m1a1 的 CH2 环 3 的氨基酸序列。

[0055] SEQ ID NO :100 是来自克隆 m1a1 的突变体 CH2 环 3 的共有氨基酸序列。

[0056] SEQ ID NO :101-105 是用于扩增第一 CH2 文库的 PCR 引物的核苷酸序列。

[0057] SEQ ID NO :106 是 m1a1 合成肽的氨基酸序列。

[0058] SEQ ID NO :107 是 m1a1 环 1 的氨基酸序列。

[0059] SEQ ID NO :108 是 m1a2 环 1 的氨基酸序列。

[0060] 详述

[0061] I. 缩写

[0062] ADCC :抗体依赖性细胞介导的细胞毒性

[0063] CDC :补体依赖性细胞毒性

[0064] CDR :互补决定区

[0065] DNA :脱氧核糖核酸

[0066] ELISA :酶联免疫吸附测定

[0067] HIV :人免疫缺陷病毒

[0068] Ig :免疫球蛋白

[0069] NK :天然杀伤细胞

[0070] RNA :核糖核酸

[0071] SDR :特异性决定残基

[0072] II. 术语

[0073] 除非另外指明,否则技术术语是根据常规用法使用。在分子生物学中常见术语的定义可见于由牛津大学出版社 1994 年出版的 Benjamin Lewin 的 Genes V (基因 V) (ISBN 0-19-854287-9); 由 Blackwell Science Ltd 在 1994 年出版的 Kendrew 等人 (编) The Encyclopedia of Molecular Biology (分子生物学百科全书) (ISBN 0-632-02182-9); 以及由 VCH Publishers, Inc. 在 1995 年出版的 Robert A. Meyers (编), Molecular Biology and Biotechnology: a Comprehensive Desk Reference (分子生物学和生物技术:综合案头参考) (ISBN 1-56081-569-8)。

[0074] 为了促进本发明的各实施方案的叙述,提供以下对具体术语的解释:

[0075] 施用:通过选定的途径将组合物引入到受试者中。例如,如果选定的途径是静脉内,那么通过将组合物引入到受试者的静脉中来施用组合物。

[0076] 动物:有生命的多细胞脊椎动物生物体,包括例如哺乳动物和禽类的分类。术语哺乳动物包括人和非人动物。类似地,术语“受试者”包括人和兽医受试者。

[0077] 抗体:包括基本上由免疫球蛋白基因或免疫球蛋白基因的片段编码的一种或多种多肽的蛋白(或蛋白复合物)。识别的免疫球蛋白基因包括 κ 、 λ 、 α 、 γ 、 δ 、 ϵ 和 μ 恒定区基因以及无数的免疫球蛋白可变区基因。轻链被分类为 κ 或 λ 。重链被分类为 γ 、 μ 、 α 、 δ 或 ϵ , 它们进而分别定义了免疫球蛋白类型: IgG、IgM、IgA、IgD 和 IgE。

[0078] 基本的免疫球蛋白(抗体)结构单元一般为四聚体。每个四聚体包含两对相同的多肽链,每对具有一条“轻”链(约 25kDa)和一条“重”链(约 50-70kDa)。每条链的 N 端定义了主要负责抗原识别的约 100 至 110 个或者以上个氨基酸的可变区。术语“可变的轻链(V_L)”和“可变的重链(V_H)”分别是指这些轻链和重链。每条轻链含有单个恒定结构域(CL),而每条重链含有三个恒定结构域 CH1、CH2 和 CH3(或者对于 IgE 和 IgM 而言四个恒定结构域)。参见图 1A 的常规免疫球蛋白分子的示意图。

[0079] 如在此所用,术语“抗体”包括完整的免疫球蛋白以及一些具有约 25 至 100kD 分子量的良好表征的片段。例如,与靶蛋白(或者与蛋白或融合蛋白内的表位)结合的 Fab、Fv 和单链 Fv(scFv)还将是该蛋白(或表位)的特异性结合剂。这些抗体片段被定义如下:(1) Fab, 含有抗体分子的单价抗原结合片段的片段,它是通过用木瓜蛋白酶消化整个抗

体得到完整的轻链和一条重链的一部分而产生；(2) Fab'，通过用胃蛋白酶处理整个抗体，随后还原得到完整的轻链和重链的一部分而获得的抗体分子的片段；每个抗体分子获得两个 Fab' 片段；(3) (Fab')₂，通过用胃蛋白酶处理整个抗体且没有随后的还原所获得的抗体的片段；(4) F(ab')₂，通过两个二硫键连在一起的两个 Fab' 片段的二聚体；(5) Fv，含有轻链的可变区和重链的可变区的表达为两条链的基因工程化片段；以及 (6) scFv，单链抗体，含有轻链的可变区、重链的可变区的基因工程化分子，所述轻链的可变区和重链的可变区通过作为基因融合单链分子的适合多肽接头相连接。制备这些片段的方法是常规的（参见例如，Harlow 和 Lane, Using Antibodies: A Laboratory Manual (使用抗体：实验室手册)，CSHL, New York, 1999)。

[0080] 抗体可以是单克隆或多克隆的。仅通过举例，单克隆抗体可以根据 Kohler 和 Milstein 的经典方法 (Nature 256:495-97, 1975) 或其衍生方法从鼠杂交瘤中制备。例如，用于单克隆抗体生产的详细步骤由 Harlow 和 Lane 描述 (Using Antibodies: A Laboratory Manual (使用抗体：实验室手册)，CSHL, New York, 1999)。

[0081] “人源化”免疫球蛋白例如人源化抗体是包括人框架区和来自非人类（例如小鼠、大鼠或合成的）免疫球蛋白的一种或多种 CDR 的免疫球蛋白。提供 CDR 的非人类免疫球蛋白命名为“供体”，而提供框架的人免疫球蛋白命名为“受体”。在一个实施方案中，所有的 CDR 是来自人源化免疫球蛋白中的供体免疫球蛋白。“人源化抗体”是包含人源化轻链和人源化重链免疫球蛋白的抗体，例如人源化单克隆抗体。人源化抗体结合至与提供 CDR 的供体抗体相同或相似的抗原上。人源化免疫球蛋白的受体框架可具有限制数目的取自供体框架的氨基酸取代。人源化分子可以具有额外的保守氨基酸取代，所述保守氨基酸取代对抗原结合或其他免疫球蛋白功能实质上没有影响。这些分子可以通过基因工程化手段构建（例如参见美国专利 5,585,089）。

[0082] 抗原：能够刺激动物中抗体或 T 细胞应答的产生的化合物、组合物或物质，包括被注射或吸收到动物中的组合物。抗原与特异性体液免疫或特异性细胞免疫的产物进行反应。

[0083] 自身免疫疾病：其中因组织损伤而起的免疫系统产生针对为正常宿主的一部分的抗原（也就是自身抗原）的免疫应答（例如，B 细胞应答或 T 细胞应答）的疾病。自身抗原可以获自宿主细胞，或者可以获自正常定殖粘膜表面的共栖生物，例如微生物（称为共栖生物）。

[0084] 影响哺乳动物的示例性自身免疫疾病包括类风湿性关节炎、幼年型少关节型关节炎 (juvenile oligoarthritis)、胶原诱导性关节炎、佐剂诱导性关节炎、干燥综合征、多发性硬化症、实验性自身免疫性脑脊髓炎、炎性肠病（例如克罗恩病、溃疡性结肠炎）、自身免疫性胃萎缩、寻常天疱疮、牛皮癣、白癜风、1 型糖尿病、非肥胖型糖尿病、重症肌无力、Grave 病、桥本氏甲状腺炎、硬化性胆管炎、硬化性涎腺炎、系统性红斑狼疮、自身免疫性血小板减少性紫癜、肺出血 - 肾炎综合征 (Goodpasture's syndrome)、Addison 病、全身性硬化症、多发性肌炎、皮肌炎、自身免疫性溶血性贫血、恶性贫血以及类似疾病。

[0085] 结合亲和力：在结合位点与配体之间（例如，抗体、CH2 结构域或 CH3 结构域与抗原或表位之间）的结合的力。结合位点 X 对配体 Y 的亲和力由解离常数 (K_d) 表示，解离常数是占据溶液中存在的半数 X 的结合位点所需的 Y 的浓度。较低的 (K_d) 表明在 X 与 Y 之

间较强或较高的亲和力相互作用,并且需要较低浓度的配体占据这些位点。一般来说,结合亲和力可以受由互补位(识别表位的分子的部分)所识别的表位中一个或多个氨基酸的改变、修饰和/或取代影响。结合亲和力可以是抗体结合抗原的亲和力。

[0086] 在一个实例中,结合亲和力是通过 Ag-ELISA 测定中的终点滴定所测量。如果修饰的/置换的表位的特异性抗体的终点滴度与未改变的表位相比相差至少 4 倍,如至少 10 倍,至少 100 倍或更大,那么结合亲和力大体上是通过由抗体互补位所识别的表位中的一个或多个氨基酸的修饰和/或取代来降低(或可测量地减少)。

[0087] CH2 或 CH3 结构域分子:获自免疫球蛋白 CH2 或 CH2 结构域的多肽(或编码多肽的核酸)。免疫球蛋白可以是 IgG、IgA、IgD、IgE 或 IgM。在本文所述的一个实施方案中,CH2 或 CH3 结构域分子包含至少一个 CDR 或其功能片段。CH2 或 CH3 结构域分子还可以包含额外的氨基酸序列,例如完整的高变环。在另一个实施方案中,CH2 或 CH3 结构域分子至少具有以 CDR 或其功能片段置换的一个或多个环区域的至少一部分。在本文所述的一些实施方案中,CH2 或 CH3 结构域在该分子的环区域中包含一个或多个突变。CH2 或 CH3 结构域的“环区域”是指位于 β -折叠的区域之间的蛋白部分(例如,每个 CH2 结构域包含七个 β -折叠 A 至 G,方向从 N 端到 C 端)。如在图 3A-3C 中所示,CH2 结构域包含六个环区域:环 1、环 2、环 3、环 A-B、环 C-D 和环 E-F。环 A-B、环 C-D 和环 E-F 分别位于 β -折叠 A 与 B 之间、C 与 D 之间、以及 E 与 F 之间。环 1、2 和 3 分别位于 β -折叠 B 与 C 之间、D 与 E 之间、以及 F 与 G 之间。对于 CH2 结构域中的环的氨基酸范围参见表 1。本文所公开的 CH2 和 CH3 结构域分子还可以包含 N 端缺失,例如约 1 个至约 7 个氨基酸的缺失。在具体的实例中,N 端缺失是 1、2、3、4、5、6 或 7 个氨基酸长度。本文所公开的 CH2 和 CH3 结构域分子还可以包含 C 端缺失,例如约 1 个至约 4 个氨基酸的缺失。在具体的实例中,C 端缺失是 1、2、3 或 4 个氨基酸长度。

[0088] CH2 和 CH3 结构域分子的尺寸较小,通常小于 15kD。CH2 和 CH3 结构域分子的尺寸可以改变,这取决于在环区域中插入的 CDR/高变氨基酸序列的长度、插入 CDR 的数量以及另一分子(如效应物分子或标记)是否与 CH2 或 CH3 结构域结合。在一些实施方案中,CH2 和 CH3 结构域分子不包含额外的恒定结构域(即 CH1 结构域或另一 CH2 结构域或 CH3 结构域)或可变结构域。在一个实施方案中,CH2 结构域是来自 IgG、IgA 或 IgD。在另一实施方案中,恒定结构域是与 IgG、IgA 或 IgD 的 CH2 结构域同源的来自 IgE 或 IgM 的 CH3 结构域。

[0089] 本文提供的 CH2 和 CH3 结构域分子可以是糖基化或非糖基化的。例如,可以将重组 CH2 结构域或 CH3 结构域表达在适当的哺乳动物细胞中以允许该分子的糖基化。

[0090] 互补决定区(CDR):在抗原受体(如免疫球蛋白和 T 细胞受体)蛋白的可变结构域中发现的短氨基酸序列,其为该受体提供对抗原的接触位点及其对特定抗原的特异性。抗原受体的每个多肽链含有三个 CDR(CDR1、CDR2 和 CDR3)。抗原受体通常是由两个多肽链(重链和轻链)组成,因此每个抗原受体具有六个能够与抗原接触的 CDR。因为与抗原受体相关的大多数序列变异在 CDR 中被发现,所以这些区域有时被称为高变结构域。

[0091] CDR 在抗原受体的环区域内(通常在 β -折叠结构的区域之间;参见图 3A-3C)被发现。这些环区域通常称为高变环。每个抗原受体包含六个高变环:H1、H2、H3、L1、L2 和 L3。例如,H1 环包含重链的 CDR1,并且 L3 环包含轻链的 CDR3。本文所述的 CH2 和 CH3 结构

域分子包含来自抗体可变结构域的移入的氨基酸。移入的氨基酸包含 CDR 的至少一部分。移入的氨基酸还可以包括额外的氨基酸序列,如完整的高变环。如本文所用,CDR 的“功能片段”是保留了结合特异性抗原的能力的 CDR 的至少一部分。

[0092] CDR 位置的编号规则由 Kabat 等人., (1991) Sequences of Proteins of Immunological Interest (免疫学目的蛋白的序列), 第 5 版, 美国卫生与公共服务部, 公共卫生服务, 国立卫生研究院, Bethesda, MD (NIH 公开号 91-3242) 描述。

[0093] 接触: 以直接物理关联来放置, 其包括以固体形式和以液体形式二者。

[0094] 简并变体: 如本文所述, CH2 或 CH3 结构域分子的“简并变体”是编码 CH2 或 CH3 结构域分子的多核苷酸, 该多核苷酸包括由于遗传密码简并的序列。存在 20 种天然氨基酸, 它们大多数被多于一个密码子指定。因此, 所有的简并核苷酸序列被包括在内, 只要由该核苷酸序列所编码的 CH2 或 CH3 结构域分子的氨基酸序列是不变的。

[0095] 结构域: 一种蛋白质结构, 该结构保留了独立于蛋白质的其余部分的三级结构。在一些情况下, 结构域具有分离的功能特性并且可被添加、移除或转移至另一蛋白而没有功能损失。

[0096] 效应物分子: 预期对分子或嵌合分子所靶向的细胞具有期望的作用的分子或嵌合分子的部分。效应物分子也称为效应部分 (EM)、治疗剂或诊断剂或类似术语。

[0097] 治疗剂包括这类化合物, 如核酸、蛋白、肽、氨基酸或衍生物、糖蛋白、放射性同位素、脂质、碳水化合物或重组病毒。核酸治疗部分和诊断部分包括反义核酸、与单链或双链 DNA 共价交联的衍生的寡核苷酸和形成三股螺旋的寡核苷酸。可选地, 与靶向部分相连的分子如 CH2 或 CH3 结构域分子可以是含有治疗组合物的囊封系统, 如脂质体或胶束, 所述治疗组合物例如药物、核酸 (如反义核酸) 或能够被屏蔽免于直接暴露于循环系统的另外的治疗部分。制备连接于抗体的脂质体的方法是本领域技术人员公知的。参见例如, 美国专利第 4, 957, 735 号; 以及 Connor 等人., Pharm. Ther. 28 :341-365, 1985。诊断剂或诊断部分包括放射性同位素或其他可检测的标记。用于此类目的的可检测的标记也是本领域公知的, 并且包括放射性同位素 (如 ^{32}P 、 ^{125}I 和 ^{131}I)、荧光团、化学发光剂和酶。

[0098] 表位: 抗原决定簇。是在有抗原性的分子 (即该分子引发特异性免疫应答) 上的特定化学基团, 或者连续或不连续的肽序列。抗体结合特定的抗原表位是基于该抗体的三维结构和匹配 (或同源) 的表位。

[0099] 表达: 核酸翻译为蛋白质。蛋白质可以被表达并留在细胞内, 成为细胞表面膜的组分, 或者被分泌到细胞外基质或介质中。

[0100] 表达控制序列: 调控异源核酸序列的表达的核酸序列, 所述异源核酸序列与之可操作地连接。当表达控制序列控制并调控核酸序列的转录和 (适当时) 翻译时, 该表达控制序列与该核酸序列可操作地连接。因此表达控制序列可包括适当的启动子、增强子、转录终止子、在编码蛋白的基因前面的起始密码子 (即 ATG)、内含子的剪接信号、允许 mRNA 的适当翻译的该基因的正确可读框的维持序列以及终止密码子。术语“控制序列”旨在最少包括其存在能够影响表达的组分, 并且还包括其存在有利的额外组分, 例如前导序列和融合配偶体序列。表达控制序列可包括启动子。

[0101] 启动子是指导核酸转录的一系列核酸控制序列。启动子包括转录起始位点附近的必要核酸序列, 如在聚合酶 II 型启动子情况下的 TATA 元件。启动子还任选地包括远端增

强子或阻抑物元件,它们可位于距离转录起始位点多达数千碱基对处。包括组成型启动子和诱导型启动子二者(参见例如,Bitter 等人,Methods in Enzymology(酶学方法)153:516-544,1987)。

[0102] 还包括足以赋予启动子依赖性基因表达的那些启动子元件,所述启动子依赖性基因表达是细胞类型特异性、组织特异性可控的,或者通过外部信号或试剂可诱导的;此类元件可位于基因的 5' 或 3' 区域中。包括组成型启动子和诱导型启动子二者(参见例如,Bitter 等人,Methods in Enzymology(酶学方法)153:516-544,1987)。例如,当在细菌系统中克隆时,可使用诱导型启动子,如细菌噬菌体 λ 的 pL、plac、ptrp、ptac(ptrp-lac 杂合启动子)以及类似启动子。在一个实施方案中,当在哺乳动物细胞系统中克隆时,可使用获自哺乳动物细胞的基因组的启动子(如金属硫蛋白启动子)或来自哺乳动物病毒的启动子(如逆转录病毒长末端重复序列;腺病毒后期启动子;痘苗病毒 7.5K 启动子)。通过重组 DNA 或合成技术产生的启动子也可以用来提供核酸序列的转录。

[0103] 可将多核苷酸插入到含有启动子序列的表达载体中,所述启动子序列促进宿主的插入基因序列的有效转录。表达载体通常含有复制起点、启动子以及允许对转化细胞的表型选择的特异性核酸序列。

[0104] 框架区:在 CDR(或高变区)之间插入的氨基酸序列。框架区包括可变的轻链框架区和可变的重链框架区。每个可变结构域包含四个框架区,通常称为 R1、FR2、FR3 和 FR4。框架区用来将 CDR 保持在适当的方向以用于抗原结合。框架区通常形成 β -折叠结构。

[0105] 真菌相关抗原(FAA):能够刺激真菌特异性 T 细胞限定的免疫应答的真菌抗原。示例性 FAA 包括但不限于来自以下的抗原:白色念珠菌(*Candida albicans*)、隐球菌属(*Cryptococcus*)(如来自新型隐球菌(*C. neoformans*)的 d25 或 MP98 或 MP88,或者其免疫片段)、芽生菌属(*Blastomyces*)(如皮炎芽生菌(*B. dermatitidis*),例如 WI-1 或其免疫片段)和组织胞浆菌属(*Histoplasma*)(如荚膜组织胞浆菌(*H. capsulatum*))。

[0106] 异源的:异源的多肽或多核苷酸是指获自不同来源或物种的多肽或多核苷酸。

[0107] 高变区:在抗体可变结构域内的序列变异性特别高的区域。高变区形成框架区的 β -折叠之间的环结构。因此,高变区也称为“高变环”。每个可变结构域包含三个高变区,在重链中通常称为 H1、H2 和 H3,而在轻链中通常称为 L1、L2 和 L3。高变环的环结构描绘在图 3A-5C 中。

[0108] 免疫应答:免疫系统的细胞如 B 细胞、T 细胞、巨噬细胞或多形核细胞(polymorphonucleocyte)对诸如抗原的刺激物的反应。免疫应答可包括参与宿主防御反应的身体的任何细胞,例如分泌干扰素或细胞因子的上皮细胞。免疫应答包括但不限于先天免疫应答或炎症。

[0109] 免疫结合物:效应物分子与抗体或者与 CH2 或 CH3 结构域分子的共价连接。效应物分子可以是可检测的标记或免疫毒素。毒素的具体、非限制性实例包括但不限于:相思豆毒蛋白,蓖麻毒蛋白,假单胞菌外毒素(PE,如 PE35、PE37、PE38 和 PE40)、白喉毒素(DT),肉毒杆菌毒素或其改性的毒素,或者直接或间接地抑制细胞生长或杀死细胞的其他毒剂。例如,PE 和 DT 是通常通过肝毒性导致死亡的高毒性化合物。然而,可通过如下方式将 PE 和 DT 改性成为用作免疫毒素的形式:移除毒素的天然靶向组分(如 PE 的结构域 Ia 和 DT 的 B 链)并且用不同的靶向部分(如 CH2 或 CH3 结构域分子)取代该天然靶向组分。在一个

实施方案中,将 CH2 或 CH3 结构域分子与效应物分子 (EM) 相接合。在另一个实施方案中,与效应物分子相接合的 CH2 或 CH3 结构域分子还与脂质或蛋白或肽的其他分子相连接以提高 CH2 或 CH3 结构域分子在体内的半衰期。该连接可以通过化学手段或重组手段。“化学手段”是指在 CH2 或 CH3 结构域分子与效应物分子之间的反应,以使得在这两个分子之间形成共价键从而形成一个分子。在 CH2 或 CH3 结构域分子与效应物分子之间可以任选地包括肽接头(短肽序列)。因为免疫结合物最初是由具有不同官能性的两个分子,如抗体和效应物分子制备,所以它们有时被称为“嵌合分子”。本文所用的术语“嵌合分子”因此是指结合(偶联)至效应物分子的靶向部分,如配体、抗体或者 CH2 或 CH3 结构域分子。

[0110] 术语“结合”、“接合”、“键合”或“连接”是指使两个多肽成为一个连续多肽分子,或者是指将放射性核苷酸或其他分子与多肽,如 CH2 或 CH3 结构域分子共价连接。在具体的上下文中,所述术语包括指配体,如抗体部分与效应物分子(“EM”)相接合。

[0111] 免疫原:能够在适当条件下刺激动物的免疫应答,如抗体或 T 细胞免疫应答的产生的化合物、组合物或物质,包括被注射或吸收到动物中的组合物。

[0112] 分离的:“分离的”生物组分(如核酸分子或蛋白)已从该组分天然存在的其他生物组分(例如,细胞的其他生物组分)中基本上分离或纯化出来,所述其他生物组分例如其其他的染色体的和染色体外的 DNA 和 RNA 以及蛋白,包括其他抗体。已被“分离的”核酸和蛋白包括通过标准的纯化方法纯化的核酸和蛋白。“分离的抗体”是其已从其他蛋白或生物组分基本上分离或纯化出来以使得保持其抗原特异性的抗体。该术语还包括通过宿主细胞中的重组表达制备的核酸和蛋白(包括 CH2 和 CH3 结构域分子)以及化学合成的核酸或蛋白或者其片段。

[0113] 标记:直接或间接地与另外的分子如抗体或者 CH2 或 CH3 结构域分子结合以促进该分子的检测的可检测的化合物或组合物。标记的具体、非限制性实例包括荧光标签、酶连接和放射性同位素。

[0114] 配体接触残基或特异性决定残基(SDR):在 CDR 内的参与接触配体或抗原的残基。配体接触残基也被称为特异性决定残基(SDR)。非配体接触残基是 CDR 中不接触配体的残基。非配体接触残基也可以是框架残基。

[0115] 纳米抗体(nAb):被工程化以使得分子特异性结合抗原的 CH2 或 CH3 结构域分子。被工程化以结合抗原的 CH2 和 CH3 结构域分子是基于抗原特异性结合抗体结构域的已知的最小分子。

[0116] 瘤形成和肿瘤:瘤形成的产物是瘤(肿瘤),它是由细胞过度分裂产生的组织异常生长。瘤也被称为“癌症”。不转移的肿瘤也被称为“良性的”。入侵周围组织和/或能够转移的肿瘤也被称为“恶性的”。

[0117] 实体肿瘤,如肉瘤和癌的实例包括纤维肉瘤、粘液肉瘤、脂肪肉瘤、软骨肉瘤、成骨性肉瘤、以及其他肉瘤、滑膜瘤、间皮瘤、尤因氏瘤、平滑肌肉瘤、横纹肌肉瘤、结肠癌、淋巴样恶性肿瘤、胰腺癌、乳腺癌、肺癌、卵巢癌、前列腺癌、肝细胞癌、鳞状细胞癌、基底细胞癌、腺癌、汗腺癌、皮脂腺癌、乳头状癌、乳头状腺癌、髓样癌、支气管原癌、肾细胞癌、肝癌、胆管癌、绒毛膜癌、维尔姆斯氏肿瘤、子宫颈癌、睾丸癌、膀胱癌和中枢神经系统肿瘤(如神经胶质瘤、星形细胞瘤、髓母细胞瘤、颅咽管瘤(craniopharyngioma)、室管膜瘤、松果体瘤、血管母细胞瘤、听神经瘤、少突神经胶质瘤、脑膜瘤(meningioma)、黑色素瘤、成神经细胞瘤和视

网膜母细胞瘤)。

[0118] 血液肿瘤的实例包括白血病,白血病包括急性白血病(如急性淋巴细胞性白血病、急性粒细胞性白血病、急性骨髓性白血病和成髓细胞性白血病、前髓细胞性白血病、骨髓单核细胞性白血病、单核细胞性及红白血病);慢性白血病(如慢性髓细胞性(粒细胞性)白血病、慢性骨髓性白血病以及慢性淋巴细胞性白血病);真性红细胞增多症;淋巴瘤;霍奇金病;非霍奇金淋巴瘤(无痛形式和高级形式);多发性骨髓瘤;沃尔丹斯特伦氏巨球蛋白血症;重链病;骨髓增生异常综合症;多毛细胞白血病;以及骨髓发育不良。

[0119] 核酸:由通过磷酸二酯键连接的核苷酸单元(核糖核苷酸、脱氧核糖核苷酸、相关的天然存在的结构变体及其合成的非天然存在的类似物)、相关的天然存在的结构变体及其合成的非天然存在的类似物构成的聚合物。因此,该术语包括如下核苷酸聚合物:其中核苷酸与它们之间的连接包括非天然存在的合成类似物,诸如例如且不限于:硫代磷酸酯、氨基磷酸酯、甲基膦酸酯、手性甲基膦酸酯、2'-O-甲基核糖核苷酸、肽-核酸(PNA)以及类似物。此类多核苷酸可以例如使用自动化DNA合成仪来合成。术语“寡核苷酸”通常是指不超过约50个核苷酸的短的多核苷酸。应理解当核苷酸序列由DNA序列(即A、T、G、C)表示时,这还包括RNA序列(即A、U、G、C),在RNA序列中“U”代替“T”。

[0120] 本文使用常规符号来描述核苷酸序列:单链核苷酸序列的左手末端是5'端;双链核苷酸序列的左手方向称为5'方向。将核苷酸以5'到3'添加到新生RNA转录物的方向也称为转录方向。具有与mRNA相同的序列的DNA链称为“编码链”;在具有与从该DNA转录的mRNA相同的序列的DNA链上并且被定位在RNA转录物的5'到5'端的序列称为“上游序列”;在具有与mRNA相同的序列的DNA链上并且是编码RNA转录物的3'到3'端的序列称为“下游序列”。

[0121] “cDNA”是指与mRNA互补或相同的单链或双链形式的DNA。

[0122] “编码”是指在多核苷酸,如基因、cDNA或mRNA中的核苷酸的特定序列充当合成生物过程中的其他聚合物或大分子的模板的固有特性以及从中产生的生物特性,所述其他聚合物或大分子具有确定的核苷酸序列(即rRNA、tRNA和mRNA)或确定的氨基酸序列。因此,如果由基因产生的mRNA的转录和翻译在细胞或其他生物系统中产生了蛋白,那么该基因编码该蛋白。用作基因或cDNA的转录模板的编码链(其核苷酸序列与mRNA序列相同并且通常被提供在序列表中)和非编码链二者可称为编码该基因或cDNA的蛋白或其他产物。除非另外指定,否则“编码氨基酸序列的核苷酸序列”包括为彼此的兼并形式并且编码相同的氨基酸序列的所有核苷酸序列。编码蛋白和RNA的核苷酸序列可包括内含子。

[0123] “重组核酸”是指具有如下核苷酸序列的核酸:所述核苷酸序列并非天然接合在一起并且可通过人工组合两个原本分离的序列区段而制备。人工组合通常是通过化学合成,或更为常见地通过人工操作分离的核酸片段(例如通过基因工程技术)来实现。重组核酸包括核酸载体,所述核酸载体包含可用来转化适合的宿主细胞的扩增的或装配的核酸。包含重组核酸的宿主细胞称为“重组宿主细胞”。然后将该基因表达在重组宿主细胞中以产生“重组多肽”。重组核酸也可起非编码功能的作用(例如,启动子、复制起点、核糖体结合位点以及类似物)。

[0124] 可操作地连接:当第一核酸序列以功能关系与第二核酸序列一起放置时,该第一核酸序列与该第二核酸序列可操作连接。例如,如果启动子影响编码序列的转录或表达,那

么该启动子与该编码序列可操作地连接。一般来说,可操作地连接的 DNA 序列是连续的,并且如果有必要接合两个蛋白编码区,则可操作地连接的 DNA 序列是在相同的阅读框内。

[0125] 病原体:引起宿主的疾病或疾患的生物原。病原体包括例如细菌、病毒、真菌、原生动物和寄生虫。病原体也称为传染原。

[0126] 致病病毒的实例包括以下病毒科中的病毒:逆转录病毒科 (Retroviridae) (例如人免疫缺陷病毒 (HIV);人 T 细胞性白血病病毒 (HTLV);小 RNA 病毒科 (Picornaviridae) (例如脊髓灰质炎病毒、甲型肝炎病毒、丙型肝炎病毒、肠道病毒、人柯萨奇病毒、鼻病毒、艾可病毒、口蹄疫病毒);杯状病毒科 (Caliciviridae) (例如引起胃肠炎的毒株);披膜病毒科 (Togaviridae) (例如马脑炎、风疹病毒);黄病毒科 (Flaviridae) (例如登革病毒、黄病毒、西尼罗病毒、圣路易斯脑炎病毒、乙型脑炎病毒以及其他脑炎病毒);冠状病毒科 (Coronaviridae) (例如冠状病毒、严重急性呼吸窘迫综合征 (SARS) 病毒);弹状病毒科 (Rhabdoviridae) (例如疱疹性口炎病毒、狂犬病毒);纤丝病毒科 (Filoviridae) (例如埃博拉病毒);副粘病毒科 (Paramyxoviridae) (例如副流感病毒、腮腺炎病毒、麻疹病毒、呼吸道合胞病毒 (RSV));正粘病毒科 (Orthomyxoviridae) (例如流感病毒);布尼亚病毒科 (Bunyaviridae) (例如汉坦病毒、辛诺瓦病毒、立夫特山谷热病毒、布尼亚病毒、静脉病毒和内罗病毒);沙粒病毒科 (Arena viridae) (出血热病毒、马秋波病毒、鸠宁病毒);呼肠孤病毒科 (Reoviridae) (例如呼肠孤病毒、环状病毒和轮状病毒);双 RNA 病毒科 (Birnaviridae) 嗜肝 DNA 病毒科 (Hepadnaviridae) (型肝炎病毒);细小病毒科 (Parvoviridae) (细小病毒);乳多空病毒科 (Papovaviridae) (乳头瘤病毒、多瘤病毒、BK 病毒);腺病毒科 (Adenoviridae) (大部分腺病毒);疱疹病毒科 (Herpesviridae) (单纯疱疹病毒 (HSV)-1 和 HSV-2;巨细胞病毒 (CMV);EB 病毒 (EBV);水痘带状疱疹病毒 (VZV);以及其他疱疹病毒,包括 HSV-6);痘病毒科 (Poxviridae) (天花病毒、牛痘病毒、痘病毒);和虹彩病毒科 (Iridoviridae) (如非洲猪瘟病毒);纤丝病毒科 (Filoviridae) (例如埃博拉病毒、马尔堡病毒);杯状病毒科 (Caliciviridae) (例如诺瓦克病毒) 和未分类病毒 (例如海绵状脑病的病原、 δ 肝炎的病原 (被认为是乙型肝炎缺陷型卫星病毒);以及星状病毒)。

[0127] 真菌病原体的实例包括但不限于:新型隐球菌 (*Cryptococcus neoformans*)、荚膜组织胞浆菌 (*Histoplasma capsulatum*)、粗球孢子菌 (*Coccidioides immitis*)、皮炎芽生菌 (*Blastomyces dermatitidis*)、沙眼衣原体 (*Chlamydia trachomatis*)、白色念珠菌 (*Candida albicans*)。

[0128] 细菌病原体的实例包括但不限于:幽门螺杆菌 (*Helicobacter pylori*)、伯氏疏螺旋体 (*Borrelia burgdorferi*)、嗜肺军团菌 (*Legionella pneumophila*)、分枝杆菌 (*Mycobacteria* spp) (如结核分枝杆菌 (*M. tuberculosis*)、禽分枝杆菌 (*M. avium*)、胞内分枝杆菌 (*M. intracellulare*)、堪萨斯分枝杆菌 (*M. kansasii*)、戈登分枝杆菌 (*M. goodii*)、金黄色葡萄球菌 (*Staphylococcus aureus*)、淋病奈瑟菌 (*Neisseria gonorrhoeae*)、脑膜炎奈瑟菌 (*Neisseria meningitidis*)、单核细胞增生李斯特菌 (*Listeria monocytogenes*)、化脓链球菌 (*Streptococcus pyogenes*) (A 群链球菌)、无乳链球菌 (*Streptococcus agalactiae*) (B 群链球菌)、链球菌 (*Streptococcus*) (草绿色种群)、粪链球菌 (*Streptococcus faecalis*)、牛链球菌 (*Streptococcus*

bovis)、链球菌(厌氧种)、肺炎链球菌(*Streptococcus pneumoniae*)、致病性弯曲杆菌(pathogenic *Campylobacter* sp.)、肠道球菌(*Enterococcus* sp.)、流感嗜血菌(*Haemophilus influenzae*)、炭疽芽孢杆菌(*Bacillus anthracis*)、白喉棒杆菌(*Corynebacterium diphtheriae*)、棒状杆菌(*Corynebacterium* sp.)、猪红斑丹毒丝菌(*Erysipelothrix rhusiopathiae*)、产气荚膜梭菌(*Clostridium perfringens*)、破伤风梭菌(*Clostridium tetani*)、产气肠杆菌(*Enterobacter aerogenes*)、肺炎克雷伯氏杆菌(*Enterobacter aerogenes*)、多杀性巴氏杆菌(*Pasturella multocida*)、拟杆菌(*Bacteroides* sp.)、具核梭杆菌(*Fusobacterium nucleatum*)、念珠状链杆菌(*Streptobacillus moniliformis*)、苍白密螺旋体(*Treponema pallidum*)、极细密螺旋体(*Treponema pertense*)、钩端螺旋体属(*Leptospira*)以及以色列放线菌(*Actinomyces israelii*)。

[0129] 其他的病原体(如原生生物)包括:恶性疟原虫(*Plasmodium falciparum*)和刚地弓形体(*Toxoplasma gondii*)。

[0130] 药学上可接受的介质:在本公开内容中有用的药学上可接受的载体(介质)是常规的。Remington's Pharmaceutical Sciences(雷明顿药物科学), E. W. Martin, Mack 出版公司, Easton, PA, 第 15 版(1975)描述了适合于一种或多种治疗性化合物或分子(如一种或多种抗体)和其他药剂的药物递送的组合物和制剂。

[0131] 一般来说,载体的种类将取决于采用的具体施用方式。例如,肠胃外制剂通常包含可注射液体,所述可注射液体包括作为介质的药学上和生理学上可接受的流体,如水、生理盐水、平衡盐溶液、右旋糖水溶液、甘油等。对于固体组合物(例如,粉末、丸剂、片剂或胶囊形式)来说,常规的无毒固体载体可包括例如:药用级的甘露醇、乳糖、淀粉或硬脂酸镁。除了生物学上中性的载体以外,待施用的药物组合物可以含有少量的无毒辅助性物质,如湿润剂或乳化剂、防腐剂和 pH 缓冲剂等,例如乙酸钠或脱水山梨糖醇单月桂酸酯。

[0132] 多肽:一种聚合物,其中单体为通过酰胺键接合在一起的氨基酸残基。当氨基酸是 α 氨基酸时,可使用 L 型光学异构体或 D 型光学异构体。如本文所用的术语“多肽”或“蛋白质”旨在涵盖任何氨基酸序列并且包括修饰的序列,如糖蛋白。术语“多肽”具体旨在覆盖天然存在的蛋白质以及重组或合成产生的蛋白质。

[0133] 术语“残基”或“氨基酸残基”包括所指的被掺入到蛋白、多肽或肽中的氨基酸。

[0134] “保守的”氨基酸取代是基本上不影响或者降低多肽的活性或抗原性的那些取代。例如,多肽可以包括至多约 1 个、至多约 2 个、至多约 5 个、至多约 10 个或至多约 15 个保守取代,并且其特异性结合与原多肽结合的抗体。术语保守的变异还包括使用取代的氨基酸置换未取代的母体氨基酸,条件是针对取代的多肽所产生的抗体产生的抗体(antibodies raised against antibodies raised)也与未取代的多肽免疫反应。保守性取代的实例在下面显示。

	原始残基	保守性置换
	Ala	Ser
	Arg	Lys
	Asn	Gln, His
	Asp	Glu
	Cys	Ser
	Gln	Asn
	Glu	Asp
	His	Asn; Gln
[0135]	Ile	Leu, Val
	Leu	Ile; Val
	Lys	Arg; Gln; Glu
	Met	Leu; Ile
	Phe	Met; Leu; Tyr
	Ser	Thr
	Thr	Ser
	Trp	Tyr
	Tyr	Trp; Phe
	Val	Ile; Leu

[0136] 保守性取代一般保持 (a) 取代区域中多肽骨架的结构, 例如折叠或螺旋构象; (b) 在靶点的分子的电荷或疏水性; 和 / 或 (c) 侧链的大小 (bulk)。

[0137] 一般预期产生蛋白质特性的最大改变的取代将是非保守的, 例如如下的改变: (a) 亲水性残基例如丝氨酰基或苏氨酰基取代 (或被取代为) 疏水性残基例如亮氨酰基、异亮氨酰基、苯丙氨酰基、缬氨酰基或丙氨酰基; (b) 半胱氨酸或脯氨酸取代 (或被取代为) 任何其他残基; (c) 具有正电性侧链的残基例如赖氨酰基、精氨酰基或组氨酰基取代 (或被取代为) 负电性残基例如谷氨酰基或天冬氨酰基; 或者 (d) 具有大侧链的残基例如苯丙氨酸取代 (或被取代为) 不具有侧链的残基例如甘氨酸。

[0138] 预防、治疗或改善疾病: “预防”疾病是指抑制疾病的完全发展。“治疗”是指在疾病开始发展后改善疾病或者病理状态的体征或症状的治疗干预。“改善”是指降低疾病的体征或症状的数量或严重程度。

[0139] 探针和引物: 探针包括连接至可检测的标记或报道分子的分离的核酸。引物是短核酸并且可以是长度为 15 个核苷酸或更长的 DNA 寡核苷酸。引物可以通过核酸杂交退火

至互补靶 DNA 链以在该引物与靶 DNA 链之间形成杂交体,并且然后其通过 DNA 聚合物酶沿着靶 DNA 链延伸。引物对可以用于扩增核酸序列,例如通过聚合酶链式反应 (PCR) 或本领域已知的其他核酸扩增方法。本领域技术人员应理解具体的探针或引物的特异性随其长度的增加而增加。因此,例如,包含 20 个保守核苷酸的引物将以比仅 15 个核苷酸的对应引物更高的特异性退火至靶。因此,为了获得更大特异性,可选择包含 20、25、30、35、40、50 或更多个连续核苷酸的探针和引物。

[0140] 纯化的:术语纯化的不需要绝对的纯度;反之,它旨在作为一个相对的术语。因此,例如,纯化的 CH2 或 CH3 结构域分子是整体或部分地从自然相关蛋白和其他污染物中分离的分子,其中该分子相对于其天然存在状态,例如,相对于其在细胞提取物或生物流体中的纯度,可被纯化至相当的程度。

[0141] 术语“纯化的”包括期望的产物,如类似物或模拟物或其他生物活性化合物,其中使其他化合物或部分与 CH2 或 CH3 结构域分子相结合以允许其他化合物的连接和/或提供用于治疗性疗法或诊断程序的制剂。

[0142] 一般来说,基本上纯化的 CH2 或 CH3 结构域分子包括在用于治疗性施用的完整药物制剂中相应化合物与其他成分混合或配制之前,制品中存在的所有大分子种类的多于 80%。其他成分可以包括药物载体、赋形剂、缓冲剂、吸收增强剂、稳定剂、防腐剂、佐剂或其他类似的共成分(co-ingredient)。更典型地,将 CH2 或 CH3 结构域分子纯化成代表在与其他制剂成分混合之前在纯化制品中存在大于 90%、通常大于 95%的所有大分子种类。在其他情况下,纯化的制品可以为实质上均质的,其中其他大分子种类小于 1%。

[0143] 重组体:重组体核酸或多肽是具有非天然存在的序列或具有通过人工组合两种原本分离的序列区段而制备的序列的核酸或多肽。人工组合通常是通过化学合成或更为常见地通过人工操作分离的核酸区段(例如通过基因工程技术)而实现。

[0144] 样品:代表整体的部分、片段或区段。该术语涵盖任何材料,包括例如从受试者获得的样品。

[0145] “生物样品”是从受试者中获得的样品,包括但不限于细胞、组织和体液。体液包括例如:唾液、痰、脑脊髓液、尿液、血液及血液的衍生物和级分,包括血清和淋巴细胞(如 B 细胞、T 细胞及其细分级分)。组织包括来自活组织检查物、尸体剖检物和病理学样本的组织以及活组织检查的组织或外科移除的组织,包括例如未固定的、冷冻的、固定在福尔马林中的和/或埋入石蜡中的组织。

[0146] 在具体的实施方案中,生物样品是从受试者如血液或血清中获得。生物样品通常是从哺乳动物中获得,所述哺乳动物如大鼠、小鼠、牛、狗、豚鼠、兔或灵长类。在一个实施方案中,灵长类是猕猴、黑猩猩或人。

[0147] 支架:如本文所用,CH2 或 CH3 结构域支架是可用作引入突变(例如到环区域中;参见图 2 和图 3A-3C)的平台以便使得抗原与该 CH2 或 CH3 结构域结合的重组 CH2 或 CH3 结构域。在一些实施方案中,支架被改变以表现出与天然 CH2 或 CH3 结构域相比提高的稳定性。在具体的实例中,诱变支架来引入半胱氨酸残基对,从而允许一个或多个非天然二硫键的形成。在一些情况下,支架是具有 N 端缺失(如约 1 个至约 7 个氨基酸的缺失)的 CH2 或 CH3 结构域。

[0148] 序列同一性:核苷酸序列或氨基酸序列之间的相似性按照这些序列之间的相似性

表达,另外称为序列同一性。序列同一性通常是按照百分比同一性(或相似性或同源性)测量;该百分比越高,这两个序列就越相似。同源物或变体当使用标准方法比对时,将拥有相对高程度的序列同一性。

[0149] 用于比较的序列比对方法在本领域内是公知的。各种程序和比对算法描述于:Smith 和 Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman 和 Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson 和 Lipman, *Proc. Natl. Acad. Sci. U. S. A.* 85:2444, 1988; Higgins 和 Sharp, *Gene* 73:237-244, 1988; Higgins 和 Sharp, *CABIOS* 5:151-153, 1989; Corpet 等人., *Nucleic Acids Research* 16:10881-10890, 1988; 以及 Pearson 和 Lipman, *Proc. Natl. Acad. Sci. U. S. A.* 85:2444, 1988. Altschul 等人., *Nature Genet.* 6:119-129, 1994。

[0150] NCBI 基本局部比对搜索工具 (BLAST™) (Altschul 等人., *J. Mol. Biol.* 215:403-410, 1990) 可从几种来源获得,包括美国国家生物技术信息中心 (NCBI, Bethesda, MD) 和供与序列分析程序 blastp、blastn、blastx、tblastn 和 tblastx 连接使用的互联网。

[0151] 特异性结合剂:基本上只与确定的靶结合的剂。因此,抗原特异性结合剂是基本上与抗原性多肽或其抗原性片段结合的剂。在一个实施方案中,特异性结合剂是特异性结合抗原性多肽或其抗原性片段的单克隆抗体或多克隆抗体或 CH2 或 CH3 结构域分子。

[0152] 涉及抗原,术语“特异性结合”是指抗体或其他配体整体或部分地与具有该抗原的细胞或组织优先结合而不与缺乏可检测的量的该抗原的细胞或组织优先结合。当然,认识到某种程度的非特异性相互作用可发生在分子与非靶细胞或非靶组织之间。尽管如此,特异性结合可以被区分为通过对抗原的特异性识别来介导。特异性结合导致抗体(或者 CH2 或 CH3 结构域分子)与具有该抗原的细胞之间的结合大大强于结合的抗体(或者 CH2 或 CH3 结构域分子)与缺乏该抗原的细胞之间的结合。特异性结合通常导致结合的抗体或者 CH2 或 CH3 结构域分子(每单位时间)对具有该抗原性多肽的细胞或组织的量相比于结合的抗体或者 CH2 或 CH3 结构域分子(每单位时间)对相应地缺乏抗原性多肽的细胞或组织的量,提高大于 2 倍,如大于 5 倍、大于 10 倍或大于 100 倍。在此类条件下特异性结合蛋白质需要依对具体蛋白的特异性而选择的抗体或者 CH2 或 CH3 结构域分子。多种免疫测定形式适合于选择与具体蛋白进行特异性免疫反应的抗体或者 CH2 或 CH3 结构域分子。例如,常规地使用固相 ELISA 免疫测定。

[0153] 受试者:有生命的多细胞生物体,包括脊椎动物生物体,该类别包括人和非人哺乳动物。

[0154] 治疗有效量:足以在以特定剂治疗的受试者中达到期望效果的该剂的量。此类剂包括本文所述的 CH2 或 CH3 结构域分子。例如,这可以是用于预防、治疗或改善 HIV 感染的 HIV 特异性 CH2 结构域分子的量。理想地,抗体的治疗有效量是足以预防、治疗或改善受试者的感染或者诸如由 HIV 感染所引起的疾病而不引起重大的细胞毒性效应的量。用于预防、改善和/或治疗受试者的剂的治疗有效量将取决于被治疗的受试者、病患的类型和严重性以及治疗组合物的施用方式。

[0155] 毒素:对细胞有细胞毒性的分子。毒素包括但不限于:相思豆毒蛋白、蓖麻毒蛋白、假单胞菌外毒素 (PE)、白喉毒素 (DT)、肉毒杆菌毒素、皂草素、局限曲菌素或白树毒素或其改性的毒素。例如,PE 和 DT 是通常通过肝毒性导致死亡的高毒性化合物。然而,可将 PE 和 DT 改性成为用作免疫毒素的形式,这是通过移除毒素的天然靶向组分(例如 PE 的结

构域 Ia 和 DT 的 B 链) 并且用不同的靶向部分(如 CH2 或 CH3 结构域分子) 取代该天然靶向组分。

[0156] 转导的: 转导的细胞是已通过分子生物学技术将核酸分子引入到其中的细胞。如本文所用, 术语转导涵盖可将核酸分子引入此种细胞中的所有技术, 包括用病毒载体转染、用质粒载体转化以及通过电穿孔、脂转染和粒子枪加速引入裸 DNA。

[0157] 肿瘤相关抗原(TAA): 能够刺激肿瘤特异性 T 细胞限定的免疫应答的肿瘤抗原。示例性 TAA 包括但不限于 RAGE-1、酪氨酸酶、MAGE-1、MAGE-2、NY-ESO-1、Melan-A/MART-1、糖蛋白(gp)75、gp100、 β -连环蛋白、PRAME、MUM-1、WT-1、CEA 以及 PR-1。其他 TAA 是本领域已知的(例如参见 Novellino 等人., *Cancer Immunol. Immunother.* 54(3):187-207, 2005) 并且包括还未鉴定出的 TAA。

[0158] 载体: 引入宿主细胞中, 由此产生转化的宿主细胞的核酸分子。载体可包括允许其在宿主细胞中复制的核酸序列, 如复制起点。载体还可以包括一种或多种可选择的标记基因和本领域已知的其他基因元件。

[0159] 病毒相关抗原(VAA): 能够刺激病毒特异性 T 细胞限定的免疫应答的病毒抗原。示例性 VAA 包括但不限于来自如下病毒的抗原: 人免疫缺陷病毒(HIV)、BK 病毒、JC 病毒、EB 病毒(EBV)、巨细胞病毒(CMV)、腺病毒、呼吸道合胞体病毒(RSV)、单纯疱疹病毒 6(HSV-6)、副流感病毒 3 或 B 型流感。

[0160] 除非另外解释, 否则本文所用的所有技术术语和科学术语具有与本发明所属领域的普通技术人员所通常理解的含义相同的含义。除非上下文清楚地另外指明, 否则单数术语包括复数指示物。类似地, 除非上下文清楚地另外指明, 否则单词“或”旨在包括“和”。所以“包含 A 或 B”表示包括 A 或 B, 或者 A 和 B。还应理解对于核酸或多肽所给出的所有的碱基大小或氨基酸大小以及所有的分子量或分子质量值都是近似值, 并且为描述而提供。尽管在本发明的实施或试验中可使用与本文所述的方法和材料相似或等同的方法和材料, 但在下面描述了适合的方法和材料。本文所提及的所有出版物、专利申请、专利和其他参考文献均通过引用以其整体并入本文。在冲突的情况下, 包括术语的解释, 将以本说明书为准。此外, 所述材料、方法和实例仅为示例性的, 且并非旨在限制。

[0161] III. 几个实施方案的概述

[0162] 常规的抗体是包含至少四条多肽链的大型多亚基蛋白复合物, 所述四条多肽链包括两条轻链和两条重链(参见图 1A 的常规免疫球蛋白的示意图)。抗体的重链和轻链含有结合抗原的可变区和提供结构支撑和效应物功能的恒定区(如 CH1、CH2 和 CH3 结构域)。抗原结合区包含两个分离的结构域, 重链可变结构域(V_H) 和轻链可变结构域(V_L)。典型的抗体如 IgG 分子具有大约 150kD 的分子量。已在蛋白酶消化后鉴定了天然存在的抗体的许多较小的抗原结合片段(例如, Fab、Fab' 和 $F(ab')_2$)。这些抗体片段具有范围在大约 50kD 至 100kD 的分子量。已使用重组方法来产生可选的抗原结合片段, 称为单链可变片段(scFv), 它是由通过合成的肽接头连接的 V_L 和 V_H 组成。scFv 分子具有大约 25-30kD 的分子量。

[0163] 然而, 在一些情况下, 抗体或抗体片段的治疗性使用可能由于抗体的大小而受限。例如, 如果抗体或抗体片段太大, 则组织渗透和表位进入可能受到限制。此外, 许多治疗性抗体是非人源的, 这能够在人类受试者中产生毒性。鉴于这些局限性, 能够特异性结合抗原

的小型人类抗体是利用抗体或其片段的诊断应用或治疗应用所期望的。

[0164] 本文描述了工程化抗体恒定结构域分子。本文公开了用作支架的重组 CH2 和 CH3 结构域分子,所述支架用于引入突变以使得抗原与该分子结合。还提供了特异性结合抗原的修饰的 CH2 和 CH3 结构域分子。在一些实施方案中,抗体恒定结构域是来自 IgG、IgA 或 IgD 的 CH2 结构域。在其他实施方案中,抗体恒定结构域是来自 IgE 或 IgM 的 CH3 结构域。公开的 CH2 和 CH3 结构域分子是小的、稳定的、可溶的,具有最小毒性乃至没有毒性,并且在一些情况下能够结合抗原。本文所述的 CH2 和 CH3 结构域分子不包含多于一个恒定结构域并且不包含免疫球蛋白可变结构域。

[0165] 本文提供了包含免疫球蛋白 CH2 或 CH3 结构域的多肽,其中所述 CH2 或 CH3 结构域包含至少一个来自异源免疫球蛋白可变结构域的互补决定区(CDR)或其功能片段(如 SDR)。还提供了在 CH2 或 CH3 结构域的一个或多个环中包含至少一个突变(如 1、2、3、4、5、6、7、8、9、10 或更多突变)的 CH2 或 CH3 结构域分子。本文所述的 CH2 或 CH3 结构域分子具有小于约 15kD 的分子量。在一些实施方案中,CH2 或 CH3 结构域分子具有约 12kD 至约 14kD 的分子量。在一些实施方案中,CH2 或 CH3 结构域包含约 1 个至约 7 个氨基酸,如 1、2、3、4、5、6 或 7 个氨基酸的 N 端平截。在一些实施方案中,CH2 或 CH3 结构域分子包含约 1 个至约 4 个氨基酸,如 1、2、3 或 4 个氨基酸的 C 端平截。

[0166] 将特定突变引入 CH2 或 CH3 结构域和 / 或异源 CDR 移入 CH2 或 CH3 结构域使得该多肽能够结合抗原。在一些实施方案中,来自异源免疫球蛋白的移入部分只包含 CDR 或其功能片段。在其他实施方案中,移入部分包含额外的序列,如高变环的全部或一部分。移入部分的长度可以不同,但通常是在 5 至 21 个氨基酸之间,包括 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 或 21 个氨基酸。在一个实施方案中,移入部分是在 8 至 15 个氨基酸之间。尽管移入部分的长度不同,但得到的 CH2 或 CH3 结构域分子特异性结合抗原。在一些实施方案中,CH2 或 CH3 结构域分子以约 10^{-6} 、约 10^{-7} 或约 10^{-8} M 的 K_d 特异性结合抗原。在一些实施方案中,多肽包含多于一个 CDR 或其功能片段,如两个或三个 CDR。

[0167] 在一些实施方案中,CH2 或 CH3 结构域的环区域的至少一部分被 CDR 或其功能片段置换。从环区域移除的氨基酸的数目可以不同。在一些实施方案中,从环区域移除的氨基酸的数目是在 1 至 10 个氨基酸之间,包括 1、2、3、4、5、6、7、8、9 或 10 个氨基酸。在其他实施方案中,移入 CDR 而没有从环区域移除氨基酸。从一个或多个 CH2 或 CH3 结构域环中移除的氨基酸的数目可以不同。本领域技术人员能够通过经验,例如通过测试 CH2 或 CH3 结构域分子的稳定性、溶解性和结合目的抗原的能力确定要移除的适当序列。

[0168] 所移入的具体 CDR 可以是来自任何免疫球蛋白可变结构域,如 V_H 结构域或 V_L 结构域的任何 CDR。在一个实施方案中,CDR 为 CDR1。在另一个实施方案中,CDR 为 CDR2。在另一个实施方案中,CDR 为 CDR3。在另一实施方案中,两个或三个或更多个 CDR 被移入到 CH2 或 CH3 结构域分子的环中。

[0169] 被 CDR 所置换的 CH2 或 CH3 结构域环(或 CDR 移入其中而未移除环序列的 CH2 结构域)可以是该 CH2 或 CH3 结构域的任何环。在一些实施方案中,环区域是选自环 1、环 2、环 3、环 A-B、环 C-D 或环 E-F。CH2 或 CH3 结构域的任何环可以被任何 CDR 置换。此外,多个环可以被 CDR 以任意组合置换。在一个实施方案中,环 1 被 CDR1 或 CDR3 置换。在另一个实施方案中,环 3 被 CDR1 或 CDR3 置换。在另一个实施方案中,环 1 和环 3 分别被 CDR1

和 CDR3 置换。在另一个实施方案中,环 1 和环 3 分别被 CDR3 和 CDR1 置换。在另一实施方案中,环 A-B 被 CDR1 置换;环 C-D 被 CDR2 置换;或者环 E-F 被 CDR3 置换。

[0170] 在优选的实施方案中,本文提供的多肽不包含可变结构域,如 V_H 结构域或 V_L 结构域。

[0171] 抗体恒定结构域可以获自任何类型的免疫球蛋白。在一个实施方案中,免疫球蛋白为 IgG。在其他实施方案中,免疫球蛋白为 IgA、IgD、IgM 或 IgE。在具体实例中,恒定结构域为来自 IgG 的 CH2 结构域。

[0172] 在一些实施方案中,结合抗原的 CH2 或 CH3 结构域具有提高该分子稳定性的额外的突变。例如,分子可以包含允许形成非天然二硫键的突变,所述非天然二硫键的形式是例如通过引入一对氨基酸取代以用半胱氨酸残基置换原始残基。在一些实例中,在恒定结构域的 N 端 A 链中引入第一氨基酸取代,而在恒定结构域的 C 端 G 链中引入第二氨基酸取代。此外,结合 CH2 和 CH3 结构域分子的抗原可以是糖基化的或非糖基化的。

[0173] 本文还提供了包含 IgG、Ig 或 IgD 的免疫球蛋白 CH2 结构域或者 IgE 和 IgM 的 CH3 结构域的多肽,其中所述 CH2 结构域或 CH3 结构域包含第一氨基酸取代和第二氨基酸取代,其中所述第一氨基酸取代和第二氨基酸取代各自用半胱氨酸残基置换原始的残基,其中所述半胱氨酸残基形成二硫键,并且其中所述多肽具有小于约 15kD 的分子量。此类 CH2 和 CH3 结构域表现出相对于未修饰的 CH2 和 CH3 结构域的稳定的稳定性,并因此用作将突变引入以使得抗原与 CH2 或 CH3 结构域结合的支架。

[0174] 在一些实施方案中,第一氨基酸取代是在 N 端 A 链中,而第二氨基酸取代是在 C 端 G 链中,这允许在 A 链与 G 链之间形成二硫键(参见图 3A-3C 环区域的示意图)。在一些实例中,恒定结构域为 IgG 的 CH2 结构域。

[0175] 在本文所述的具体实例中,第一氨基酸取代是 L12 至 C12,并且第二氨基酸取代是 K104 至 C104(关于 SEQ ID NO:5 编号)。在其他实例中,第一氨基酸取代是 V10 至 C10,并且第二氨基酸取代是 K104 至 C104(关于 SEQ ID NO:5 编号)。

[0176] CH2 和 CH3 结构域支架可包含例如提高稳定性或增强溶解度和表达的额外的突变。在一些实施方案中,CH2 或 CH3 结构域包含约 1 个至约 7 个氨基酸的 N 端平截。在具体的实例中,N 端平截是 1、2、3、4、5、6 或 7 个氨基酸长度。在一些实施方案中,CH2 或 CH3 结构域支架包含约 1 个至约 4 个氨基酸的 C 端平截。在具体的实例中,C 端平截是 1、2、3 或 4 个氨基酸。

[0177] 在一些实施方案中,还诱变 CH2 或 CH3 结构域支架以赋予抗原结合。在具体实施方案中,(i)CH2 或 CH3 结构域的至少一个环被突变;(ii)CH2 或 CH3 结构域的环区域的至少一部分被来自异源免疫球蛋白可变结构域的 CDR 或其片段置换;或者(iii)二者兼有。

[0178] 此外,CH2 或 CH3 结构域可以是非糖基化的或糖基化的。例如,可以在哺乳动物细胞中表达重组 CH2 或 CH3 结构域以允许翻译后修饰,如糖基化。

[0179] 在一些实施方案中,抗原来自病原体,如病毒或细菌。在一个实施方案中,病原体是 HIV。在其他实施方案中,抗原是癌症特异性抗原或癌症相关蛋白。在其他实施方案中,抗原与自身免疫疾病相关,例如 TNF- α 。

[0180] 在一些实施方案中,CH2 或 CH3 结构域分子结合肿瘤抗原。肿瘤抗原可以是本领域公知的肿瘤相关抗原。

[0181] 本文还提供了包含本文所述的 CH2 或 CH3 结构域分子的组合物。在一些实施方案中,所述组合物包含 CH2 结构域或 CH3 结构域以及药学上可接受的载体。

[0182] 本文还提供了编码公开的 CH2 或 CH3 结构域分子的核酸分子、包含该核酸序列的载体以及包含该载体的细胞。

[0183] 在一些实施方案中,工程化 CH2 或 CH3 结构域分子包含 Fc 受体结合位点并且能够结合至少一个 Fc 受体。在具体实例中,Fc 受体是新生儿 Fc 受体。结合 Fc 受体的能力将效应物功能赋予给 CH2 或 CH3 结构域分子,诸如例如 ADCC。在其他实施方案中,工程化 CH2 或 CH3 结构域结合能够活化补体系统的补体相关分子,如 C1q。在另外其他实施方案中,将 CH2 或 CH3 结构域分子结合至效应物分子,所述效应物分子包括但不限于治疗部分、诊断部分或检测部分。

[0184] 还提供了使用 CH2 或 CH3 结构域分子制备药物的方法。在一个实施方案中,药物用于治疗 HIV 感染。在另一个实施方案中,药物用于治疗癌症。在另一个实施方案中,药物用于治疗自身免疫疾患或炎性疾患。

[0185] 本文所述的 CH2 或 CH3 结构域分子可被工程化以特异性结合任何期望的抗原。鉴定和选择抗原特异性 CH2 或 CH3 结构域分子的方法可以使用本领域已知的任何适合的技术,例如通过使用噬菌体展示文库来实现。

[0186] 本文提供了一种鉴定特异性结合靶抗原的重组 CH2 结构域或 CH3 结构域的方法。所述方法包括:(a) 提供在它们的表面上展示重组 CH2 或 CH3 结构域的粒子的文库,其中所述 CH2 或 CH3 结构域具有小于约 15kD 的分子量;(b) 将所述粒子的文库与靶抗原接触来选择特异性结合所述靶抗原的粒子;以及(c) 从表达特异性结合所述靶抗原的 CH2 或 CH3 结构域的粒子中克隆 CH2 或 CH3 结构域核酸分子,由此鉴定特异性结合所述靶抗原的 CH2 或 CH3 结构域。在一些实施方案中,文库是通过如下步骤产生:(i) 提供编码 CH2 或 CH3 结构域的遗传多样性群体的核酸分子的文库,其中所述遗传多样性群体是通过将突变引入到所述 CH2 或 CH3 结构域的一个或多个环区域中来提供;以及(ii) 在重组宿主细胞中表达核酸分子的文库,从而将所述 CH2 结构域或 CH3 结构域表达在所述粒子的表面上并且由所述粒子的遗传材料编码所述 CH2 或 CH3 结构域核酸分子。在一些实施方案中,CH2 或 CH3 结构域包含约 1 个至约 7 个氨基酸的 N 端缺失。在一些实施方案中,粒子为噬菌体粒子。

[0187] 在一些实施方案中,噬菌体文库表达重组 CH2 结构域,例如 IgGCH2 结构域。在一些实施方案中,CH2 结构域或 CH3 结构域包含在环 1 中的至少一个突变、或在环 2 中的至少一个突变、或在环 3 中的至少一个突变、或在环 A-b 中的至少一个突变、或在环 C-D 中的至少一个突变、或在环 E-F 中的至少一个突变、或其任意组合。

[0188] 可使用任何适合的重组宿主细胞来产生噬菌体粒子。此类宿主细胞是本领域公知的。在一些实例中,重组宿主细胞是 TG1 细胞。

[0189] 本文还提供了一种制备重组 CH2 或 CH3 结构域的文库的方法,所述方法包括(i) 将突变引入 CH2 结构域或 CH3 结构域支架的一个或多个环区域中,或者(ii) 用来自异源免疫球蛋白可变结构域的 CDR 或其功能片段置换所述 CH2 结构域或 CH3 结构域支架的环区域的一部分,或者(iii) 二者兼有,其中所述支架包含 IgG、IgA 或 IgD 的分离的免疫球蛋白 CH2 结构域,或者 IgE 或 IgM 的 CH3 结构域。

[0190] 在一些实施方案中,CH2 或 CH3 结构域支架还包含约 1 个至约 7 个氨基酸,例如约

1、2、3、4、5、6 或 7 个氨基酸的 N 端平截。在一些实施方案中,CH2 或 CH3 结构域支架还包含约 1 个至约 4 个氨基酸,例如约 1、2、3 或 4 个氨基酸的 C 端平截。

[0191] 在一些情况下,CH2 或 CH3 结构域支架还包含稳定分子的额外的突变。在该方法的一些实施方案中,所述 CH2 或 CH3 结构域支架还包含第一氨基酸取代和第二氨基酸取代,其中所述第一和第二氨基酸取代各自用半胱氨酸残基取代原始的残基,其中所述半胱氨酸残基形成二硫键。

[0192] 还提供了一种鉴定特异性结合靶抗原的重组 CH2 结构域或 CH3 结构域的方法,所述方法包括使通过本文公开的方法产生的文库与所述靶抗原接触以选择特异性结合所述靶抗原的重组 CH2 或 CH3 结构域。

[0193] 还提供了 CH2 或 CH3 结构域分子的文库,例如噬菌体展示文库。文库包含具有一个或多个突变、移入的 CDR、高变环或其功能片段的 CH2 或 CH3 结构域分子。可使用包含突变残基的文库来鉴定具有期望的抗原结合亲和力的 CH2 或 CH3 结构域分子和 / 或来鉴定具有减少的免疫原性的 CH2 或 CH3 结构域分子。

[0194] 还提供了包含本文所公开的 CH2 或 CH3 结构域分子的试剂盒。在一个实施方案中,CH2 或 CH3 结构域分子被标记(例如用荧光标记、放射性标记或酶标记)。在另一实施方案中,试剂盒包括公开 CH2 或 CH3 结构域分子的使用方法的说明材料。说明材料可能是书面的,电子形式的(例如计算机软盘或光盘)或者可能是可视的(例如视频文件)。试剂盒还可以包括促进对试剂盒设计的特殊应用的额外组分。因此,例如试剂盒可额外地含有检测标记的手段(例如用于酶标记的酶底物、检测荧光标记的滤光片套件、适当的第二标记如第二抗体,等)。试剂盒可以额外地包括常规用于实施具体方法的缓冲剂和其他试剂。此类试剂盒和适当的内容物是本领域的技术人员公知的。

[0195] IV. 工程化抗体恒定结构域

[0196] 本文所述的工程化抗体恒定结构域分子尺寸很小(通常小于 15kD),这为检测、诊断和治疗提供了显著优势。例如,分子的小尺寸允许更大的表位进入和更好的组织渗透。如在图 5C 中所示,本文所提供的 CH2 结构域抗体具有低于其他类型的抗体和抗体片段,诸如 scFv、Fab 和 IgG 分子的分子量。它们还小于 V_H 结构域抗体。

[0197] 如本文所述,CH2 或 CH3 结构域分子能够在不存在包括可变结构域或其他恒定结构域在内的其他免疫球蛋白结构域下有效地结合抗原。例如,CH2 或 CH3 结构域分子能够以约 10^{-6} 、约 10^{-7} 、约 10^{-8} 、或约 10^{-9} 或更小的 kD 特异性结合抗原。

[0198] 本文所述的特异性结合抗原的 CH2 或 CH3 结构域包含来自免疫球蛋白可变结构域的至少一条异源氨基酸序列和 / 或包含至少一个突变。在 CH2 或 CH3 结构域中移入的异源氨基酸序列包含至少一个 CDR 或其功能片段(例如来自特异性结合目的抗原的抗体的 SDR)。移入的氨基酸序列还可以含有从 CDR 向 N 端和 / 或向 C 端延伸的额外的氨基酸序列,例如包含高变环的其他氨基酸。因此,在一些实施方案中,工程化 CH2 或 CH3 结构域分子包含来自异源免疫球蛋白可变结构域的完整高变环。工程化 CH2 和 CH3 结构域还可以包含第二或第三 CDR 或高变环。移入的 CDR 或高变环的长度可以不同。适当的长度可以通过经验确定,例如通过表达工程化 CH2 或 CH3 结构域并评价蛋白的稳定性和可溶性,以及通过确定结合亲和力。决定蛋白可溶性和评估抗原结合亲和力的蛋白表达的方法是本领域公知的。如本文所述,已确定高达 21 个氨基酸长度的序列可成功移入 CH2 结构域。

[0199] 人 CH2 结构域包含六个环区域：环 1、环 2、环 3、环 A-B、环 C-D 和环 E-F。来自异源免疫球蛋白可变结构域的 CDR 和 / 或高变环可以任意组合移入一个或多个任何这些环中（例如参见图 5A-5C）。

[0200] 人 γ 1CH2 结构域的氨基酸序列作为 SEQ ID NO :5 列出。包含每个环区域的氨基酸残基示于以下表 1 中。以 CH2 的第一个残基的编号 1 开始对氨基酸位置编号。

[0201] 表 1

[0202] CH2 结构域的环的氨基酸位置

[0203]

环	氨基酸位置 (SEQ ID NO :5)
环 A-B	14-27
环 1	35-43
环 C-D	54-62
环 2	67-69
环 E-F	78-88
环 3	96-100

[0204] 人 V_H 结构域的氨基酸序列示于图 1B 中并且作为 SEQ ID NO :1 列出。包含每个 CDR 和高变环的氨基酸残基示于以下表 2 中。

[0205] 表 2

[0206] 高变环的氨基酸位置

[0207]

CDR/ 环	氨基酸位置 (SEQ ID NO :1)
H1/CDR1	27-36
H2/CDR2	50-68
H3/CDR3	99-109

[0208] 在一个示例性实施方案中，用来自人抗体的 V_H 结构域的高变环 H1/CDR1 的 10 个氨基酸置换来自 CH2 结构域的环 1 的九个氨基酸。在其他示例性实施方案中，用来自人抗体的 V_H 结构域的高变环 H3/CDR3 的十二或十三个氨基酸置换来自 CH2 结构域的环 3 的六个氨基酸。在另一个示例性实施方案中，用来自人抗体的 V_H 结构域的高变环 H1/CDR1 的 10 个氨基酸置换来自 CH2 结构域的环 3 的六个氨基酸。在其他示例性实施方案中，用来自人抗体的 V_H 结构域的高变环 H3/CDR3 的十二或十三个氨基酸置换来自 CH2 结构域的环 1 的九个氨基酸。

[0209] 在其他实施方案中，用来自异源抗体的一个或多个 CDR 或高变环的 5、6、7、8、9、

10、11、12、13、14、15、16、17、18、19、20 或 21 个氨基酸以任意组合置换环 1、2、3、A-B、C-D 或 E-F 中的一个或多个的 0、1、2、3、4、5、6、7、8、9 或 10 个氨基酸。CDR 或高变环可以来自 V_L 或 V_H 结构域（参见图 3A-3C）。

[0210] 所移入的一个或多个环或者一个或多个 CDR 可以来自任何目的抗体。此类抗体包括但不限于病原体特异性抗体和癌症特异性抗体。病原体特异性抗体包括例如特异性结合抗原的抗体，所述抗原来自诸如病毒、细菌或真菌、原生动物或寄生虫的病原体。在一个示例性实施方案中，抗体特异性结合 HIV-1。癌症特异性抗体包括特异性识别被癌细胞表达（例如在细胞表面上）但不被其他非癌细胞表达的抗原的抗体。癌症特异性抗体包括但不限于识别以下的抗体：肺癌、乳腺癌、前列腺癌、肝癌、膀胱癌、甲状腺癌、肾癌、胰腺癌、结肠直肠癌、皮肤癌、黑色素瘤、成神经细胞瘤、尤因氏肉瘤、白血病或淋巴瘤的细胞或组织。

[0211] 在一些实施方案中，工程化 CH2 或 CH3 结构域分子包含具有已知特异性的 CDR/ 高变序列。或者，工程化 CH2 结构域分子可包含一条或多条随机的 CDR 肽序列。可以进行 CDR 的突变分析来鉴定具有提高的结合亲和力和 / 或降低的免疫原性的 CH2 结构域分子。此外，可以产生包含随机的或突变的 CDR 肽序列的 CH2 或 CH3 结构域分子的文库来鉴定以高亲和力结合如以下所述的特定目的抗原的 CH2 或 CH3 结构域分子。

[0212] 例如出于治疗、诊断或检测的目的，本文所提供的 CH2 和 CH3 结构域分子还可以包含效应物分子。例如，效应物分子可以包括毒素和可检测的标记，例如放射性标记、酶或荧光标志物。以下描述了可与 CH2 和 CH3 结构域分子一起使用的效应物分子类型的其他细节（参见“抗体恒定结构域分子的效应物功能”）。

[0213] V. 抗体恒定结构域分子文库

[0214] 本文还提供了包含随机插入的或突变的 CDR 氨基酸序列的工程化 CH2 或 CH3 结构域分子的文库。所述文库可被用来筛选对目的特定抗原具有高亲和力的 CH2 或 CH3 结构域分子。在一个实施方案中，文库是噬菌体展示文库。抗体噬菌体展示文库和产生此类文库的方法是本领域公知的（参见例如，美国专利第 6,828,422 号和第 7,195,866 号，通过引用并入本文）。

[0215] 已描述了多肽（包括抗体）的文库的开发（美国专利第 6,828,422 号）。为了产生多肽的文库（例如 CH2 或 CH3 结构域分子的文库），首先必须产生适合于创建文库的核酸序列。为了产生此类随机的核酸序列，通常使用易错 PCR。在至少一个环中随机引入突变。例如，鉴定了同源蛋白的集合（例如两个或三个或更多），建立了蛋白序列的数据库，并且将蛋白序列彼此进行比对。在 CH2 结构域分子的情况中，鉴定了人 CH2 结构域序列的集合并将其用来创建数据库。将该数据库用来限定在序列和 / 或结构排列中显示高度相似性的蛋白序列的亚组。对于每个亚组而言，推导出包含至少一个共有序列的多肽序列，所述共有序列代表这一亚组（例如 CH2 结构域的亚组）的成员。多肽序列的完整集合代表同源蛋白（CH2 结构域）的集合的完整的结构库。可以根据这些人工多肽序列的结构特性对其分析以鉴定在所述多肽序列内的氨基酸之间或在所述多肽序列与其他多肽序列之间不利的相互作用。因此可以通过改变共有序列将此类相互作用除去。

[0216] 接下来，分析多肽序列以识别子元件，包括结构域、环、 β -折叠、 α -螺旋和 / 或 CDR。将氨基酸序列反向翻译成对应的编码核酸序列，该编码核酸序列适于计划用于表达所述核酸序列的宿主的密码子使用。建立一组切割位点以使得编码如以上所述鉴定的子元件

的每个子序列侧接核酸序列内没有第二次出现的两个位点。这可以通过如下方式实现：通过鉴定已经侧接子序列的切割位点，或通过改变一个或多个核苷酸以产生切割位点，以及通过从基因的剩余部分中除去该位点。切割位点对所有对应的子元件或子序列应该是共同的，这允许产生在核酸序列中的子序列的完全模块化排列和在对应多肽中的子元件的完全模块化排列。

[0217] 上述核酸序列使用本领域公知的几种方法的任意一种合成，诸如例如通过全基因合成或通过基于 PCR 的方法。

[0218] 在一个实施方案中，将核酸序列克隆到载体中。载体可以是本领域公知的测序载体、表达载体或展示载体（例如噬菌体展示）。载体可以包含在不同或相同的操作子中的一个核酸序列、或者两个或更多个核酸序列。如果在相同的操纵子中，核酸序列可以被分开克隆或者作为连续序列克隆。

[0219] 在一个实施方案中，核酸序列的一个或多个子序列（例如环）被不同的序列置换。这可以通过使用如下方式实现：例如通过适当的限制性酶，使用与子序列相邻或在子序列末端的切割位点将子序列切除，并由与已切割的核酸序列相容的不同序列置换子序列。在进一步的实施方案中，置换原来的一个或多个子序列的不同序列（也称为“移入的序列”）是来自异源抗体的基因组序列或重排的基因组序列，例如 CDR、SDR 或高变环。在一些实施方案中，异源序列是随机序列。随机序列的引入将变异性引入到多肽（或 CH2 结构域分子）中以创建文库。可以使用本领域公知的许多方法的任一种来产生随机序列，例如通过使用在自动化寡核酸合成过程中使用单核苷酸或三核苷酸的混合物，或者通过易错 PCR 进行。根据在已知蛋白序列中确定位置的氨基酸分布，随机序列可以是完全随机的，或者偏向于某些密码子，或偏离某些密码子。此外，随机子序列的集合可以包含不同数目的密码子，产生具有不同长度的子元件的集合。

[0220] 可以在本领域公知的适当条件下从适合的载体中表达核酸序列。在一个实施方案中，筛选从核酸序列中表达的多肽。还可以对多肽进行优化。筛选可以通过使用本领域公知的任何方法进行，例如噬菌体展示、选择性感染的噬菌体、筛选结合的多核糖体技术、酶活性或蛋白稳定性的测定系统。可以通过对核酸序列或氨基酸序列测序或通过质谱来鉴定具有期望特性的多肽（例如 CH2 结构域分子）。筛选的多肽的期望特性可以是例如对靶分子的最佳亲和力或特异性。

[0221] 在一些实施方案中，可以使用噬菌粒载体来同时表达大量的核酸序列，例如编码 CH2 或 CH3 结构域分子的文库的那些核酸序列（参见例如，美国专利申请公开号 2008/0312101）。可以使用已知可应用于噬菌体的任何筛选测定法筛选表达 CH2 和 CH3 结构域的噬菌体粒子的文库。例如，可将噬菌体暴露于可溶的或固定的（例如在板上或在珠上）纯化抗原或者暴露于全细胞、组织或动物，以便鉴定粘附于在复合物结构中且特别在生理上或治疗上相关的位置中存在的靶（例如，与癌细胞或与病毒粒子上的抗原结合）的噬菌体。

[0222] 可以将其中已克隆、表达并基于其对特定配体的结合而特异性分离异源序列的所选噬菌粒载体从细菌细胞中提取，并测序、PCR 扩增和 / 或重新克隆到另一适当的载体，例如用于在细菌、植物、酵母或哺乳动物细胞中大规模的重组生产。

[0223] 与特定靶抗原的相互作用的检测可以通过应用标准的淘选方法或通过应用于

评估在展示的 CH2 或 CH3 结合分子与其靶抗原之间的相互作用的更为复杂的生物物理学技术来进行,所述生物物理学技术例如基于荧光的光谱法或显微术、磷酸酶反应或其他高通量技术。

[0224] 特异性结合靶抗原的表达 CH2 或 CH3 结构域的噬菌体粒子被选定后,可以根据本领域已知的方法对重组的噬菌体和相关的 DNA 序列分离并表征(例如,使用限制性酶从噬菌粒载体中分离、直接测序和/或通过 PCR 扩增)。然后可以将这些序列转移到更适合的载体中,用于进一步修饰和/或表达达到原核或真核宿主细胞。编码 CH2 或 CH3 结构域的 DNA 序列一旦插入到适合的载体中即可通过任何适合的手段(转化、转染、接合、原生质体融合、电穿孔、磷酸钙沉淀、直接微注射等等)引入到适当的宿主细胞中来转化细胞。

[0225] 然后可以使用这种 DNA 分子的集合来创建 CH2 或 CH3 结构域分子的文库。可以使用上述方法来优化 CH2 或 CH3 结构域分子的亲和力。可以将文库用来鉴定与靶结合的一种或多种 CH2 或 CH3 结构域分子。期望的 CH2 或 CH3 结构域分子的鉴定包括表达 CH2 或 CH3 结构域分子,并且然后筛选它们来分离以期望的亲和力与给定的靶分子结合的一种或多种分子。如果必要, DNA 分子的模块设计允许切除一个或多个遗传子序列并用一个或多个编码结构子元件的第二子序列置换。然后可以重复表达和筛选的步骤直到产生具有期望的亲和力的 CH2 或 CH3 结构域分子。

[0226] 在一个实施方案中是其中使用切割位点由序列的随机集合(文库)置换一个或多个遗传亚基(例如,一个或多个 CH2 或 CH3 结构域环区域)的方法。然后针对任何选定的抗原筛选得到的文库。选择、收集具有期望的特性(例如具有期望的结合亲和力)的 CH2 或 CH3 结构域分子并可将其用作下个文库的起始材料。

[0227] 在另一个实施方案中,可以通过提供编码如以上所述的多肽和额外部分二者的 DNA 序列来产生融合蛋白。此类部分包括免疫毒素、酶、效应物分子、治疗分子、标记或标签(例如用于检测和/或纯化)。

[0228] 本文还提供了核酸序列、含有核酸序列的载体、含有载体的宿主细胞以及根据上述方法产生的多肽。

[0229] 还提供了包含根据上述方法的核酸序列、重组载体、多肽和/或载体中的一种或多种的试剂盒,以及用于产生多肽的适合的宿主细胞。

[0230] VI. 编码抗体恒定结构域分子的核酸

[0231] 编码 CH2 或 CH3 结构域分子和/或免疫毒素的核酸序列可以通过包括例如适当序列的克隆或通过直接的化学合成在内的任何适合的方法来制备,所述化学合成通过如下方法,例如:Narang 等人., Meth. Enzymol. 68:90-99, 1979 的磷酸三酯法;Brown 等人., Meth. Enzymol. 68:109-151, 1979 的磷酸二酯法;Beaucage 等人, Tetra. Lett. 22:1859-1862, 1981 的二乙基磷酰胺酸法;由 Beaucage&Caruthers, Tetra. Letts. 22(20):1859-1862, 1981 描述的固相亚磷酰胺三酯法,该方法使用如在例如 Needham-VanDevanter 等人, Nucl. Acids Res. 12:6159-6168, 1984 中所述的自动化分析仪;以及美国专利第 4,458,066 号的固相方法。化学合成产生了单链寡核苷酸。这可以通过与互补序列杂交或者通过使用单链作模板用 DNA 聚合酶进行的聚合反应而被转换为双链 DNA。技术人员将认识到虽然 DNA 的化学合成一般限制为约 100 个碱基的序列,但是较长序列可以通过连接较短序列来获得。

[0232] 编码 CH2 或 CH3 结构域分子或编码包括 CH2 或 CH3 结构域分子的免疫毒素的

示例性核酸编码序列可通过克隆技术制备。适宜的克隆和测序技术的实例以及足以指导技术人员通过许多克隆实践的指示说明参见以上的 Sambrook 等人、以上的 Berger 和 Kimmel (编) 以及以上的 Ausubel。来自生物试剂制造商和实验设备制造商的产品信息也提供了有用的信息。此类制造商包括 SIGMA Chemical Company (Saint Louis, MO)、R&D Systems (Minneapolis, MN)、Pharmacia Amersham (Piscataway, NJ)、CLONTECH Laboratories, Inc. (Palo Alto, CA)、ChemGenes Corp.、Aldrich Chemical Company (Milwaukee, WI)、Glen Research, Inc.、GIBCO BRL Life Technologies, Inc. (Gaithersburg, MD)、FlukaChemica-Biochemika Analytika (Fluka Chemie AG, Buchs, Switzerland)、Invitrogen (San Diego, CA) 和 Applied Biosystems (Foster City, CA) 以及技术人员已知的许多其他的商业来源。

[0233] 还可以通过扩增方法制备核酸。扩增方法包括聚合酶链式反应 (PCR)、连接酶链式反应 (LCR)、基于转录的扩增系统 (TAS)、自主序列复制系统 (3SR)。种类广泛的克隆方法、宿主细胞和体外扩增方法学是技术人员公知的。

[0234] 在一个实例中,使用的 CH2 结构域分子是通过将编码该 CH2 结构域分子的 cDNA 插入到包含编码效应物分子 (EM) 的 cDNA 的载体中来制备。进行这种插入以使得可变区和 EM 被读在框内从而使得产生一种连续多肽。因此,编码的多肽含有功能 CH2 结构域区域和功能 EM 区域。在一个实例中,将编码效应物分子 (例如但不限于细胞毒素) 的 cDNA 连接至 CH2 结构域分子以使得 EM 定位在该 CH2 结构域分子的羧基端。在一个实例中,将编码假单胞菌外毒素 (“PE”) 的 cDNA (其被突变以消除或减少非特异性结合) 连接至 CH2 结构域分子以使得 EM 定位在 CH2 结构域分子的氨基端。

[0235] 将编码 CH2 结构域分子 (或免疫毒素) 的核酸分离和克隆后,可将蛋白表达在重组工程化细胞,例如细菌、植物、酵母、昆虫或哺乳动物细胞中。例如,可以通过将 DNA 转移到适合的宿主细胞中在体内表达编码 CH2 结构域分子的一种或多种 DNA 序列。细胞可以是原核的或真核的。该术语还包括任何目标宿主细胞的子代。应理解所有的子代可以不与亲代细胞完全相同,因为可能存在复制过程中发生的突变。本领域已知稳定转移的方法,即将外源 DNA 连续地保持在宿主中。或者,可以在体外表达编码免疫毒素、抗体或其片段的 DNA 序列。

[0236] 编码 CH2 或 CH3 结构域分子的多核苷酸序列能够与表达控制序列可操作地连接。连接与编码序列可操作地连接的表达控制序列以使得在与该表达控制序列相容的条件下实现该编码序列的表达。表达控制序列可包括但不限于适当的启动子、增强子、转录终止子、在编码蛋白的基因前面的起始密码子 (例如 ATG)、内含子的剪接信号、允许 mRNA 的适当翻译的该基因的正确阅读框的维持以及终止密码子。

[0237] 可以将编码 CH2 或 CH3 结构域分子的多核苷酸序列插入到表达载体中,所述表达载体包括但不限于能够被操作以允许序列的插入或引入并且能够被表达在原核生物或真核生物中的质粒、病毒或其他载体。宿主可以包括微生物、酵母、昆虫和哺乳动物生物体。在原核生物中表达具有真核序列或病毒序列的 DNA 序列的方法是本领域公知的。能够在宿主中表达和复制的有生物功能的病毒和质粒 DNA 载体是本领域已知的。

[0238] 用重组 DNA 转化宿主细胞可以通过本领域技术人员已知的常规技术进行。当宿主为能够摄取 DNA 的原核 (例如大肠杆菌) 时,感受态细胞,可以由指数生长期之后收获的并

且随后通过 CaCl_2 方法使用本领域公知的程序来处理的细胞制备。或者,可以使用 MgCl_2 或 RbCl 。转化还可以必要时在形成宿主细胞的原生质之后进行,或者通过电穿孔进行。

[0239] 当宿主为真核生物时,可以使用诸如磷酸钙共沉淀法这类转染 DNA 的方法、诸如微注射这类常规机械程序、电穿孔、包围在脂质体中的质粒的插入或病毒载体。真核细胞还可以用编码免疫毒素、抗体或其片段的多核苷酸序列和编码可选择的表型的第二外源 DNA 分子(例如单纯疱疹病胸腺激酶基因)共转化。另一方法是使用诸如猿猴病毒 40(SV40)或牛乳头瘤病毒的真核病毒载体来瞬时感染或转化真核细胞并表达蛋白(参见例如, *Eukaryotic Viral Vectors*(真核病毒载体),冷泉港实验室,Gluzman 编.,1982)。本领域技术人员可以容易地使用表达系统,例如在细胞中用于产生蛋白的质粒和载体,所述细胞包括高等真核生物细胞,例如 COS、CHO、HeLa 和骨髓瘤细胞系。

[0240] 可以通过包括制备型色谱和免疫分离的常规手段进行重组表达的多肽(例如 CH2 结构域分子)的分离和纯化。被表达后,重组表达的多肽可以根据本领域的标准程序纯化,所述标准程序例如硫酸铵沉淀、亲和柱、柱层析以及类似程序(一般参见 R. Scopes, *Protein Purification*(蛋白质纯化),Springer-Verlag,N.Y.,1982)。本文公开了具有至少约 90%至 95%的同质性的基本上纯的组合物,并且出于药用目的,可以使用 98%至 99%或更高同质性。如果在治疗上使用,一旦被纯化,部分地或达到所期望的同质性后,多肽应该基本上不含内毒素。

[0241] 用于从诸如大肠杆菌的细菌中表达蛋白和/或重折叠成适当的活性形式的方法已被描述并且是公知的,并且可应用于本文所公开的抗体。参见 Buchner 等人, *Anal. Biochem.* 205:263-270,1992;Pluckthun, *Biotechnology* 9:545,1991;Huse 等人, *Science* 246:1275,1989;以及 Ward 等人, *Nature* 341:544,1989,全部通过引用并入本文。

[0242] 通常,来自大肠杆菌或其他细菌的有功能的异源蛋白可以从包涵体中分离,并且需要使用强变性剂增溶和随后的重折叠。在增溶步骤过程中,如本领域中公知的,必须存在还原剂以分离二硫键。具有还原剂的示例性缓冲剂为:0.1M Tris pH 8、6M 胍、2mM EDTA、0.3MDTE(二硫赤藓糖醇)。可以通过将变性和还原的蛋白稀释(例如 100 倍)到重折叠缓冲剂中来实现复性。示例性的缓冲剂为 0.1M Tris, pH8.0、0.5M L-精氨酸、8mM 氧化谷胱甘肽(GSSG)和 2mM EDTA。

[0243] 除了重组的方法之外,还可以使用标准的肽合成整体上或部分地构建本文公开的 CH2 和 CH3 结构域分子。长度小于约 50 个氨基酸的多肽的固相合成可以通过将该序列的 C 端氨基酸连接至不可溶的支撑体、随后顺序添加该序列中的剩余氨基酸来实现。用于固相合成的技术描述于 Barany&Merrifield, *The Peptides:Analysis, Synthesis, Biology*(肽:分析、合成生物学). 第 2 卷:Special Methods in Peptide Synthesis(肽合成中的具体方法),A 部分. 第 3-284 页;Merrifield 等人, *J. Am. Chem. Soc.* 85:2149-2156, 1963;以及 Stewart 等人, *Solid Phase Peptide Synthesis*(固相肽合成),第二版,Pierce Chem. Co., Rockford, Ill.,1984。更长的蛋白可以通过缩合较短片段的氨基端和羧基端来合成。通过激活羧基末端形成肽键的方法(例如,通过使用偶联试剂 N,N'-二环己基碳二亚胺)是本领域公知的。

[0244] VII. 使用抗体恒定结构域分子用于诊断或治疗

[0245] CH2 和 CH3 结构域分子具有用于诊断和/或治疗许多抗体可用的疾病或病症中的

任一种的巨大潜力。例如,CH2 或 CH3 结构域分子可以用于治疗癌症、传染性疾病(例如病毒、细菌、真菌或寄生虫的感染)、自身免疫疾病、炎性疾患或者抗体或其片段可以用作治疗剂的任何其他疾病或病症。

[0246] 在一些实施方案中,传染性疾病是由病毒引起,诸如来自以下科之一的病毒:逆转录病毒科(例如人免疫缺陷病毒(HIV);人T细胞性白血病病毒(HTLV);小RNA病毒科(例如脊髓灰质炎病毒、甲型肝炎病毒、丙型肝炎病毒、肠道病毒、人柯萨奇病毒、鼻病毒、艾可病毒、口蹄疫病毒);杯状病毒科(例如引起胃肠炎的毒株);披膜病毒科(例如马脑炎、风疹病毒);黄病毒科(例如登革病毒、黄病毒、西尼罗病毒、圣路易斯脑炎病毒、乙型脑炎病毒以及其他脑炎病毒);冠状病毒科(例如冠状病毒、严重急性呼吸窘迫综合征(SARS)病毒);弹状病毒科(例如疱疹性口炎病毒、狂犬病毒);纤丝病毒科(例如埃博拉病毒);副粘病毒科(例如副流感病毒、腮腺炎病毒、麻疹病毒、呼吸道合胞病毒(RSV));正粘病毒科(例如流感病毒);布尼亚病毒科(例如汉坦病毒、辛诺瓦病毒、立夫特山谷热病毒、布尼亚病毒、静脉病毒和内罗病毒);沙粒病毒科(出血热病毒、马秋波病毒、鸠宁病毒);呼肠孤病毒科(例如呼肠孤病毒、环状病毒和轮状病毒);双RNA病毒科;嗜肝DNA病毒科(乙型肝炎病毒);细小病毒科(细小病毒);乳多空病毒科(乳头瘤病毒、多瘤病毒、BK病毒);腺病毒科(大部分腺病毒);疱疹病毒科(单纯疱疹病毒(HSV)-1和HSV-2;巨细胞病毒(CMV);EB病毒(EBV);水痘带状疱疹病毒(VZV);以及其他疱疹病毒,包括HSV-6);痘病毒科(天花病毒、牛痘病毒、痘病毒);和虹彩病毒科(如非洲猪瘟病毒);纤丝病毒科(例如埃博拉病毒、马尔堡病毒);杯状病毒科(例如诺瓦克病毒)和未分类病毒(例如海绵状脑病的病原、 δ 肝炎的病原(被认为是乙型肝炎缺陷型卫星病毒);以及星状病毒)。

[0247] 在其他实施方案中,传染性疾病由一种类型的细菌引起,所述细菌例如为:幽门螺杆菌、伯氏疏螺旋体、嗜肺军团菌、分枝杆菌(如结核分枝杆菌、禽分枝杆菌、胞内分枝杆菌、堪萨斯分枝杆菌、戈登分枝杆菌)、金黄色葡萄球菌、淋病奈瑟菌、脑膜炎奈瑟菌、单核细胞增生李斯特菌、化脓链球菌(A群链球菌)、无乳链球菌(B群链球菌)、链球菌(草绿色种群)、粪链球菌、牛链球菌、链球菌(厌氧种)、肺炎链球菌、致病性弯曲杆菌、肠道球菌、流感嗜血菌、炭疽芽孢杆菌、白喉棒杆菌、棒状杆菌、猪红斑丹毒丝菌、产气荚膜梭菌、破伤风梭菌、产气肠杆菌、肺炎克雷伯氏杆菌、多杀性巴氏杆菌、拟杆菌、具核梭杆菌、念珠状链杆菌、苍白密螺旋体、极细密螺旋体、钩端螺旋体属以及以色列放线菌。

[0248] 在其他实施方案中,传染性疾病是由真菌引起的,例如:新型隐球菌、荚膜组织胞浆菌、粗球孢子菌、皮炎芽生菌、沙眼衣原体或白色念珠菌。在其他实施方案中,传染性疾病是由寄生虫引起的,例如恶性疟原虫和刚地弓形体引起。

[0249] 在一些实施方案中,癌症为实体肿瘤或血原性癌。在具体实例中,实体肿瘤为肉瘤或癌,例如:纤维肉瘤、粘液肉瘤、脂肪肉瘤、软骨肉瘤、成骨性肉瘤、或另外的肉瘤、滑膜瘤、间皮瘤、尤因氏瘤、平滑肌肉瘤、横纹肌肉瘤、结肠癌、淋巴样恶性肿瘤、胰腺癌、乳腺癌、肺癌、卵巢癌、前列腺癌、肝细胞癌、鳞状细胞癌、基底细胞癌、腺癌、汗腺癌、皮脂腺癌、乳头状癌、乳头状腺癌、髓样癌、支气管原癌、肾细胞癌、肝癌、胆管癌、绒毛膜癌、维尔姆斯氏肿瘤、子宫颈癌、睾丸癌、膀胱癌、或中枢神经系统肿瘤(如神经胶质瘤、星形细胞瘤、髓母细胞瘤、颅咽管瘤、室管膜瘤、松果体瘤、血管母细胞瘤、听神经瘤、少突神经胶质瘤、脑膜瘤、黑色素瘤、成神经细胞瘤或视网膜母细胞瘤)。

[0250] 在一些实例中,血源性癌症是白血病,例如:急性白血病(如急性淋巴细胞性白血病、急性粒细胞性白血病、急性骨髓性白血病和成髓细胞性白血病、前髓细胞性白血病、骨髓单核细胞性白血病、单核细胞性白血病及红白血病);慢性白血病(如慢性髓细胞性(粒细胞性)白血病、慢性骨髓性白血病以及慢性淋巴细胞性白血病);真性红细胞增多症;淋巴瘤;霍奇金氏病;非霍奇金氏淋巴瘤(无痛形式和高级形式);多发性骨髓瘤;沃尔丹斯特伦氏巨球蛋白血症;重链病;骨髓增生异常综合症;多毛细胞白血病;或骨髓发育不良。

[0251] 在一些实施方案中,CH2或CH3结构域分子特异性结合肿瘤抗原。肿瘤抗原是本领域公知的并且包括例如:癌胚抗原(CEA)、 β -人体绒毛膜促性激素(β -HCG)、 α 甲胎蛋白(AFP)、凝集素反应性AFP(AFP-L3)、甲状腺球蛋白、RAGE-1、MN-CA IX、人端粒酶逆转录酶(hTERT)、RU1、RU2(AS)、肠羧基酯酶、mut hsp70-2、M-CSF、前列腺酶(prostase)、前列腺特异性抗原(PSA)、PAP、NY-ESO-1、LAGE-1a、p53、prostein、PSMA、Her2/neu、存活素和端粒末端转移酶、前列腺癌肿瘤抗原-1(PCTA-1)、黑色素瘤相关抗原(MAGE)、ELF2M、中性白细胞弹性蛋白酶、ephrinB2和CD22。CH2或CH3结构域分子还能够结合任何癌症相关蛋白,如IGF-I、IGF-II、IGR-IR或间皮素(mesothelin)。另外的肿瘤相关抗原在以下表3中提供。

[0252] 表3:示例性肿瘤及其肿瘤抗原

肿瘤	肿瘤相关的靶抗原
急性骨髓性白血病	Wilms 瘤 1 (WT1)、优先表达的黑色素瘤抗原 (PRAME)、PR1、蛋白酶 3、弹性蛋白酶、组织蛋白酶 G
慢性骨髓性白血病	WT1、PRAME、PR1、蛋白酶 3、弹性蛋白酶、组织蛋白酶 G
骨髓异常增生综合征	WT1、PRAME、PR1、蛋白酶 3、弹性蛋白酶、组织蛋白酶 G
急性成淋巴细胞性白血病	PRAME
慢性淋巴细胞性白血病	存活素
[0253] 非何杰金氏淋巴瘤	存活素
多发性骨髓瘤	纽约食管 1 (New York esophageous 1) (NY-Eso1)
恶性黑色素瘤	MAGE、MART、酪氨酸酶、PRAME GP100
乳腺癌	WT1、赫赛汀 (herceptin)
肺癌	WT1
前列腺癌	列腺特异性抗原(PSA)
结肠癌	癌胚抗原(CEA)
肾细胞癌 (RCC)	成纤维细胞生长因子 5 (FGF-5)

[0254] 在一些实施方案中,自身免疫疾病是类风湿性关节炎、幼年型少关节型关节炎、胶原诱导性关节炎、佐剂诱导性关节炎、干燥综合症、多发性硬化症、实验性自身免疫脑脊髓炎、炎性肠病(例如克罗恩病、溃疡性结肠炎)、自身免疫性胃萎缩、寻常天疱疮、牛皮癣、白癜风、1型糖尿病、非肥胖型糖尿病、重症肌无力、Grave病、桥本氏甲状腺炎、硬化性胆管

炎、硬化性涎腺炎、系统性红斑狼疮、自身免疫性血小板减少性紫癜、肺出血 - 肾炎综合征、Addison 病、全身性硬化症、多发性肌炎、皮炎、自身免疫性溶血性贫血或恶性贫血。

[0255] CH2 和 CH3 结构域分子的广泛效用至少部分上是因为它们尺寸小,这允许有效渗透组织,包括实体肿瘤和淋巴样组织(其中 HIV 复制),并且还允许有效中和快速进化以避免被宿主免疫系统所产生的免疫球蛋白中和的病毒(例如 HIV)。由于工程化 CH2 或 CH3 结构域分子易于产生针对目的抗原的高亲和力结合抗体,所以它们对于治疗也是有用的。此外,如本文所述,CH2 或 CH3 结构域分子还可以包含具有治疗特性的效应物分子(诸如例如,药物、酶或毒素)。

[0256] 如本文所述,CH2 或 CH3 结构域分子可以被工程化以包含来自对病原体(诸如 HIV)具有特异性的抗体的一个或多个 CDR。X5 是对 HIV-1 具有特异性的中和抗体(Moulard 等人,Proc. Natl. Acad. Sci. U. S. A. 99 :6913-6918,2002)。当从完整的免疫球蛋白(IgG1)或 Fab 转换为 scFv 抗体时,X5 的中和活性已经显示被显著地提高,所述 scFv 抗体只含有重链和轻链的可变结构域(Labrijn 等人,J. Virol. 77 :10557-10565,2003)。相信这种效果是由于对进入 X5 表位的尺寸限制。CH2 和 CH3 结构域分子小于 scFv 抗体,产生以下假设:工程化 CH2 结构域分子(包含一个或多个 X5 CDR)因其进入表位的能力而将具有增强的中和能力。

[0257] CH2 和 CH3 结构域分子通常作为包含一种或多种药学上可接受的载体的组合物施用给受试者。此类载体由施用的具体组合物以及由用来施用该组合物的具体方法来部分地决定。因此,存在种类广泛的适宜的本公开内容的药物组合物的制剂。

[0258] 用于肠胃外施用的制品包括无菌的水或非水的溶液、悬浮液和乳剂。非水溶剂的实例是丙二醇、聚乙二醇、植物油(如橄榄油)以及可注射的有机酯(如油酸乙酯)。含水的载体包括水、醇/水的溶液、乳剂或悬浮液,包括盐水和缓冲介质。肠胃外介质包括氯化钠溶液、林格氏右旋糖、右旋糖和氯化钠、乳酸钠林格氏溶液或不挥发性油。静脉内介质包括流体和营养补充剂、电解质补充剂(例如基于林格氏右旋糖的电解质补充剂)等。防腐剂和添加剂也可以存在,诸如例如抗微生物剂、抗氧化剂、螯合剂和惰性气体等。

[0259] 用于局部施用的制剂可以包括软膏、洗液、乳膏、凝胶、滴剂、栓剂、喷雾剂、液体和粉末。常规的药物载体、水性、粉末或油性基底、增稠剂等可能是必要的或令人期望的。

[0260] 用于口服施用的组合物包括粉末或颗粒、在水或非水介质中的悬浮液或溶液、胶囊、小药囊或片剂。增稠剂、调味剂、稀释剂、乳化剂、分散助剂或粘合剂可能是期望的。

[0261] 所述组合物中的一些可能作为药学上可接受的酸加成盐或碱加成盐来施用,所述药学上可接受的酸加成盐或碱加成盐是通过与无机酸和有机酸反应或通过与无机碱和有机碱反应而形成,所述无机酸例如盐酸、氢溴酸、高氯酸、硝酸、硫氰酸、硫酸和磷酸,所述有机酸例如甲酸、乙酸、丙酸、乙醇酸、乳酸、丙酮酸、草酸、丙二酸、琥珀酸、马来酸和富马酸,所述无机碱例如氢氧化钠、氢氧化铵、氢氧化钾,所述有机碱例如单烷基、二烷基、三烷基胺和芳基胺以及取代的乙醇胺。

[0262] 施用可以通过单剂量或多剂量来完成。所需的剂量在受试者之间将不同,这取决于受试者的物种、年龄、重量、一般状况、要治疗的具体的出血性疾患或发作、所使用的具体 CH2 或 CH3 结构域分子及其施用方式。适当的剂量可以通过本领域普通技术人员仅使用常规实验来确定。

[0263] 本文提供了包括治疗有效量的单独的或与药学上可接受的载体组合的工程化 CH2 或 CH3 结构域分子的药物组合物。药学上可接受的载体包括但不限于盐水、缓冲盐水、右旋糖、水、甘油、乙醇以及它们的组合。载体和组合物可以是无菌的,并且制剂适合于施用方式。组合物还可以含有少量的湿润剂或乳化剂或 pH 缓冲剂。组合物可以是液体溶液、悬浮液、乳剂、片剂、丸剂、胶囊、缓释制剂或粉末。可以将组合物与传统的粘合剂和载体(甘油三酯)一起配制为栓剂。口服制剂可以包括标准载体,例如药用级别的甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素和碳酸镁。可以使用任何常见的药物载体,例如无菌盐水溶液或麻油。所述介质还可以含有常规的药物辅料,诸如例如:调整渗透压的药学上可接受的盐、缓冲液、防腐剂等。可与本文提供的组合物和方法一起使用的其他介质是生理盐水和麻油。

[0264] VIII. 抗体恒定结构域分子用于检测的用途

[0265] 确定多肽存在或不存在的方法是本领域公知的。例如,可以将特异性结合剂例如 CH2 结构域分子与其他化合物结合,所述其他化合物包括但不限于酶、磁珠、胶体磁珠、半抗原、荧光染料、金属化合物、放射性化合物或药物。CH2 或 CH3 结构域分子还可以用在免疫测定中,所述免疫测定例如但不限于放射性免疫测定(RIA)、酶联免疫吸附测定(ELISA)、免疫组织化学测定、蛋白质印迹或免疫沉淀测定。这些测定法是本领域公知的(参见 Harlow&Lane, Antibodies, A Laboratory Manual (抗体, 实验室手册), Cold Spring Harbor Publications, New York (1988) 的免疫测定形式的描述)。

[0266] 在一个实施方案中,提供了包含免疫测定的诊断试剂盒。尽管免疫测定的细节可以随着采用的具体形式变化,但是用于检测生物样品中抗原的方法一般包括使该生物样品接触 CH2 或 CH3 结构域分子的步骤,所述 CH2 或 CH3 结构域分子在免疫反应条件下与目的抗原特异性反应。允许 CH2 或 CH3 结构域分子在免疫反应条件下特异性结合以形成免疫复合物,并且直接或间接地检测该免疫复合物(结合的抗原)的存在。

[0267] 本文所公开的 CH2 或 CH3 结构域分子还可以用于荧光激活细胞分选术(FACS)。在其他更为复杂的检测水平之外, FACS 测定采用多个颜色通道、低角度和钝光散射检测通道以及阻抗通道,以分离或分选细胞(参见美国专利第 5,061,620 号)。FACS 可以用来通过使该细胞与适当标记的 CH2 或 CH3 结构域分子接触分选抗原阳性的细胞。然而,可以采用不同效力的其他技术来纯化和分离期望的细胞群体。所采用的分离技术应该使保留收集的细胞级分的存活率最大化。所采用的具体技术将当然取决于分离的效率、方法的细胞毒性、分离的容易性和速度以及需要的设备和/或技术技能。

[0268] 另外的分离程序可以包括使用 CH2 或 CH3 结构域分子包被的磁珠的磁性分离、亲和色谱、与 CH2 或 CH3 结构域分子接合或与补体结合使用的细胞毒性剂、以及利用连接至固体基质的抗体或者 CH2 或 CH3 结构域分子的“淘选”,或者另外的便利技术。特异性结合剂与磁珠和其他固体基质(例如琼脂糖珠、聚苯乙烯珠、中空纤维膜和塑料培养皿)的连接允许直接分离。可以通过简单地将固体支撑体从细胞悬浮液中物理分离来将由诸如 CH2 或 CH3 结构域分子的特异性结合剂结合的细胞从该细胞悬浮液中移除。细胞与固相连接的抗体或者 CH2 或 CH3 结构域分子一起孵育的精确条件和持续时间将取决于具体针对所采用的系统的若干因素。然而,适当条件的选择是本领域公知的。

[0269] 然后,可以在允许表达目的抗原的细胞与固相连接的结合剂结合足够时间之后,用生理缓冲液将未结合的细胞洗脱或洗出。然后通过任何适当的方法将结合的细胞从固相

分离,并使用本领域公知的方法定量,所述适当的方法主要取决于固相的性质和所采用的抗体或 CH2 或 CH3 结构域分子。在一个具体的非限制性实例中,通过 FACS 定量从固相分离的结合细胞。

[0270] 如本领域所知,可以将 CH2 或 CH3 结构域分子与生物素或荧光染料结合,所述生物毒素可以用与支撑体结合的亲和素或链亲和素除去,荧光染料可以与 FACS 一起使用以能够进行细胞分离和定量。

[0271] CH2 或 CH3 结构域分子可以与其他化合物结合,所述其他化合物包括但不限于:酶、顺磁性珠、胶体顺磁性珠、半抗原、荧光染料、金属化合物、放射性化合物或药物。可与 CH2 或 CH3 结构域分子结合的酶包括但不限于碱性磷酸酶、过氧化物酶、尿素酶和 β -半乳糖苷酶。可以与 CH2 结构域分子结合的荧光染料包括但不限于:异硫氰酸荧光素、异硫氰酸四甲基若丹明 (tetramethyl rhodamine isothiocyanate)、藻红蛋白、别藻蓝蛋白和德克萨斯红。对于可以与抗体结合的其他荧光染料参见 Haugland, R. P., *Molecular Probes: Handbook of Fluorescent Probes and Research Chemicals* (分子探针:荧光探针和研究化学品手册) (1992-1994)。可与 CH2 或 CH3 结构域分子结合的金属化合物包括但不限于铁蛋白、胶体金及特别是胶体超顺磁性珠。可与 CH2 或 CH3 结构域分子结合的半抗原包括但不限于生物素、异羟基洋地黄毒甙元、噁唑酮 (oxazalone) 和硝基苯酚。其他试剂是本领域公知的。

[0272] IX. 抗体恒定结构域分子的效应物功能

[0273] 工程化 CH2 或 CH3 结构域能够结合 Fc 受体和 / 或补体相关分子,如 C1q,这允许多种效应物功能,包括抗体依赖性细胞介导的细胞毒性 (ADCC)、补体依赖性细胞毒性 (CDC)、吞噬作用、调理作用和调理吞噬。在一些实施方案中,本文所述的 CH2 或 CH3 结构域分子包含一个或多个 Fc 受体的结合位点,因此能够使得这些分子介导多种效应物功能 (参见以下表 4)。如果效应物功能是不期望的,那么可以诱变一个或多个 Fc 结合位点以阻止这些功能。

[0274] 抗体-抗原复合物与免疫系统的细胞相互作用导致广泛的反应,包括多种效应物功能和免疫调节信号。通过抗体的 Fc 结构域或免疫复合物与特定的细胞表面受体 Fc 结构域的结合启动这些相互作用。Fc 受体家族的每个成员通过 Fc 结构域上的识别结构域来识别一种或多种同种型的免疫球蛋白。Fc 受体是通过其对免疫球蛋白亚型的特异性来确定 (例如, IgG 的 Fc 受体被称为 Fc γ R) (美国授予前公开号 2006-0134709)。

[0275] Fc 受体是在一些免疫系统细胞的表面上发现的糖蛋白,所述免疫系统细胞包括单核细胞、巨噬细胞、嗜中性细胞、嗜酸性细胞、肥大细胞、自然杀伤细胞、B 细胞和树突状细胞。Fc 受体表现出多种细胞表达形式和效应物功能 (参见表 4)。Fc 受体允许免疫细胞与连接至微生物或微生物感染细胞的表面的抗体结合,帮助这些细胞鉴定和消除微生物病原体。Fc 受体在它们的 Fc 区域结合抗体,即为激活拥有 Fc 受体的细胞的相互作用。

[0276] 表 4

[0277] Fc 受体的细胞分布和效应物功能

[0278]

受体名称	细胞分布	效应物功能
FcγRI (CD64)	巨噬细胞 嗜中性细胞 嗜酸性细胞 树突状细胞	吞噬作用 细胞活化 激活呼吸爆发 诱导微生物杀伤
FcγRIIA (CD32)	巨噬细胞 嗜中性细胞 嗜酸性细胞 血小板 朗格汉斯细胞	吞噬作用 脱粒（嗜酸性细胞）
FcγRIIB1 (CD32)	B 细胞 肥大细胞	无吞噬作用 抑制细胞活性
FcγRIIB2 (CD32)	巨噬细胞 嗜中性细胞 嗜酸性细胞	吞噬作用 抑制细胞活性
FcγRIIIA (CD16a)	NK 细胞	诱导 ADCC

[0279]

受体名称	细胞分布	效应物功能
FcγRIIIB (CD16b)	嗜酸性细胞 巨噬细胞 嗜中性细胞 肥大细胞 囊泡树突状细胞	诱导微生物杀伤
FcϵRI	肥大细胞 嗜酸性细胞 嗜碱性细胞 朗格汉斯细胞	脱粒
FcϵRII (CD23)	B 细胞 嗜酸性细胞 朗格汉斯细胞	可能的粘附分子
FcαRI (CD89)	单核细胞 巨噬细胞 嗜中性细胞 嗜酸性细胞	吞噬作用 诱导微生物杀伤
Fcα/μR	B 细胞 系膜细胞 巨噬细胞	内吞作用 诱导微生物杀伤
FcRn	单核细胞 巨噬细胞 树突状细胞 上皮细胞 内皮细胞 肝细胞	通过胎盘将 IgG 从母亲转移至胎儿 以母乳将 IgG 从母亲转移至婴儿 保护 IgG 免于降解

[0280] 吞噬细胞的活化是Fc受体引起的最常见功能。例如,巨噬细胞通过其Fc γ 受体的衔接后的吞噬作用开始摄入和杀伤覆有IgG的病原体。另一涉及Fc受体的过程称作抗体依赖性细胞介导的细胞毒性(ADCC)。在ADCC期间,在自然杀伤(NK)细胞表面上的Fc γ RIII受体刺激NK细胞以便从NK细胞的颗粒中释放细胞毒性分子来杀伤抗体覆盖的靶细胞。然

而, $\text{Fc}\epsilon\text{RI}$ 具有不同的功能。 $\text{Fc}\epsilon\text{RI}$ 是粒细胞上参与变态反应和防御寄生虫感染的 Fc 受体。当适当的变应抗原或寄生虫存在时, IgE 分子和它们在粒细胞表面上的 Fc 受体中的至少两个交联将触发该细胞从其颗粒中快速释放形成的介质。

[0281] 此外, IgG 和 IgM 抗体的 Fc 结构域能够结合 C1q , C1q 是补体激活经典途径的组分。当 IgG 或 IgM 抗体与病原体的表面结合时, C1q 能够结合其 Fc 区域, 这启动了补体级联, 最终导致炎症细胞的募集和对病原体的调理及杀伤。

[0282] 为了进一步向 CH2 或 CH3 结构域分子提供官能性, 可以使用本领域技术人员已知的许多的手段将效应物分子 (例如, 治疗、诊断或检测的部分) 连接至 CH2 或 CH3 结构域分子。示例性效应物分子包括但不限于放射性标记、荧光标志物或毒素。可以使用共价和非共价的连接手段。用于将效应物分子连接至抗体的程序根据效应物的化学结构而改变。多肽通常含有各种官能团, 例如羧酸 (COOH)、游离胺 ($-\text{NH}_2$) 或巯基 ($-\text{SH}$) 基团, 所述官能团可用于与抗体上适合的官能团反应以导致效应物分子的结合。或者, 使抗体衍生以暴露或连接另外的反应官能团。这种衍生作用可能涉及许多接头分子中的任一种的连接, 例如可从 Pierce Chemical Company, Rockford, IL 获得的那些接头分子。接头可以是用来接合抗体与效应物分子的任何分子。接头能够与抗体和效应物分子形成共价键。适合的接头是本领域技术人员公知的并且包括但不限于直链或支链的碳接头、杂环碳接头或肽接头。当抗体和效应物分子是多肽时, 接头可以通过组成氨基酸的侧基 (例如通过二硫键与半胱氨酸连接) 与该氨基酸接合, 或者与末端氨基酸的 α 碳氨基和羧基接合。

[0283] 在一些情况下, 在免疫结合物到达其靶点时, 期望使效应物分子从抗体游离。因此, 在这些情况下, 免疫结合物将包含在靶点附近可切割的连接。切割接头以从抗体释放效应物分子可以通过免疫结合物在靶细胞内部或靶点附近经历的酶活性或条件来促进。

[0284] 考虑到已报道的用于将各种放射诊断化合物、放射治疗化合物、标记 (例如, 酶或荧光分子) 药物、毒素和其他剂连接至抗体的大量方法, 本领域技术人员将能够确定用于将给定的剂连结至 CH2 或 CH3 结构域分子的适合方法。

[0285] 治疗剂包括各种药物 (例如长春碱、道诺霉素等)、以及诸如细胞毒素 (例如天然或改性的假单胞菌外毒素或白喉毒素) 的效应物分子、本身含有药理学组合物的囊封剂 (例如脂质体)、靶部分和配体。具体治疗剂的选择取决于具体的靶分子或靶细胞以及期望引发的生物效应。因此, 例如, 治疗剂可以是用来导致特定的靶细胞死亡的具有细胞毒性的效应物分子。相反地, 如果只希望引发非致死性生物反应, 可以将治疗剂与非致死性药理剂或含有非致死性药理剂的脂质体结合。

[0286] 可以将毒素与用作免疫毒素的 CH2 或 CH3 结构域分子一起使用。示例性毒素包括假单胞菌外毒素 (PE)、蓖麻毒蛋白、相思豆毒素、白喉毒素和它们的亚基、核毒素 (ribotoxin)、核糖核酸酶、皂草素和刺孢霉素以及肉毒杆菌毒素 A 至 F。这些毒素是本领域公知的并且许多可以容易地从商业来源获得 (例如 Sigma Chemical Company, St. Louis, MO)。

[0287] 白喉毒素是从白喉棒杆菌分离的。通常, 对免疫毒素中使用的白喉毒素突变以减少或消除非特异性毒性。自从 20 世纪七十年代为人们了解了 (Laird 和 Groman, J. Virol. 19:220, 1976) 称为 CRM107 的突变体, 并且被用在人临床试验中, 所述突变体具有完全的酶活性但具有显著减少的非特异性毒性。参见美国专利第 5, 792, 458 号和美国专利

第 5,208,021 号。如本文所用,术语“白喉毒素”适当是指天然白喉毒素或指保留了酶活性但已被修饰以减少非特异性毒性的白喉毒素。

[0288] 蓖麻毒蛋白是来自蓖麻 (*Ricinus communis*) 的凝集素 RCA60。术语“蓖麻毒蛋白”也指其毒素变体。例如,参见美国专利第 5,079,163 号和美国专利第 4,689,401 号。蓖麻凝集素 (RCA) 以两种形式存在,这两种形式根据其大约 65kD 和 120kD 的分子量分别被称为 RCA₆₀ 和 RCA₁₂₀ (Nicholson & Blaustein, *J. Biochim. Biophys. Acta* 266 :543, 1972)。A 链负责使蛋白合成失活和杀伤细胞。B 链使蓖麻毒蛋白与细胞表面半乳糖残基结合并促进 A 链运输到胞质溶胶中 (Olsnes 等人., *Nature* 249 :627-631, 1974 和美国专利第 3,060,165 号)。

[0289] 已经将核糖核酸酶与靶分子结合用作免疫毒素 (参见 Suzuki 等人., *Nat. Biotech.* 17 :265-70, 1999)。示例性核毒素例如 α -八叠球菌和局限曲菌素在例如 Rathore 等人., *Gene* 190 :31-5, 1997; 以及 Goyal 和 Batra, *Biochem* 345 Pt. 2 :247-54, 2000 中讨论。刺孢霉素首先从棘孢小单孢菌 (*Micromonospora echinospora*) 中分离,并且是引起导致凋亡的 DNA 中的双链断裂烯二炔抗肿瘤抗生素家族的成员 (参见例如, Lee 等人., *J. Antibiot* 42 :1070-87, 1989)。该药物是临床试验中免疫毒素的毒性部分 (参见例如, Gillespie 等人., *Ann Oncol* 11 :735-41, 2000)。

[0290] 相思豆毒素包括来自相思豆 (*Abrus precatorius*) 的毒性凝集素。毒性成分相思豆毒素 a、b、c 和 d 具有约 63kD 至 67kD 的分子量,并且由两个二硫键连接的多肽链 A 和 B 组成。A 链抑制蛋白合成; B 链 (相思豆毒素 -b) 与 D-半乳糖残基结合 (参见 Funatsu 等人., *Agr. Biol. Chem.* 52 :1095, 1988; 和 Olsnes, *Methods Enzymol.* 50 :330-335, 1978)。

[0291] 提供以下实施例来说明某些具体的特征和/或实施方案。这些实施例不应该被解释为将本发明限制于所述的具体特征或实施方案。

实施例

[0292] 实施例 1: 抗体 CH2 结构域的文库的产生,其中所述抗体 CH2 结构域带有含有氨基酸残基随机突变为四个残基 Y、S、A 或 D 中的任何一个的环

[0293] 在这个实施例中,构建了突变的 CH2 结构域,其中环 1 被 10 个随机安排的 Y、S、A 或 D 残基加上该环的 C 端末端的另外一个 G 置换。类似地,环 3 被 6 个随机排列的 Y、S、A 或 D 残基加上该环的 C 端末端的另外的 G 置换。分三步产生 DNA 文库。

[0294] 首先,使用 CH2DNA 来产生分别含有突变的环 1 和环 2 的两个片段:片段 1 和片段 2。片段 1 是使用 N 端引物 (5' GCA CTG GCT GGTTC GCT ACC GT GGCC CAGGC GGCC GCA CCT GAA CTC CTG 3'; SEQ ID NO:6) 和环 1 反向引物 (5' CAC GTA CCA GTT GAA CTT GCCAKM AKM AKM AKM AKM AKM AKM AKM AKM CAC CACCAC GCA TGT GAC 3'; SEQ ID NO:7) 通过 PCR 扩增产生,其中 K = G 或 T,并且 M = A 或 C。片段 2 是通过使用环 1 正向引物 (5' AAG TTCAAC TGG TAC GTG 3'; SEQ ID NO:8) 和环 3 反向引物 (5' GAT GGTTTT CTC GAT GGG GCC AKM AKM AKM AKM AKM AKM GTTGA GAC CTT GCA CTT G 3'; SEQ ID NO:9) 产生的。

[0295] 第二,通过使用重叠延伸 (SOE) PCR 的剪接使这两个片段接合。在 SOE PCR 的第二步骤期间,除了该 N 端引物外还使用 C 端引物 (5' GGT GCA GAA GAT GGT GGT GGCC GGCCT

GGCC TTT GGC TTTGGA GAT GGT TTT CTC GAT G 3' ;SEQ ID NO :10) 以便在 DNA 的两端引入下一克隆步骤所需的限制位点 Sfi1。

[0296] 第三, 扩增的突变 CH2 片段用 Sfi1 消化并且将其连入用相同的酶消化的噬菌粒载体中。使用 Amicon Ultra-4 centricon 通过用双蒸水洗涤三次对连接产物脱盐, 然后通过电穿孔转化 TG1 细胞。

[0297] 以下显示了来自转化的 TG1 细胞的 20 个随机选择的克隆序列 (表 5), 显示通过四个残基 Y、S、D 和 A 成功产生了具有随机的环 1 和环 3 的 CH2 突变体。

[0298] 表 5

[0299] 具有随机的环 1 和环 3 的突变 CH2 序列的片段

[0300] (X2-22 表示克隆的名称)

[0301] 环 1

[0302] x9 PEVTCVVV YYDSAAAYAY GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0303] (SEQ ID NO :11)

[0304] x14 PEVTCVVV YYSASAAASA GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0305] (SEQ ID NO :12)

[0306] x13 PEVTCVVV YDSYASSDD GKFNWYVDG VEVHNAKTKPRKEQYNSTYR

[0307] (SEQ ID NO :13)

[0308] x15 PEVTCVVV AYSDDAAYD GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0309] (SEQ ID NO :14)

[0310] x10 PEVTCVVV DADDYYYYY GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0311] (SEQ ID NO :15)

[0312] x2 PEVTCVVV DAYYDADYYY GKFNWYVDGVEVHNAKTKP REEQYNSTYR

[0313] (SEQ ID NO :16)

[0314] x11 PEVTCVVV DAAYDYSY GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0315] (SEQ ID NO :17)

[0316] x19 PEVTCVVV DYSDDAYAD GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0317] (SEQ ID NO :18)

[0318] x16 PEVTCVVV SYSDSYSYSA GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0319] (SEQ ID NO :19)

[0320] x4 PEVTCVVV DAYADDASA GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0321] (SEQ ID NO :20)

[0322] x17 PEVTCVVV SYSDSDYDD GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0323] (SEQ ID NO :21)

[0324] x12 PEVTCVVV DDSYYSYDD GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0325] (SEQ ID NO :22)

[0326] x22 PEVTCVVV YDASDYADAY GKFNWYVDG VEVHNAKTKPREEQYNSTYR

[0327] (SEQ ID NO :23)

[0328] x8 PEVTCVVV ADAAAYAYAD GKFNWYVDGVEVHNAKTKP REEQYNSTYR

[0329] (SEQ ID NO :24)

- [0330] x7 PEVTCVVV ASDSSDDYD GKFNWYVDG VEVHNAKTKPREEQYNSTYR
- [0331] (SEQ ID NO :25)
- [0332] x5 PEVTCVVV AAAAADADYY SKFNWYVDG VEVHNAKTKPREEQYNSTYR
- [0333] (SEQ ID NO :26)
- [0334] x20 PEVTCVVV YDDAAYADDY GKFNWYVDG VEVHNAKTKPREEQYNSTYR
- [0335] (SEQ ID NO :27)
- [0336] x21 PEVTCVVV SADASDYD GKFNWYVDG VEVHNADTKPREEQYNSTYR
- [0337] (SEQ ID NO :28)
- [0338] x23 PEVTCVVV DDDAADAYYY GKFNWYVDG VEVHNAKTKPREEQYNSTYR
- [0339] (SEQ ID NO :29)
- [0340] x3 PEVTCVVV YSDDDYDYA GKFCWYVDG VEVHNAKTKPREEHYNSTYR
- [0341] (SEQ ID NO :30)
- [0342] 环 3
- [0343] x9 VVSVLTVLHQ DWLNGKEYKC KVSN AASAYSGPIEKTISKA K(SEQ ID NO :31)
- [0344] x14 VVSVLTVLHQ DWLNGKEYKC KVSN ADDADAGPIEKTISKA K(SEQ ID NO :32)
- [0345] x13 VVSVLTVLHQ DWLNGKEYKC KVSN AADAYAGPIEKTISKA K(SEQ ID NO :33)
- [0346] x15 VVSVLTVLHQ DWLNGKEYKC KVSN AADYSDGPIEKTISKA K(SEQ ID NO :34)
- [0347] x10 VVSVLTVLHQ DWLNGKEYKC KVSN AADAADGPIEKTISKA K(SEQ ID NO :35)
- [0348] x2 VVSVLTVLHQ DWLNGKEYKC KVSN DASASSGPIEKTISKA K(SEQ ID NO :36)
- [0349] x11 VVSVLTVLHQ DWLNGKEYKC KVSN DDYAASGPIEKTISKA K(SEQ ID NO :37)
- [0350] x19 VVSVLTVLHQ DWLNGKEYKC KVSN DAYASDGPIEKTISKA K(SEQ ID NO :38)
- [0351] x16 VVSVLTVLHQ DWLNGKEYKC KVSN DADDASGPIEKTISKA K(SEQ ID NO :39)
- [0352] x4 VVSVLTVLHQ DWLNGKEYKC KVSN AADDDSGPIEKTISKA K(SEQ ID NO :40)
- [0353] x17 VVSVLTVLHQ DWLNGKEYKC KVSN ADAYAYGPIEKTISKA K(SEQ ID NO :41)
- [0354] x12 VVSVLTVLHQ DWLNGKEYKC KVSN ADDYDYGPIEKTISKA K(SEQ ID NO :42)
- [0355] x22 VVSVLTVLHQ DWLNGKEYKC KVSN YSDSAAGPIEKTISKA K(SEQ ID NO :43)
- [0356] x8 VVSVLTVLHQ DWLNGKEYKC KVSN YAASAYGPIEKTISKA K(SEQ ID NO :44)
- [0357] x7 VVSVLTVLHQ DWLNGKEYKC KVSN YDDDADGPIEKTISKA K(SEQ ID NO :45)
- [0358] x5 VVSVLTVLHQ DWLNGKEYKC KVSN YYDYDYGPIEKTISKA K(SEQ ID NO :46)
- [0359] x20 VVSVLTVLHH DWMNGKEYKC EVSN DADSADGPIKKTISKA K(SEQ ID NO :47)
- [0360] x21 VVSVLTVLHH DWLNGEYKC KVSN DASDDAGPIEKTIS. A K(SEQ ID NO :48)
- [0361] x23 VVSVLTVLHQ DWLNGKEYKC KVSN ADDAYAGPIEKTISKA K(SEQ ID NO :49)
- [0362] x3 VVSVLTVLHH YWMNGEDYKC EVSN DSYSDDGPIKKTISKA K(SEQ ID NO :50)
- [0363] 实施例 2 :来自人抗体的 CDR3 移入到 CH2 支架中
- [0364] 在这个实施例中,通过置换环 A-B 和环 E-F 将来自抗体文库的人 VH CDR3(H3) 移入到 CH2 中。首先,使用五次 PCR 由 H3 置换环 A-B。首先两次 PCR 通过使用以下引物产生了没有环 A-B 的两个 CH2 片段:对于片段 1- 正向 1 引物 (5' TAG CGA TTC GCT ACC GTG GCCAG GCG GCC CCT GAA CTC CTG GGG GGA CC 3' ;SEQ ID NO ;51) 和反向 1 引物 (5'TCC CCC CAG GAG TTC AGG TGC 3' ;SEQ IDNO ;52),对于片段 2- 正向 2 引物 (5' TGC GTG GTG

GTG GAC GTGAGC 3';SEQ ID NO:53) 和反向 2 引物 (5' TAG GCA TGC ATC TGCATG GTG GCC GGC CTG GCC TTT GGC TTT GGA GAT GGT TTTCTC GAT GG 3' ;SEQ ID NO:54)。正向 1 引物和反向 2 引物含有终产物的 N 端和 C 端所需的 SfiI 的限制位点。反向 1 引物和正向 2 引物含有随后的 SOE PCR 所需的末端序列。第三次 PCR 使用作为模板的抗体 VH 文库和三种引物的两种混合物,其各自被设计用于扩增不同的 H3。正向引物的混合物含有 H3 正向引物 1:5' GAA CTC CTG GGG GGACCG GCY AYR TAT TAC TGT GYG 3' (SEQ ID NO:55)、H3 正向引物 2:5' GAA CTC CTG GGG GGA CCG GCY TTR TAT TAC TGT GYG 3' (SEQ ID NO:56), 以及 H3 正向引物 3:5' GAA CTC CTG GGG GGACCG GCY GTR TAT TAC TGT GYG 3' (SEQ ID NO:57)。反向引物的混合物含有 H3 反向引物 1:5'GCT CAC GTC CAC CAC CAC GCAGGT GCC CTG GCC CCA 3' (SEQ ID NO:58)、H3 反向引物 2:5'GCTCAC GTC CAC CAC CAC GCA GGT GCC ACG GCC CCA 3' (SEQ IDNO:59), 以及 H3 反向引物 3:5'GCT CAC GTC CAC CAC CAC GCAGGT GCC AYG GCC CCA 3' (SEQ ID NO:60)。第三次 PCR 产生了含有具有末端序列 H3 的片段的混合物,所述末端序列被设计为与反向 1 引物和正向 2 引物的对应末端序列相重叠。将这两个 CH2 片段和含有 H3 的片段用作 SOE PCR 中的引物和模板以产生其中环 AB 被 H3 置换的片段。通过使用正向 1 引物和反向 2 引物扩增这种片段混合物。扩增的片段用 SfiI 消化并且将其连入用相同的酶消化的噬菌粒载体 (pComb3X 或 pZUD) 中。使用 Amicon Ultra-4 离心超滤管通过用双蒸水洗涤三次对连接产物进行脱盐,然后通过电穿孔转化 TG1 细胞。

[0365] 可以将类似程序用于置换环 E-F,除外的是:对于扩增片段 1,使用另一引物 - 反向引物 12(5' GTA CGT GCT GTT GTA CTG CTC 3';SEQ ID NO:61) 取代反向引物 1;对于扩增片段 2,使用另一引物 - 正向引物 22(5' AAG GTC TCC AAC AAA GCC CTC 3';SEQ ID NO:62) 取代正向引物 2;并且对于扩增 H3, H3 引物是不同的。在这样的情况下,正向引物的混合物含有 H3 正向引物 12:5' GAG CAG TAC AACAGC ACG TAC GCA GCY AYR TAT TAC TGT GYG 3' (SEQ ID NO:63), H3 正向引物 22:5' GAG CAG TAC AAC AGC ACG TAC GCAGCY TTR TAT TAC TGT GYG 3' (SEQ ID NO:64), 以及 H3 正向引物 32:5' GAG CAG TAC AAC AGC ACG TAC GCA GCY GTR TAT TACTGT GYG 3' (SEQ ID NO:65)。在这样的情况下反向引物的混合物含有 H3 反向引物 12:5' GAG GGC TTT GTT GGA GAC CTT GGT TCC CTGGCC CCA 3' (SEQ ID NO:66), H3 反向引物 22:GAG GGC TTT GTTGGA GAC CTT GGT GCC ACG GCC CCA 3' (SEQ ID NO:67), 以及 H3 反向引物 32:5' GAG GGC TTT GTT GGA GAC CTT GGT GCCAYG GCC CCA 3' (SEQ ID NO:68)。最后,两个环 A-B 和 E-F 都可以用 VH H3 置换。在这样的情况下,在环 A-B 被 H3 置换后,在得到的片段中环 E-F 被随机重组的 H3 置换。

[0366] 以下示出了具有被 H3 置换的两个环的从转化的 TG1 细胞随机选择的 19 个克隆的序列 (表 6),表明 H3 的成功移植。图 4 示出这些克隆中若干克隆的蛋白表达。突变分子的条带位置以箭头指示。

[0367] 表 6

[0368] 具有植入的 H3 的片段

[0369] #2-38(分别为 SEQ ID NO:69-87) 表示克隆的名称

[0370] H3 H3

[0371] #2: AVYYCV. KVPVGY. WGRGT & AVYYCA. DVEASSPADFGY. WGRGT

[0372] #3: AMYYCA. RDHGVD TAMAGPWFYD. . WGRGT & AVYYCV. RGTGWELLVIDC. WGRGT

[0373] #6 :AVYYCA. RGSSGWGFDP. WGQGT & ATYYCA. RDRGY. WGRGT
[0374] #8 :AVYYCA. RRMPEGDSSGTSYYFDY. WGQGT & ALYYCA. REEKGDYDY. WGQGT
[0375] #13 :AMYYCA. IHSFDY. WGQGT & AVYYCA. KVLSGWFDHYFDS. ... WGQGT
[0376] #15 :AVYYCA. RDRVPDGVWSADS. WGQGT & AVYYCA. SKPPVSNWFDP. WGQGT
[0377] #16 :AIYYCV. KAGYNFDAFDH. WGRGT & AMYYCA. GDTAMVIFDY. WGQGT
[0378] #17 :ATYYCA. SGSSGCSY. WGQGT & ATYYCA. RGGYSSGWYHWYFDL. WGRGT
[0379] #22 :AVYYCA. ASVGAPSDFDY. WGQGT & ATYYCA. TTPDSNYGY. WGQGT
[0380] #23 :ALYYCA. KGQYGDHDY. WGQGT & AVYYCA. KEEEGAVLG. WGRGT
[0381] #25 :ATYYCA. REGTVVTPYFVY. WGQGT & AVYYCA. MGGHGSGSYLSGY. ... WGQGT
[0382] #26 :AVYYCA. RERYGALDY. WGRGT & AVYYCA. GLLHEGSGY. WGQGT
[0383] #28 :AIYYCA. ARGQGNSWWFDP. WGQGT & AIYYCA. TQVGHGD. WGQGT
[0384] #30 :ALYYCA. RAYSAYQYSFDS. WGRGT & AVYYCA. RREYNWNHNWFDP. ... WGQGT
[0385] #31 :ATYYCA. RRGGDYGDYFFDY. WGQGT & AIYYCA. RSRGSSFDY. WGQGT
[0386] #33 :AMYYCA. RDLYSNYVDY. WGQGT & AVYYCA. RGPWQQLVNWFPD. ... WGQGT
[0387] #34 :ATYYCA. SLTGTTSY. WGQGT & ALYYCA. RATWGYQFDC. WGQGT
[0388] #36 :AIYYCA. RESSSSFDY. WGQGT & AVYYCA. RMSGGRWIFDH. WGQGT
[0389] #38 :AVYYCA. RGWELDY. WGQGT & AVYYCA. KTGQFDY. WGQGT

[0390] 实施例 3 :稳定的 CH2 突变体的工程化和表征

[0391] 在这个实施例中,鉴定出与亲代野生型 CH2 相比表现出提高的稳定性的两个 CH2 突变体。因为 CH2 框架已经被链 B 与链 F 之间的内部二硫键稳定,所以假设在其他链之间的额外的二硫键可以提供 CH2 稳定性的整体提高。链 A 和链 G 中的几个位置被突变,其中一个产生了非常稳定的突变 CH2,命名为 m01,其中 L(在序列 GPSVFLFPPKPKDTL (SEQ ID NO : 88) 中)和第一个 K(在序列 EKTISKAK (SEQ ID NO :89) 中)被突变成 C。另一个突变体,命名为 m02,其中 V(在 GPSVFLFPPKPKDTL (SEQ ID NO :90) 中)和 EKTISKAK (SEQ ID NO :91) 中的第一个 K 被突变成 C,与亲代 CH2 相比表现出稳定性提高,但是比 m01 的稳定性低。

[0392] 材料和方法

[0393] CH2 结构域的克隆、表达和纯化。将人 γ 1CH2 克隆到细菌表达载体中并且用于转化大肠杆菌菌株 HB2151 细胞,使该细胞在 37℃ 下的 SB 培养基中生长至光密度为 $OD_{600} \sim 0.6-0.8$ 。用 1mM IPTG 在 37℃ 诱导表达持续 12-16 小时。收集细菌细胞并且重悬在 1 : 10(缓冲液 A 的体积 : 培养物体积)的缓冲液 A 中 (50mM Tris • Cl, 450mM NaCl, pH8.0)。将硫酸多粘菌素 B (Sigma-Aldrich, MO) (0.5mu/ml) 添加至该悬浮液中 (1 : 1000 硫酸多粘菌素 B 体积 : 培养物体积)。随后通过在 4℃ 以 15,000rpm 离心 45min 使细胞裂解物澄清,并且通过 SDS-PAGE 和 Western 测试其表达。通过使用 1ml 的 HiTrap Chelating HP Ni-NTA 柱 (GE Healthcare, NJ) 纯化澄清的上清液。在用缓冲液 B (50mM Tris • Cl, 450mM NaCl, 200mM 咪唑, pH 8.0) 洗脱后,通过 Amicon Ultra-15 离心式过滤设备 (MILLIPORE, MA) 除去咪唑,并且将纯化的蛋白保持在缓冲液 A 或 PBS (9.0g/L NaCl, 144mg/L KH_2PO_4 , 795mg/L Na_2HPO_4 , pH 7.4) 中。通过 SDS-PAGE 检查蛋白纯度并通过测量 UV 吸光度确定蛋白浓度。

[0394] CH2 突变体的设计和质粒构建。为了设计 CH2 突变体,使用 Fc 晶体结构。通过以计算机程序 VMD 18.6 分析结构来选择五个突变体 V10/E103 至 C10/C103、F11/K104 至 C11/

C104、L12/T105 至 C11/C105、L12/K104 至 C12/C104 以及 V10/K104 至 C10/C104 用于表征 (Humphrey 等人, J Mol. Graph. 14 :33-38, 1996)。它们是通过基于 PCR 的定点诱变制得并且被克隆到细菌表达载体中。通过直接测序证实这些克隆并将其用于转化大肠杆菌菌株 HB2151。以与野生型 CH2 相似地表达并纯化这些突变体。

[0395] 尺寸排阻色谱法。将纯化的 CH2、CH2m01 和 CH2m02 载入到运行在 ÄKTA BASIC pH/C 色谱系统 (GE Healthcare, NJ) 上的 Hiload26/60 Superdex 75HR 10/30 柱 (GE Healthcare, NJ) 中以评估寡聚体的形成。选择缓冲液 A 作为移动相。使用由醛缩酶 (158kDa)、牛血清白蛋白 (67kDa)、卵白蛋白 (44kDa)、糜蛋白酶原 A (25kDa) 以及核糖核酸酶 A (17kDa) 组成的凝胶过滤标准品来确定 CH2、CH2 m01 和 CH2 m02 的分子量。

[0396] 通过质谱确定二硫键。经 Voyager 4700 MALDI-TOF/TOF 质谱 (Applied Biosystems, CA), 通过比较在 (A) 所有 SH 基团的还原和烷基化和 (B) 原始的游离 SH 基团的烷基化而二硫键没有还原之后的分子量, 确定纯化的 CH2、CH2 m01 和 CH2 m02 中的二硫键的总数。还原使用 TCEP 进行, 并且烷基化用碘乙酰胺进行。

[0397] 圆二色性 (CD)。CH2、CH2 m01 和 CH2 m02 的二级结构通过圆二色性 (CD) 光谱法确定。将纯化的蛋白以 0.49mg/ml 的终浓度溶解在 PBS 中, 并且在 AVIV Model 202 CD 分光计 (Aviv Biomedical, NJ) 上记录 CD 光谱。使用用于天然结构测定的 0.1cm 光程的比色池在 25°C 记录波谱。通过记录 25-90°C 温度范围内 1°C/min 加热速率的 CD 信号测量 216nm 的热动力学稳定性。在加热后, 记录 90°C 的波谱。对于重折叠的评估, 将所有样品保持在 4°C 过夜并且在 25°C 再次测量。用外部探头传感器记录温度并且通过校准计算出微量比色池内部的温度 - 比通过外部传感器测量的温度低约 2-3°C (对于 20°C 至 80°C 的温度, 其范围是 1.9°C 至 3.8°C)。

[0398] 差示扫描量热法 (DSC)。用 VP-DSC 微热量计 (MicroCal, Northampton, MA) 进一步监测 CH2、CH2 m01 和 CH2 m02 的热稳定性。在 PBS (pH 7.4) 中三种蛋白的浓度为 1.5mg/ml。所采用的加热速率为 1°C/min 并且从 25°C 至 100°C 进行扫描。

[0399] 分光荧光法。在荧光计 Fluoromax-3 (HORIBA Jobin Yvon, NJ) 上记录了 CH2、m01 和 m02 的内荧光。使用 10 µg/ml 的蛋白浓度用 280nm 的激发波长进行内荧光测量, 并记录了 25°C 从 320nm 到 370nm 的发射谱。使用了存在 0mM 至 8mM 尿素的缓冲液 A。对于所有样品, 荧光光谱校正了溶液 (缓冲液 + 变性剂) 的背景荧光。将 340nm 的荧光强度用于解折叠评估。

[0400] 核磁共振 (NMR)。对于 NMR 实验, 使大肠杆菌首先生长在 2×YT 中。将单个菌落接种在 3mL 2×YT 中约 3 小时, 然后检查浑浊度并且将细菌转移至 1 升 2×YT 培养基中进一步在 37°C 生长直至达到 OD₆₀₀ ~ 0.8-0.9。然后将细胞培养物离心除去 2×YT 培养基并用 M9 基本培养基代替 2×YT 培养基, M9 基本培养基分别以 ¹⁵N NH₄Cl 和 ¹³C 葡萄糖作为唯一的 ¹⁵N 和 ¹³C 来源 (17)。将细胞在 30°C 孵育过夜, 并用 1mMIPTG 诱导。将收获的细胞悬浮在冰上的 TES 缓冲液 (1L 培养物 10mL 缓冲液) 中持续 1h。通过在冰上添加 1.5 体积 TES/5 持续 4 小时, 诱导释放周质蛋白的渗透压震扰。然后在 4°C 将上清液在透析缓冲液 (50mM Tris • Cl, 0.5M NaCl) 中透析过夜。通过以上所述的初始纯化的方法纯化蛋白。然后将含有显著量的蛋白的级分装载在 Sephacryl S-200 柱 (GE Healthcare, NJ) 上用于进一步纯化。将分离级分的样品收集在缓冲液 A 中。

[0401] 在 25℃ 下于 5-mm Shigemi 管 (Shigemi Inc, PA) 内的 pH 7.8 的 40mM Tris • Cl 缓冲液中进行 NMR 实验, 该缓冲液含有在 95 % H₂O/5 % D₂O 中的 64mM NaCl 以及大约 300 μl 的样品体积, 具有的蛋白浓度为 0.5-0.8mM。使用配有低温探头的 Bruker Avance 600MHz 仪器 (Bruker Instruments, MA) 进行 NMR 实验。将水翻转 (water-flipback) 序列用于 ¹H-¹⁵N HSQC 和 {¹H}-¹⁵N NOE 实验以使在酰胺的质子与水的质子之间的交换最小化 (Grzesiek 和 Bax, J. Am. Chem. Soc. 115 :12593, 1994)。对于 8012Hz 谱宽的获取维度 (acquisitiondimension), 记录具有 1024 个复点 (complex point) 的 ¹H-¹⁵N HSQC 谱, 且对于间接 (t₁) 维度, 记录具有 256 个复点的 ¹H-¹⁵N HSQC 谱。{¹H}-¹⁵N NOE 实验是用相似数目的点通过记录两组谱进行, 在 3 秒和 4 秒重复延迟分别有和没有质子饱和 (Gong 和 Ishima, J. Biomol. NMR 37 :147-157, 2007)。NOE 值的不确定性从这两个谱和峰高的均方差噪声 (r. m. s. d. noise) 估计。

[0402] 信号分配 (Signal assignment) 是基于 CH2 结构域的 HNCA、CBCACONH、CCONH 实验, 以及 CH2m01 结构域的 HNCACB 和 CBCACONH 和 ¹³C, ¹⁵N 同时评估 NOESY (Kay 等人, J. Magn. Reson. 89 :496-514, 1990 ;Muhandiram 和 Kay, J. Magn. Res. Series B. 103 :203-216, 1994)。使用 nmrPipe 对 NMR 数据加工和分析 (Delaglio 等人., J. Biomol. NMR 6 :277-293, 1995 ;Masse 和 Keller, J. Magn. Reson. 174 :133-151, 2005)。针对 CH2 主链结构上的化学位移改变的颜色意义, 计算出标准化的化学位移改变 $\delta_{norm} = \sqrt{(\delta_{Ca})^2 + (\gamma_{Ca}/\gamma_N)^2 (\delta_N)^2}$ 、其平均值和标准差 (s. d.), 并将其分成四类: $\delta_{norm} > 3.0$ (红色)、 $3.0 > \delta_{norm} > 2.0$ (橙色)、 $2.0 > \delta_{norm} > 1.0$ (黄色)、以及 (4) $\delta_{norm} < 1$ (蓝色)。

[0403] 结果

[0404] 分离的、未糖基化的人 γ1 CH2 结构域是相对稳定的。将人 γ1 重链 CH2 (图 5A) 克隆在细菌表达载体中, 如以上所述地表达和纯化。人 γ1 CH2 以高水平表达为可溶性蛋白 (超过 10mg 每升的细菌培养物) 并且是高度可溶的 (超过 10mg/ml)。如通过尺寸排阻色谱所测定的, 它在 pH 7.4 的 PBS 中为单体 (图 5B) (Prabakaran 等人., Acta Crystallogr. B. 64 :1062-1067, 2008)。人 γ1 CH2 的 SDS-PAGE 揭示约 14-15kDa 的表观分子量 (MW), 该值接近于计算出的 MW (14.7kDa, 包括 His 和 FLAG 标签)。如所预期的, 它比 scFv、Fab 和 IgG1 的 MW 小得多 (图 5C)。

[0405] 之前已发现分离的未糖基化的小鼠 CH2 结构域在生理相关的温度下是相对不稳定的 (如通过圆二色性 (CD) 所测量的 T_m = 41℃ (Feige 等人., J. Mol. Biol. 344 :107-118, 2004)。人 CH2 的序列与小鼠 CH2 的序列不同, 这可产生不同的稳定性 (图 5A)。为了测试人 γ1 CH2 的热力学稳定性, 使用 CD 和差示扫描量热法 (DSC) 二者。如通过 CD 所测量的, CH2 的二级结构在 25℃ 是由 β 链组成。CH2 解折叠在约 42℃ 开始并且在约 62℃ 完成 (图 6A), 具有 54.1 ± 1.2℃ 的计算的 T_m (图 6A), 解折叠是可逆的 (图 6A)。通过 DSC 获得类似结果 (T_m = 55.4℃, 图 6B)。因此人 γ1 CH2 比其小鼠相应物显著地更稳定。

[0406] 具有额外二硫键的工程化人 γ1 CH2 结构域的设计和产生。为了进一步改善人 CH2 的稳定性, 工程化了在 N 端 A 链与 C 端 G 链之间的额外的二硫键。推论认为限制这两个链的自由度可能导致解折叠程度的降低。最初, 基于在完整的 Fc 中的 CH2 晶体结构设计这些突变体, 该晶体结构非常类似于近来已经被报道的分离的 CH2 的晶体结构, 尽管在一些环中和在末端存在某些差异 (Prabakaran 等人., Acta Crystallogr. B. 64 :1062-1067, 2008)。

基于在具有已知结构的蛋白的两个 C α -碳之间的距离 (Dani 等人., Protein Eng. 16 : 187-193, 2003 ;Pellequer 和 Chen, Proteins 65 :192-202, 2006) 和键的方向, 选择了被 Cys 取代的五对氨基酸 :V10/E103、F11/K104、L12/T105、L12/K104 和 V10/K104 (编号开始于 1 :Ala, 对应于 γ 1 重链中的编号 231) (图 5A ;SEQ ID NO :5)。分别命名为 m01 和 m02 的两个突变体 (L12/K104 至 C12/C104, 在 L12 与 K104 之间的 C α s 距离 = 6.53 Å, 并且 V10/K104 至 C10/C104, 在 V10 与 K104 之间的 C α s 距离 = 7.25 Å) (图 7) 是高度可溶的并且以相当于或高于 CH2 的水平表达 (图 8)。

[0407] 通过质谱证实了额外二硫键的存在。如所预期的, 在 CH2 中二硫键的数目为一, 而在突变体 m01 和 m02 中二硫键数目为二 (表 7)。选择这些突变体用于进一步表征。

[0408] 表 7

[0409] 通过质谱确定的二硫键数目

[0410]

蛋白	完整 (Da)	变性 (D)	还原(R)	还原/烷基化	烷基化	N_{cys}	N_{SH}	-s-s-的数目
		(Da)	(Da)	(R/A) (Da)	(A) (Da)			
CH2	14707.3607	14714.5977	14710.9160	14822.6719	14708.5791	2	0	1
CH2	14674.3447	14677.7539	14676.3398	14899.9238	14669.0400	4	0	2
m01								
CH2	14688.9561	14686.2461	14695.6543	14901.8076	14686.1230	4	0	2
m02								

[0411] $N_{cys} = (M_{R/A}) / 57$

[0412] $N_{SH} = (M_A - M_D) / 57$

[0413] 二硫键 (-s-s-) 的数目 = $(N_{cys} - N_{SH}) / 2$

[0414] M01 和 m02 比 CH2 显著地更稳定。通过 CD 和 DSC 测量 m01、m02 和 CH2 的热力学稳定性, 而它们针对化学剂的稳定性是通过使用尿素和荧光分光光度法测定。在所有的情况下, 这两个突变体比 CH2 稳定得多 (图 9)。CH2、m01 和 m02 的 CD 谱显示它们在 25°C 具有高 β 折叠含量 (图 9A 和 9B)。随着温度升高 β 折叠结构逐渐被破坏 (图 9C)。在 90°C, 该结构处于解折叠状态 (图 9A 和 9B)。也如之前报道, 通过二态模型拟合 S 形曲线 (Feige 等人 1., J. Mol. Biol. 344 :107-118, 2004)。值得注意地, m01 和 m02 的 50% 解折叠发生的温度 (分别是 $T_m = 73.8 \pm 1.7^\circ\text{C}$ 和 $65.3 \pm 0.6^\circ\text{C}$) 显著高于天然 CH2 ($54.1 \pm 1.2^\circ\text{C}$) (图 9C)。CH2 和 m01 可逆地重折叠; 然而 m02 仅部分地重折叠 (图 9A 和图 6A 相比于图 9B)。

[0415] 通过 DSC 获得了类似结果。m01 和 m02 的解链温度比天然 CH2 的解链温度高得多, 也分别升高了约 20°C 和 10°C (图 9D)。令人感兴趣的是, m02 的解折叠比 CH2 和 m01 的解折叠更宽并且具有更低的峰。这种现象可能通过 m02 中二聚体的存在引起。

[0416] m01 和 m02 针对化学引导的解折叠的稳定性也高于 CH2 (图 9E)。使用尿素作为化学剂来测量内荧光光谱。也可通过二态模型拟合解折叠对尿素浓度的依赖性。m01 和 m02 的 50% 解折叠发生的尿素浓度 (分别为 6.8M 和 5.8M) 比 CH2 (4.2M) 高。

[0417] 观察到 m01 的唯一单体级分,而 m02 含有少量的高分子物质,通过 SEC 测定大部分为二聚体(图 10)。因为 m01 的优异特性,选择它用于进一步表征。还测试了平截 CH2 的稳定性和平截 m01 的稳定性,其中前七个 N 端残基(SEQ ID NO:5 的残基 1-7)被删除。这些平截蛋白表现出高稳定性。通过 CD 测量的 50%解折叠温度(分别为 62°C 和 79°C)显著高于(分别 8°C 和 5°C)对应的 CH2 和 m01(分别为 54°C 和 74°C)的 50%解折叠温度(图 10B)。

[0418] M01 的结构保守。为了检验由半胱氨酸突变引起的结构摄动,对 CH2 结构域和 m01 突变体进行了溶液 NMR 实验。¹H-¹⁵N HSQC 谱一般显示氮原子与它们直接结合的质子之间的关联并且提供蛋白质主链的“指纹”。每个 CH2 和 m01 的 ¹H-¹⁵N HSQC 谱(在相同的实验条件中记录)都表现出一组峰,表明蛋白在溶液中完全折叠。在蛋白的结构区域中,主链 ¹⁵N、C' 和 C_α 的化学位移在两种蛋白中指定为约 75%。在 m01 中,残基 Cys 的 C_α 和 C_β 的测量的化学位移分别为 57.6ppm 和 37.7ppm,而 Cys104 的 C_α 和 C_β 的化学位移分别为 34.2ppm 和 54.5ppm。这些值属于氧化的半胱氨酸残基的预期范围(Sharma 和 Rajarathnam, J. Biomol. NMR 18:165-171, 2000),显示在 m01 突变体中形成了另外的二硫键。

[0419] N 和 C_α 的总体主链化学位移的比较也显示在 CH2 与 m01 之间蛋白结构的总体相似性。然而,在残基 Cys31 和 Cys91 附近以及在新引入的 Cys 残基 12 和 104 附近观察到化学位移的改变。这是出乎意料的,因为新引入的二硫桥通过连接相同 β 折叠中的邻近 β 链与 Cys31-Cys91 桥而靠近天然 Cys31-Cys91。在 CH2 m01 中新引入的二硫键最有可能影响在 Cys31 与 Cys91 之间的天然二硫键的微观环境。

[0420] CH2 与 m01 的相对高的环柔韧性和刚性框架。为了确定环在 CH2 和 m01 二者中是否具有柔韧性,记录了 ¹⁵N-¹H NOE。如由高 NOE 值(高于 0.7)所指示,确定框架是刚性的;相反环在平均更具柔韧性。CH2 和 m01 的局部动力学相当,显示处于天然状态的 m01 的构象熵非常类似于 CH2 的构象熵。最可能的是 CH2 结构域的本质结构和动力学被保持,而热稳定性随着半胱氨酸突变的引入而提高。环柔韧性的提高也表明 CH2 和 m01 二者都可以被用作在环中移入残基或突变残基的支架。

[0421] 实施例 4:对 HIV 具有特异性的 CH2 结构域分子

[0422] 这个实施例描述了基于人 IgG1 的 CH2 结构域的环构建合成的噬菌体文库以鉴定特异地结合 HIV 包膜的 CH2 分子。

[0423] 材料和方法

[0424] 引物、肽和蛋白。在这个研究中使用的所有引物都购自 Invitrogen(Carlsbad, CA)。生物素标记的肽购自 Sigma(St. Louis, MO)。Balgpl20-CD4 由 Tim Fouts(马里兰大学, Baltimore, MD) 友情提供并且其他的 gp 120/140 由 Christopher Broder(USUHS, Bethesda, MD) 提供。SCD4 是通过 AIDS 研究与试剂项目获得的。

[0425] 文库构建。使用重叠 PCR 来向环 1 和环 3 引入突变以产生第一基于 CH2 的文库。使用 N 端引物 ACGT GGCC CAGGC GGCC GCA CCTGAA CTC CTG(SEQ ID NO:101) 和环 1 引物 CAC GTA CCA GTTGAA CTT GCC AKM AKM AKM AKM AKM AKM AKM AKM AKMAKM CAC CAC CAC GCA TGT GAC(SEQ ID NO:7) 来产生环 1 中含有突变的半个 CH2 的 N 端。使用环 1 连接引物 AAG TTC AAC TGGTAC GTG(SEQ ID NO:8) 和环 3 引物 GAT GGT TTT CTC GAT GGGGCC AKM AKM AKM AKM AKM AKM GTT GGA GAC CTT GCACTT G(SEQ ID NO:9) 来产生环 3 中具有

突变的其余的 CH2。然后通过重叠 PCR 步骤将这两个片段组合并且用 N 端引物和 C 端引物 ACGTGGCC GGCCT GGCC TTT GGC TTT GGA GAT GGT TTT CTC GAT G (SEQ ID NO:102) 扩增, 将 SfiI 位点 (下划线标出的) 引入到 CH2 片段的两个末端。为了产生基于从第一个文库分离的结合物的二级文库, 使用环 2 引物 GCT GAC CAC ACG GTA ADH ADH ADH GTA CTGCTC CTC CCG (SEQ ID NO:103) 和上述 N 端引物来将环 2 的突变引入到一级结合物。使用环 2 连接引物 TAC CGT GTG GTC AGC (SEQ ID NO:104) 和环 3 引物 (2) GGA GAT GGT TTT CTC GAT GGG ADH TGGADH ADH ADH GTT GGA GAC CTT GCA (SEQ ID NO:105) 来将突变引入到一级结合物。通过重叠 PCR 步骤将这两个片段接合, 并使用上述用于扩增的相同一对 N 端和 C 端引物扩增。使 PCR 片段经历 SfiI 消化并将其连接至载体。将连接的产物脱盐并转化到适合于电穿孔仪 (Bio-Rad, Hercules, CA) 的电感受态 TG1 细胞中。由得到的转化体制备噬菌体文库。

[0426] 淘选。在 4℃ 将 Ba1 gp120-CD4、Ba1 gp120 以及 BSA 直接在 PBS 缓冲液中包被到 Maxisorp 板 (Nunc, Denmark), 过夜用于板形式的淘选。将对应 CH2 文库的大约 10^{13} 个噬菌体粒子悬浮在具有 2% 奶粉的 PBS 中, 并且施用至以蛋白包被的孔中。在室温下 2 小时后, 第一轮每个孔洗涤 5 次并且随后一轮洗涤 10 次, 然后噬菌体用处于指数生长期的 TG1 细胞复苏。对第一个文库的每种抗原进行总计五轮的淘选。对基于一级结合物的第二个文库, 进行三轮淘选。然后使用单克隆 ELISA 来选择阳性克隆。对每种抗原筛选两百个克隆。只选择在单克隆 ELISA 中展示 OD₄₀₅ > 2.0 的克隆用于噬菌粒的制备和测序。

[0427] CH2 的表达和重折叠。将如以上所述选择的克隆转化到大肠杆菌菌株 HB2151 中用于表达。简言之, 将单个克隆接种在补充了 100 单位 amp 的 2xYT 中并且在摇动下在 37℃ 培养。当 OD₆₀₀ 达到 0.5 时, 加入 IPTG 以达到 1mM 的终浓度, 并将培养物继续摇动另外 3-5 小时。然后收集细胞, 用 PBS 中多粘菌素 B (Sigma, St Louis) 裂解, 并且使上清液经历用于 CH2 克隆的可溶级分的 Ni-NTA 琼脂糖小珠 (Qiagen, Hilden, Germany) 纯化。然后将沉淀重悬在含有 25mM Tris, HCl, pH8.0、6M 尿素、0.5M NaCl 的缓冲液中并使其经历短暂超声处理。通过离心收集上清液并使其经历 Ni-NTA 琼脂糖小珠 (Qiagen) 纯化。使通过沉淀获得的 CH2 经历针对两次 PBS 改变的过夜透析然后过滤通过 0.2 μm 的低蛋白结合滤器 (Pal, Ann Arbor, MI)。

[0428] ELISA。将不同的蛋白抗原以 1-4 μg/ml 的浓度范围稀释在 PBS 缓冲液中, 并将其在 4℃ 包被到 96 孔板过夜。然后用 PBS+奶粉的缓冲液将板封闭。将不同浓度的 CH2 克隆稀释在相同的封闭缓冲液中并且施加到 ELISA 板。除非另外指明, 否则在大多数 ELISA 中使用小鼠抗 His-HRP 来检测在每个 CH2 克隆 C 端的 His 标签。然后将 ABTS 加至每个孔并且在之后的 5-10 分钟取得 OD₄₀₅。

[0429] 凝胶过滤分析。将纯化的和过滤的 CH2 蛋白的样品在用 PBS 预平衡的 Superdex7510/300GL 柱 (GE Healthcare, Piscataway, NJ) 上分析。用分子量标准物校准该柱。将 CH2 样品以 0.5ml/min 的流速从该柱上洗脱。

[0430] 假病毒中和测定。基本上如之前所述地进行 HIV Env 假型病毒的制备和中和 (Choudhry 等人., Virology 363:79-90, 2007)。

[0431] 结果

[0432] 基于人 CH2 文库的设计和构建。假设 CH2 环的有限诱变可能没有显著影响许多突

变体的折叠和稳定性,并且可以被用于产生潜在结合物的大型文库。首先,进行环 1(L1) 和环 3(L3) 的诱变,因为它们是在分子相同侧上的最长(分别为 9 个和 5 个残基)的两个环(环 BC、DE 和 FG 在本文分别称为 L1、L2 和 L3;两个螺旋 AB 和 EF 分别称为 H1 和 H2;并且环 CD 称为 L0)(图 11)(Radaev 等人., J. Biol. Chem. 276:16469-16477, 2001)。选择 CDR 中四个经常出现的残基(A、Y、D 和 S)来随即置换所有的 L1 和 L3 残基并且添加一个额外的残基。还将一个额外的残基(G)添加到每个环的 C 端末端以增加柔韧性(图 11)。之前已观察到这四个残基(有时只有两个)足以在不同的框架内构建特异性结合表面(Fellouse 等人., Proc. Natl. Acad. Sci. USA101:12467-12472, 2004;Koide 等人., Proc. Natl. Acad. Sci. USA104:6632-6637, 2007)。计算出的这个文库的理论多样性为 $4^{16} = 4.3 \times 10^9$ 。然而,由于通过 PCR 产生的可能的突变(见以下),多样性可能显著高达文库的大小(5×10^{10})。如通过分析 100 个随机选择的克隆所表明的,大多数突变体(可能多于 80%)具有正确的阅读框。

[0433] 结合物的鉴定和序列分析。为了测试文库并选择可能有用的结合物,将与两个结构域 CD4 融合的来自 Bal 分离物(isolate)的 HIV-1 包膜糖蛋白 gp120(表示为 gp120_{Bal}-CD4)被用作抗原。在五轮淘选之后,通过噬菌体 ELISA 筛选 200 个克隆,并且分离 15 个具有最高信号的克隆用于进一步表征。三个克隆 m1a1、m1a2 和 m1a3 分别由(15 个序列之中的)7、5 和 2 序列支配代表,表现出特异性富集。它们具有主要由 D 和 Y 组成的类似 L1 序列但是它们的 L3 非常不同。最丰富的克隆 m1a1 和 m1a2 具有 L1 中若干改变(分别为 L1 中的两个 F 和在 G 之前的缺失),所述改变明显是由于 PCR 错误。针对 Bal gp120-CD4 所选择的克隆的环 1 和环 3 序列显示在以下表 8 中。这些结果表明基于 CH2 的支架能够支撑具有不同的 L1 和 L3 的噬菌体展示的结合物;新鉴定的 HIV-1 特异性结合物如以下所述被进一步表征。

[0434] 表 8

[0435] CH2 的环 1 和环 3 序列

[0436]

克隆	环 1 序列	SEQ ID NO:	环 3 序列	SEQ ID NO:
m1a1	DYDYDSYDFG	107	SDSAASG	96 (aa 4-10)
m1a2	DYDYDSYYD..G	108	DDYAADG	96 (aa 4-10)
m1a3	DYDYDSYYDYG	94 (aa 4-14)	YDYADDG	96 (aa 4-10)
m1a3 [*]	DYDYDSYYDYG	94 (aa 4-14)	SDYDSSG	96 (aa 4-10)
野生型 CH2	DVSHEDPEV	93 (aa 4-12)	KALPA	95 (aa 4-8)

[0437] *m1a3' 克隆具有与 m1a3 相同的环 1 序列但具有不同的环 3 序列

[0438] 可溶性 nAb 的表达及其结合的表征。大多数表达的 CH2 结构域分子（称为“nAb”）在包涵体中被发现（图 12A）并且如以上所述地重折叠，每升细菌培养物平均产生 10-30mg。如通过 ELISA 所测量，纯化的 nAb 与淘选抗原（gp120-CD4）特异性结合，具有范围从 500nM(m1a1 和 m1a2) 到低 μ M(m1a3) 的 EC50（图 12B）。对于从上清液中纯化的 nAb，获得了相似的结果。这些结果显示 m1a1、m1a2 和 m1a3 保留了其以可溶形式（而非噬菌体展示形式）的结合活性，并且从包涵体中重折叠并未影响这些分子。选择具有最高亲和力的两个克隆 m1a1 和 m1a2 用于进一步表征。

[0439] 为了测试它们的交叉活性，单独使用或在与可溶性 CD4 的复合物中使用四种（Ba1、JRFL、R2 和 89.6）重组 HIV-1 包膜糖蛋白。如在图 14 中所示，m1a1 以不同的程度与所有蛋白结合。尽管 m1a1 在与 CD4 的复合物中与 Ba1 gp120 结合，但如对 CD4 诱导的（CD4i）抗体所预期，其单独与 gp120 结合非常弱，其与其他蛋白的结合未显著地受 CD4 存在的影响。单独的 Env 信号降低是不显著的，并且可能是因为当与 sCD4 混合时被 gp120 包覆稍微减少。对 m1a2 获得类似结果。这些数据显示由这些抗体识别的表位对一个分离物（Ba1）是 CD4i，但对其他分离物则不是。

[0440] 为了进一步表征它们的表位，测试了 m1a1 与已了解的 CD4i 抗体（scFv X5 和结构域抗体 m36）的竞争。两种 CD4i 抗体均明显地与 m1a2 竞争。因此，m1a1 识别由其他高效交叉反应的 CD4i 抗体所共享的新颖保守表位，但是与这些抗体相反，其通过 gp120 与 CD4 相互作用的暴露明显依赖于分离物。

[0441] 环 1 决定结合特异性。为了确定来自 CH2 克隆的环对特异性结合的不同贡献，产生了两个杂合克隆：m1a1CH2 和 m1a2CH2。将来自于 m1a1 和 m1a2 的 L1 植入到 CH2 上，置换了原始 L1。如通过 ELISA 所测量的，这些杂合抗体相比于 m1a1，以大约相同但略微降低的亲和力与 gp120-CD4 结合（图 13A），表明 L3 不是结合的关键。为了找出该支架是否整体上是结合所需的，分离测试了作为合成肽（DYDYDSYDFDG；SEQ ID NO:109）的 m1a1 L1。生物素标记的肽不结合。还通过在链 A 和链 G 之间产生额外的二硫键来测试支架中的相对微小的构象变化对结合的影响。如在实施例 3 中所述，这样的 S-S 键显著增加了 CH2 的稳定性，并且如 NMR 所测量，不显著影响任何 CH2 残基的移动性和微环境。得到的抗体 m1a1ss 不结合任何一个（图 13B）。这些数据表明 m1a1 的结合活性需要支架，并且尽管 L3 中的改变可能不影响其活性，但是在支架构象中相对小的改变可以除去活性。

[0442] m1a1 和 m1a2 对 HIV-1 假病毒的中和。为了评价 m1a1 和 m1a2 的中和活性，使用了细胞系 / 假病毒测定和一组九个 HIV-1 分离物。这些分离物中的七个在某种程度上被一种或两种抗体抑制（图 14A）。这两种抗体有差别地抑制两个分离物（89.6 和 IIIB），并且其程度与五个其他分离物大致相同（图 14A）。如从它们相对适中的结合亲和力所预期，它们的效力与这里用作阳性对照的高效抑制剂 C34 相比是相对适中的。这些结果提供了如下概念的证据：可从基于 CH2 支架的文库中选出功能性结合物。

[0443] 通过第二个环和第三个环的诱变进一步改进抗体（图 15）。它们在 SDS 凝胶上大多数以单体运行（图 16A）。突变体之一 m1b3 在凝胶过滤中主要为单体（图 16B）。它们特异地结合（图 16C）和中和不同程度的 HIV-1（图 17）。它们还与 scFv X5 和 m36 竞争，表明它们靶向 HIV-1 gp120 的高度保守区域（图 18）。

[0444] 本公开内容提供了包含至少一个突变或至少一个 CDR 或其功能片段的抗体恒定结构域分子。本公开内容还提供了包含抗体恒定结构域分子的组合物及其用途。明显的是所述方法的精确细节可以被改变或改良,而不背离本发明的精神。我们要求保护属于以下权利要求的范围和精神的所有这些改良方案 and 变化方案。

[0001]

序列表

<110> 美国政府健康及人类服务部

<120> 工程化抗体恒定结构域分子

<130> SCT103419-20

<150> US 61/063,245

<151> 2008-01-31

<160> 108

<170> PatentIn version 3.5

<210> 1

<211> 129

<212> PRT

<213> Homo sapiens

<220>

<221> MISC_FEATURE

<222> (50)..(50)

<223> X = any amino acid

<220>

<221> MISC_FEATURE

<222> (53)..(53)

<223> X = any amino acid

<220>

<221> MISC_FEATURE

<222> (101)..(115)

<223> X = any amino acid

<400> 1

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
20 25 30

[0002]

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Xaa Ile Ser Xaa Lys Thr Asp Gly Gly Thr Tyr Tyr Tyr Ala Asp
50 55 60

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

Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr
85 90 95

Tyr Cys Ala Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Xaa Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125

Ser

<210> 2

<211> 118

<212> PRT

<213> Homo sapiens

<400> 2

His Val Gln Leu Gly Glu Ser Gly Gly Gly Val Val Arg Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
20 25 30

[0003]

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Arg Asn Ala Lys Asn Ser Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Arg Gly Gly Val Trp Ala Phe Asp Ile Trp Gly Gln Gly Thr
100 105 110

Met Val Thr Val Ser Ser
115

<210> 3
<211> 119
<212> PRT
<213> Homo sapiens

<400> 3

Gln Val Gln Lys Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Ile Phe Ser Ser Tyr
20 25 30

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

[0004]

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Asp Leu Gly Arg Gln Gln Leu Ser Asp Tyr Trp Gly Gln Gly
100 105 110

Thr Leu Val Thr Val Ser Ser
115

<210> 4

<211> 115

<212> PRT

<213> Homo sapiens

<400> 4

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
20 25 30

Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
35 40 45

Gly Ile Ile Asn Pro Ser Gly Gly Ser Thr Ser Tyr Ala Gln Lys Phe
50 55 60

Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Thr Ser Thr Val Tyr

[0005]

65	70	75	80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys			
	85	90	95

Ala Thr Pro Tyr Gly Met Asp Val Trp Gly Gln Gly Thr Thr Val Thr			
100	105	110	

Val Ser Ser
115

<210> 5
<211> 110
<212> PRT
<213> Homo sapiens

<400> 5

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys			
1	5	10	15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val			
20	25	30	

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr			
35	40	45	

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu			
50	55	60	

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His			
65	70	75	80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys			
85	90	95	

[0006]

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys
100 105 110

<210> 6

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 6

gcactggctg gtttcgctac cgtggcccag gcggccgcac ctgaactcct g 51

<210> 7

<211> 69

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 7

cacgtaccag ttgaacttgc cakmakmakm akmakmakma kmakmakmak mcaccaccac 60

gcatgtgac 69

<210> 8

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 8

aagttcaact ggtacgtg 18

[0007]

<210> 9
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic oligonucleotide

<400> 9
gatggttttc tcgatggggc cakmakmakm akmakmakmg ttggagacct tgcacttg 58

<210> 10
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic oligonucleotide

<400> 10
ggtgcagaag atggtggtgg cgggcctggc ctttggcttt ggagatggtt ttctcgatg 59

<210> 11
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 11

Pro Glu Val Thr Cys Val Val Val Tyr Tyr Asp Ser Ala Ala Ala Tyr
1 5 10 15

Ala Tyr Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0008]

<210> 12

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 12

Pro Glu Val Thr Cys Val Val Val Tyr Tyr Ser Ala Ser Ala Ala Ala
1 5 10 15

Ser Ala Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 13

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 13

Pro Glu Val Thr Cys Val Val Val Tyr Asp Ser Asp Tyr Ala Ser Ser
1 5 10 15

Asp Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Lys Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0009]

<210> 14
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 14

Pro Glu Val Thr Cys Val Val Val Ala Tyr Ser Asp Asp Ala Ala Ala
1 5 10 15

Tyr Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 15
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 15

Pro Glu Val Thr Cys Val Val Val Asp Ala Asp Asp Asp Tyr Tyr Tyr
1 5 10 15

Tyr Tyr Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0010]

<210> 16
<211> 48
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 16

Pro Glu Val Thr Cys Val Val Val Asp Asp Ala Tyr Tyr Asp Ala Asp
1 5 10 15

Tyr Tyr Tyr Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His
20 25 30

Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 17
<211> 45
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 17

Pro Glu Val Thr Cys Val Val Val Asp Ala Ala Tyr Asp Tyr Ser Tyr
1 5 10 15

Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
20 25 30

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0011]

<210> 18

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 18

Pro Glu Val Thr Cys Val Val Val Asp Tyr Asp Ser Asp Asp Ala Tyr
1 5 10 15

Ala Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 19

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 19

Pro Glu Val Thr Cys Val Val Val Ser Tyr Tyr Asp Ser Asp Ser Tyr
1 5 10 15

Ser Ala Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0012]

<210> 20
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 20

Pro Glu Val Thr Cys Val Val Val Asp Asp Ala Tyr Ala Asp Asp Ala
1 5 10 15

Ser Ala Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 21
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 21

Pro Glu Val Thr Cys Val Val Val Ser Tyr Tyr Ser Asp Ser Asp Tyr
1 5 10 15

Asp Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0013]

<210> 22
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 22

Pro Glu Val Thr Cys Val Val Val Asp Asp Asp Ser Tyr Tyr Ser Tyr
1 5 10 15

Asp Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 23
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 23

Pro Glu Val Thr Cys Val Val Val Tyr Asp Ala Ser Asp Tyr Ala Asp
1 5 10 15

Ala Tyr Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0014]

<210> 24
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 24

Pro Glu Val Thr Cys Val Val Val Ala Asp Ala Ala Ala Tyr Ala Tyr
1 5 10 15

Ala Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 25
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 25

Pro Glu Val Thr Cys Val Val Val Ala Ser Asp Ser Ser Asp Asp Tyr
1 5 10 15

Asp Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
20 25 30

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0015]

<210> 26

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 26

Pro Glu Val Thr Cys Val Val Val Ala Ala Ala Ala Ala Asp Ala Asp
1 5 10 15

Tyr Tyr Ser Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 27

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 27

Pro Glu Val Thr Cys Val Val Val Tyr Asp Asp Ala Ala Tyr Ala Asp
1 5 10 15

Asp Tyr Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0016]

<210> 28
<211> 45
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 28

Pro Glu Val Thr Cys Val Val Val Ser Ala Asp Ala Ser Asp Tyr Asp
1 5 10 15

Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Asp
20 25 30

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 29
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 29

Pro Glu Val Thr Cys Val Val Val Asp Asp Asp Ala Ala Asp Ala Tyr
1 5 10 15

Tyr Tyr Gly Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg
35 40 45

[0017]

<210> 30

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 30

Pro Glu Val Thr Cys Val Val Val Tyr Asp Ser Asp Asp Asp Tyr Asp
1 5 10 15

Tyr Ala Gly Lys Phe Cys Trp Tyr Val Asp Gly Val Glu Val His Asn
20 25 30

Ala Lys Thr Lys Pro Arg Glu Glu His Tyr Asn Ser Thr Tyr Arg
35 40 45

<210> 31

<211> 41

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 31

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Ala Ser Ala Tyr Ser Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0018]

<210> 32
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 32

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Asp Asp Ala Asp Ala Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 33
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 33

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Ala Asp Ala Tyr Ala Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0019]

<210> 34
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 34

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Ala Asp Tyr Ser Asp Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 35
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 35

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Ala Asp Ala Ala Asp Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0020]

<210> 36
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 36

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Asp Ala Ser Ala Ser Ser Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 37
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 37

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Asp Asp Tyr Ala Ala Ser Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0021]

<210> 38
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 38

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Asp Ala Tyr Ala Ser Asp Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 39
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 39

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Asp Ala Asp Asp Ala Ser Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0022]

<210> 40
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 40

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Ala Asp Asp Asp Ser Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 41
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 41

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Asp Ala Tyr Ala Tyr Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0023]

<210> 42

<211> 41

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 42

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Asp Asp Tyr Asp Tyr Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 43

<211> 41

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 43

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Tyr Ser Asp Ser Ala Ala Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0024]

<210> 44
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 44

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Tyr Ala Ala Ser Ala Tyr Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 45
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 45

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Tyr Asp Asp Asp Ala Asp Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0025]

<210> 46
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 46

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Tyr Tyr Asp Tyr Asp Tyr Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 47
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 47

Val Val Ser Val Leu Thr Val Leu His His Asp Trp Met Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Glu Val Ser Asn Asp Ala Asp Ser Ala Asp Gly Pro
20 25 30

Ile Lys Lys Thr Ile Ser Lys Ala Lys
35 40

[0026]

<210> 48
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 48

Val Val Ser Val Leu Thr Val Leu His His Asp Trp Leu Asn Gly Glu
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Asp Ala Ser Asp Asp Ala Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Ala Lys
35 40

<210> 49
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 49

Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys
1 5 10 15

Glu Tyr Lys Cys Lys Val Ser Asn Ala Asp Asp Ala Tyr Ala Gly Pro
20 25 30

Ile Glu Lys Thr Ile Ser Lys Ala Lys
35 40

[0027]

<210> 50

<211> 41

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 50

Val Val Ser Val Leu Thr Val Leu His His Tyr Trp Met Asn Gly Glu
1 5 10 15

Asp Tyr Lys Cys Glu Val Ser Asn Asp Ser Tyr Ser Asp Asp Gly Pro
20 25 30

Ile Lys Lys Thr Ile Ser Lys Ala Lys
35 40

<210> 51

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 51

tagcgattcg ctaccgtggc ccaggcggcc cctgaactcc tggggggacc

50

<210> 52

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

[0028]

<400> 52	
tccccccagg agttcagggtg c	21
<210> 53	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 53	
tgcgtggtgg tggacgtgag c	21
<210> 54	
<211> 62	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 54	
taggcatgca tctgcatggt ggccggcctg gcctttggct ttggagatgg ttttctcgat	60
gg	62
<210> 55	
<211> 36	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 55	
gaactcctgg ggggaccggc yayrtattac tgtgyg	36
<210> 56	
<211> 36	

[0029]

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 56

gaactcctgg ggggaccggc yttrtattac tgtgyg

36

<210> 57

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 57

gaactcctgg ggggaccggc ygtrtattac tgtgyg

36

<210> 58

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 58

gctcacgtcc accaccacgc aggtgccctg gcccga

36

<210> 59

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 59

gctcacgtcc accaccacgc aggtgccacg gcccga

36

[0030]

<210>	60	
<211>	36	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic oligonucleotide	
<400>	60	
	gctcacgtcc accaccacgc aggtgccayg gcccga	36
<210>	61	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic oligonucleotide	
<400>	61	
	gtacgtgctg ttgtactgct c	21
<210>	62	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic oligonucleotide	
<400>	62	
	aaggtctcca acaaagccct c	21
<210>	63	
<211>	42	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
[0031]		

<223> Synthetic oligonucleotide

<400> 63

gagcagtaca acagcacgta cgcagcyayr tattactgtg yg

42

<210> 64

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 64

gagcagtaca acagcacgta cgcagcyttr tattactgtg yg

42

<210> 65

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 65

gagcagtaca acagcacgta cgcagcygtr tattactgtg yg

42

<210> 66

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 66

gagggctttg ttggagacct tggttccctg gcccga

36

<210> 67

<211> 36

[0032]

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 67

gagggctttg ttggagacct tgggtgccacg gcccca

36

<210> 68

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 68

gagggctttg ttggagacct tgggtgccayg gcccca

36

<210> 69

<211> 40

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 69

Ala Val Tyr Tyr Cys Val Lys Val Pro Val Gly Tyr Trp Gly Arg Gly
1 5 10 15

Thr Ala Val Tyr Tyr Cys Ala Asp Val Glu Ala Ser Ser Pro Ala Asp
20 25 30

Phe Gly Tyr Trp Gly Arg Gly Thr
35 40

[0033]

<210> 70
<211> 50
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 70

Ala Met Tyr Tyr Cys Ala Arg Asp His Gly Val Asp Thr Ala Met Ala
1 5 10 15

Gly Pro Trp Phe Asp Tyr Trp Gly Arg Gly Thr Ala Val Tyr Tyr Cys
20 25 30

Val Arg Gly Thr Gly Trp Glu Leu Leu Val Ile Asp Cys Trp Gly Arg
35 40 45

Gly Thr
50

<210> 71
<211> 38
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 71

Ala Val Tyr Tyr Cys Ala Arg Gly Ser Ser Gly Trp Gly Trp Phe Asp
1 5 10 15

Pro Trp Gly Gln Gly Thr Ala Thr Tyr Tyr Cys Ala Arg Asp Arg Gly
20 25 30

[0034]

Tyr Trp Gly Arg Gly Thr
35

<210> 72
<211> 48
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 72

Ala Val Tyr Tyr Cys Ala Arg Arg Met Pro Glu Gly Asp Ser Ser Gly
1 5 10 15

Thr Ser Tyr Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Ala Leu Tyr Tyr
20 25 30

Cys Ala Arg Glu Glu Lys Gly Asp Tyr Asp Tyr Trp Gly Gln Gly Thr
35 40 45

<210> 73
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 73

Ala Met Tyr Tyr Cys Ala Ile His Ser Phe Asp Tyr Trp Gly Gln Gly
1 5 10 15

Thr Ala Val Tyr Tyr Cys Ala Lys Val Leu Ser Gly Trp Phe Asp His
20 25 30

[0035]

Tyr Phe Asp Ser Trp Gly Gln Gly Thr
35 40

<210> 74
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 74

Ala Val Tyr Tyr Cys Ala Arg Asp Arg Val Pro Asp Gly Val Trp Ser
1 5 10 15

Ala Asp Ser Trp Gly Gln Gly Thr Ala Val Tyr Tyr Cys Ala Ser Lys
20 25 30

Pro Pro Val Ser Asn Trp Phe Asp Pro Trp Gly Gln Gly Thr
35 40 45

<210> 75
<211> 43
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 75

Ala Ile Tyr Tyr Cys Val Lys Ala Gly Tyr Asn Phe Asp Ala Phe Asp
1 5 10 15

His Trp Gly Arg Gly Thr Ala Met Tyr Tyr Cys Ala Gly Asp Thr Ala
20 25 30

[0036]

Met Val Ile Phe Asp Tyr Trp Gly Gln Gly Thr
35 40

<210> 76
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 76

Ala Thr Tyr Tyr Cys Ala Ser Gly Ser Ser Gly Cys Ser Asp Tyr Trp
1 5 10 15

Gly Gln Gly Thr Ala Thr Tyr Tyr Cys Ala Arg Gly Gly Tyr Ser Ser
20 25 30

Gly Trp Tyr His Trp Tyr Phe Asp Leu Trp Gly Arg Gly Thr
35 40 45

<210> 77
<211> 42
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 77

Ala Val Tyr Tyr Cys Ala Ala Ser Val Gly Ala Pro Ser Asp Phe Asp
1 5 10 15

Tyr Trp Gly Gln Gly Thr Ala Thr Tyr Tyr Cys Ala Thr Thr Pro Asp
20 25 30

[0037]

Ser Asn Tyr Gly Tyr Trp Gly Gln Gly Thr
35 40

<210> 78

<211> 40

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 78

Ala Leu Tyr Tyr Cys Ala Lys Gly Gln Tyr Gly Asp His Asp Tyr Trp
1 5 10 15

Gly Gln Gly Thr Ala Val Tyr Tyr Cys Ala Lys Glu Glu Glu Gly Ala
20 25 30

Val Leu Gly Trp Gly Arg Gly Thr
35 40

<210> 79

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 79

Ala Thr Tyr Tyr Cys Ala Arg Glu Gly Thr Val Val Thr Pro Tyr Phe
1 5 10 15

Val Tyr Trp Gly Gln Gly Thr Ala Val Tyr Tyr Cys Ala Met Gly Gly
20 25 30

[0038]

His Gly Ser Gly Ser Tyr Leu Ser Gly Tyr Trp Gly Gln Gly Thr
35 40 45

<210> 80

<211> 41

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 80

Ala Val Tyr Tyr Cys Ala Arg Glu Arg Tyr Gly Ala Leu Asp Tyr Trp
1 5 10 15

Gly Arg Gly Thr Ala Val Tyr Tyr Cys Ala Gly Gly Leu Leu His Glu
20 25 30

Gly Ser Gly Tyr Trp Gly Gln Gly Thr
35 40

<210> 81

<211> 41

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 81

Ala Ile Tyr Tyr Cys Ala Ala Arg Gly Gln Gly Asn Ser Trp Trp Phe
1 5 10 15

Asp Pro Trp Gly Gln Gly Thr Ala Ile Tyr Tyr Cys Ala Thr Gln Val
20 25 30

[0039]

Gly His Gly Asp Trp Gly Gln Gly Thr
35 40

<210> 82
<211> 47
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 82

Ala Leu Tyr Tyr Cys Ala Arg Ala Tyr Ser Ala Tyr Gln Tyr Ser Phe
1 5 10 15

Asp Ser Trp Gly Arg Gly Thr Ala Val Tyr Tyr Cys Ala Arg Arg Glu
20 25 30

Tyr Asn Trp Asn His Asn Trp Phe Asp Pro Trp Gly Gln Gly Thr
35 40 45

<210> 83
<211> 44
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 83

Ala Thr Tyr Tyr Cys Ala Arg Arg Gly Asp Asp Tyr Gly Asp Tyr Phe
1 5 10 15

Phe Asp Tyr Trp Gly Gln Gly Thr Ala Ile Tyr Tyr Cys Ala Arg Ser
20 25 30

[0040]

Arg Gly Ser Ser Phe Asp Tyr Trp Gly Gln Gly Thr
35 40

<210> 84

<211> 45

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 84

Ala Met Tyr Tyr Cys Ala Arg Asp Leu Tyr Ser Asn Tyr Val Asp Tyr
1 5 10 15

Trp Gly Gln Gly Thr Ala Val Tyr Tyr Cys Ala Arg Gly Pro Trp Gln
20 25 30

Gln Leu Val Asn Trp Phe Asp Pro Trp Gly Gln Gly Thr
35 40 45

<210> 85

<211> 40

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 85

Ala Thr Tyr Tyr Cys Ala Ser Leu Thr Gly Thr Thr Ser Tyr Trp Gly
1 5 10 15

Gln Gly Thr Ala Leu Tyr Tyr Cys Ala Arg Ala Thr Trp Gly Tyr Gln
20 25 30

[0041]

Phe Asp Cys Trp Gly Gln Gly Thr
35 40

<210> 86

<211> 42

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 86

Ala Ile Tyr Tyr Cys Ala Arg Glu Ser Ser Ser Ser Phe Asp Tyr Trp
1 5 10 15

Gly Gln Gly Thr Ala Val Tyr Tyr Cys Ala Arg Met Ser Gly Gly Arg
20 25 30

Trp Ile Phe Asp His Trp Gly Gln Gly Thr
35 40

<210> 87

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 87

Ala Val Tyr Tyr Cys Ala Arg Gly Trp Glu Leu Asp Tyr Trp Gly Gln
1 5 10 15

Gly Thr Ala Val Tyr Tyr Cys Ala Lys Thr Gly Gln Phe Asp Tyr Trp
20 25 30

[0042]

Gly Gln Gly Thr

35

<210> 88

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 88

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

1

5

10

15

<210> 89

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 89

Glu Lys Thr Ile Ser Lys Ala Lys

1

5

<210> 90

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 90

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

1

5

10

15

[0043]

<210> 91
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 91

Glu Lys Thr Ile Ser Lys Ala Lys
1 5

<210> 92
<211> 107
<212> PRT
<213> Mus musculus

<400> 92

Val Pro Glu Val Ser Ser Val Phe Ile Phe Pro Pro Lys Pro Lys Asp
1 5 10 15

Val Leu Thr Ile Thr Leu Thr Pro Lys Val Thr Cys Val Val Val Asp
20 25 30

Ile Ser Lys Asp Asp Pro Glu Val Gln Phe Ser Trp Phe Val Asp Asp
35 40 45

Val Glu Val His Thr Ala Gln Thr Gln Pro Arg Glu Glu Gln Phe Asn
50 55 60

Ser Thr Phe Arg Ser Val Ser Glu Leu Pro Ile Met His Gln Asp Trp
65 70 75 80

Leu Asn Gly Lys Glu Phe Lys Cys Arg Val Asn Ser Ala Ala Phe Pro

[0044]

85

90

95

Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys
100 105

<210> 93
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 93

Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn
1 5 10 15

<210> 94
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<220>
<221> MISC_FEATURE
<222> (4)..(13)
<223> X = Tyr, Ala, Asp or Ser

<400> 94

Val Val Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Lys Phe
1 5 10 15

Asn

[0045]

<210> 95
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 95

Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
1 5 10

<210> 96
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<220>
<221> MISC_FEATURE
<222> (4)..(9)
<223> X = Tyr, Ala, Asp or Ser

<400> 96

Val Ser Asn Xaa Xaa Xaa Xaa Xaa Xaa Gly Pro Ile Glu
1 5 10

<210> 97
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 97

[0046]

Gln Tyr Asn Ser Thr Tyr Arg

1

5

<210> 98

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<220>

<221> MISC_FEATURE

<222> (3)..(5)

<223> X = Asn, Thr, Ile, Asp, Ala, Val, Ser, Tyr or Phe

<400> 98

Gln Tyr Xaa Xaa Xaa Tyr Arg

1

5

<210> 99

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 99

Ser Asn Lys Ala Leu Pro Ala Pro Ile

1

5

<210> 100

<211> 9

<212> PRT

<213> Artificial Sequence

[0047]

<220>

<223> Synthetic polypeptide

<220>

<221> MISC_FEATURE

<222> (3)..(5)

<223> X = Asn, Thr, Ile, Asp, Ala, Val, Ser, Tyr or Phe

<220>

<221> MISC_FEATURE

<222> (7)..(7)

<223> X = Asn, Thr, Ile, Asp, Ala, Val, Ser, Tyr or Phe

<400> 100

Ser Asn Xaa Xaa Xaa Pro Xaa Pro Ile

1

5

<210> 101

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 101

acgtggccca ggccggccgca cctgaactcc tg

32

<210> 102

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic oligonucleotide

<400> 102

acgtggccgg cctggccttt ggctttggag atggttttct cgatg

45

[0048]

<210> 103
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic oligonucleotide

<400> 103
gctgaccaca cggtaadhad hadhgtactg ctcctcccg 39

<210> 104
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic oligonucleotide

<400> 104
taccgtgtgg tcagc 15

<210> 105
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic oligonucleotide

<400> 105
ggagatgggtt ttctcgatgg gadhtggadh adhadhgttg gagaccttgc a 51

<210> 106
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

[0049]

<400> 106

Asp Tyr Asp Tyr Asp Ser Tyr Phe Asp Phe Gly
1 5 10

<210> 107

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 107

Asp Tyr Asp Tyr Asp Ser Tyr Phe Asp Phe Gly
1 5 10

<210> 108

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic polypeptide

<400> 108

Asp Tyr Asp Tyr Asp Ser Tyr Tyr Asp Gly
1 5 10

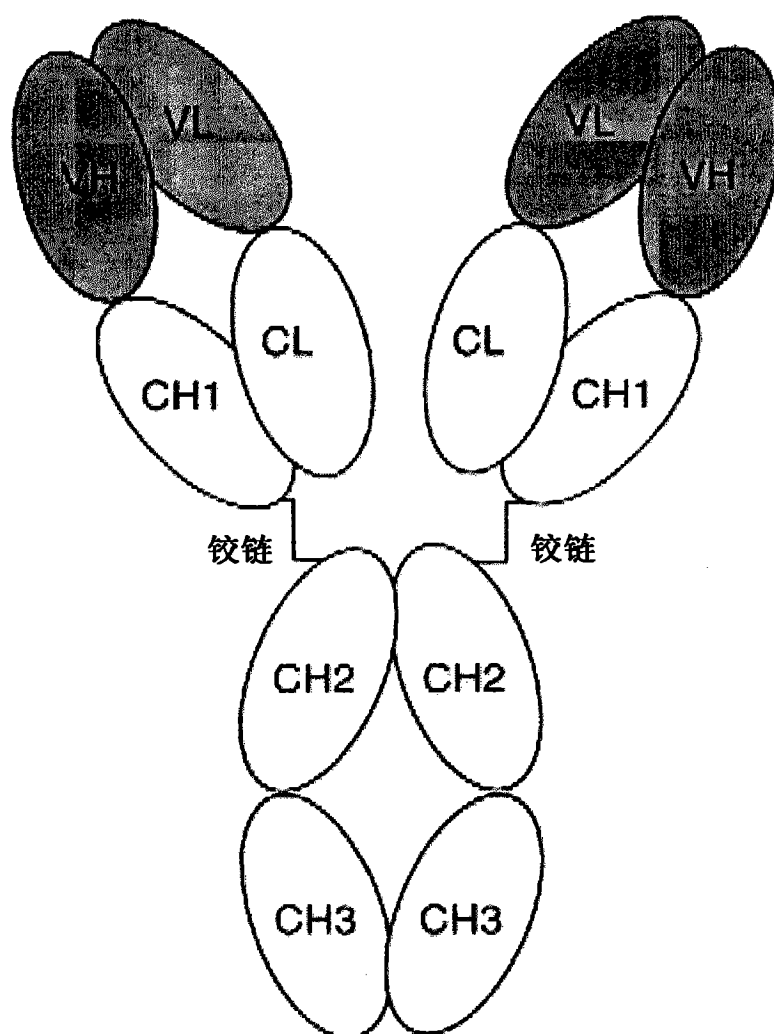


图 1A

	H1										H2									
	CDR1										CDR2									
人	EVQLV	ESGGGLVQ	PGGSLRL	S	CAASGFTFS	S	--YAMS	WVRQ	APGKGLEWVS	XISXKTDGGTYYYADSVKGG										
	5	10	15	20	25	30	35	40	45	50	abc	55	60	65						
Ab #1	HVQLGESGGGV	VRPGGSLRL	S	CAASGFTFS	S	--YAMS	WVRQ	APGKGLEWVS	AISG	--SGGSTYYADSVKGG										
Ab #2	QVQKVESGGGLVQ	PGGSLRL	S	CAASGFTFS	S	--YAMS	WVRQ	APGKGLEWVS	AISG	--SGGSTYYADSVKGG										
Ab #3	QVQLVQSGAEVKKP	ASVKV	SC	KASGYTFT	S	--YYMH	WVRQ	APGQGLEWVG	IINP	--SGGSTSYAQKFEQG										

	H3									
	CDR3									
人	RFTISR	DNS	-KNTLYLQ	MNSLRAEDTAVYYCAR	xxxxxxx	xxxxxxx	xxxxxxx	-----FDY	WGQGT	LVTVSS (SEQ ID NO: 1)
	70	75	80	abc	85	90	94	100	b d f h j	105 110
Ab #1	RFTISR	RNA	-KNSLYLQ	MNSLRAEDTAVYYCAR	RGVWA	-----FDI	WGQGT	MVTVSS		(SEQ ID NO: 2)
Ab #2	RFTISR	DNS	-KNTLYLQ	MNSLRAEDTAVYYCAK	DLGRQQL	-----SDY	WGQGT	LVTVSS		(SEQ ID NO: 3)
Ab #3	RVTMTR	DT	S	-TSTVYME	LSSLRSEDTAVYYCAT	PYG	-----MDV	WGQGT	TVTVSS	(SEQ ID NO: 4)

图 1B

102

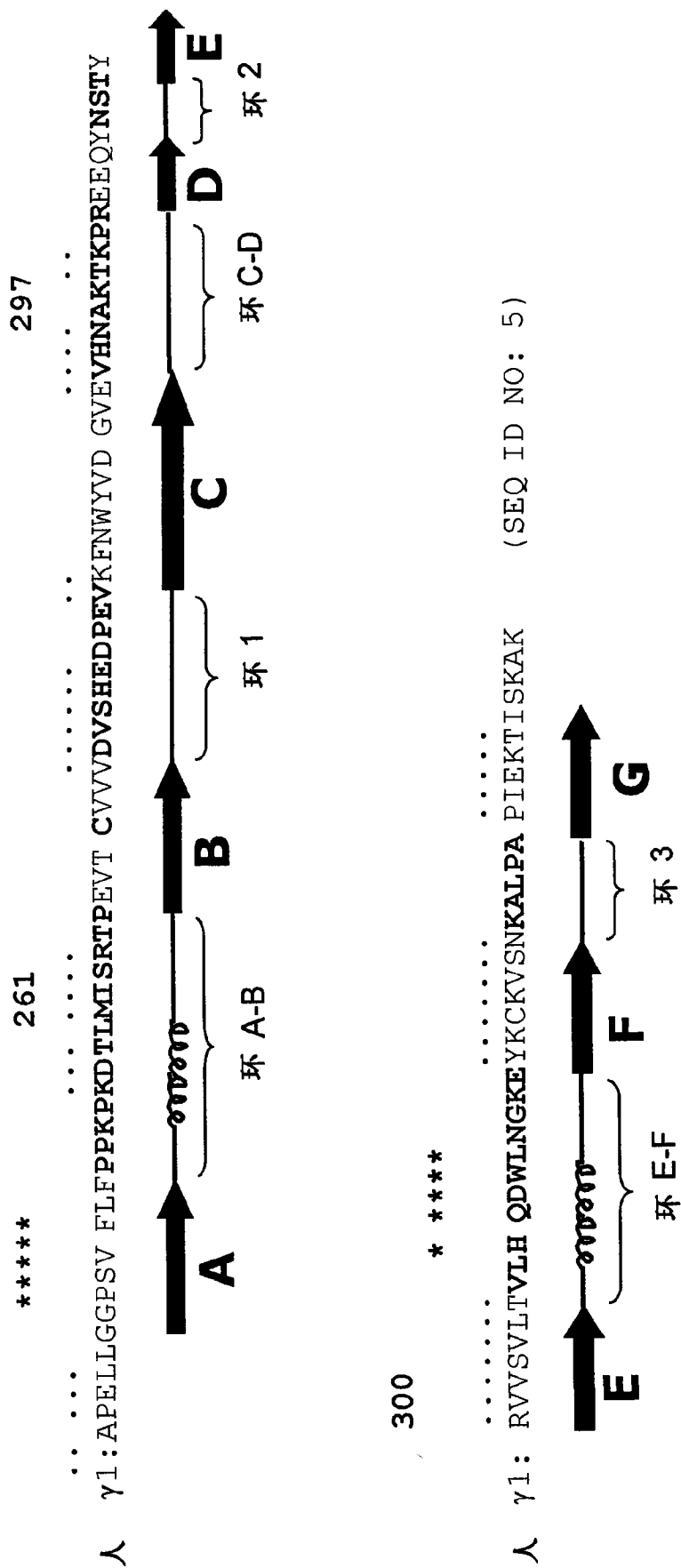


图 2

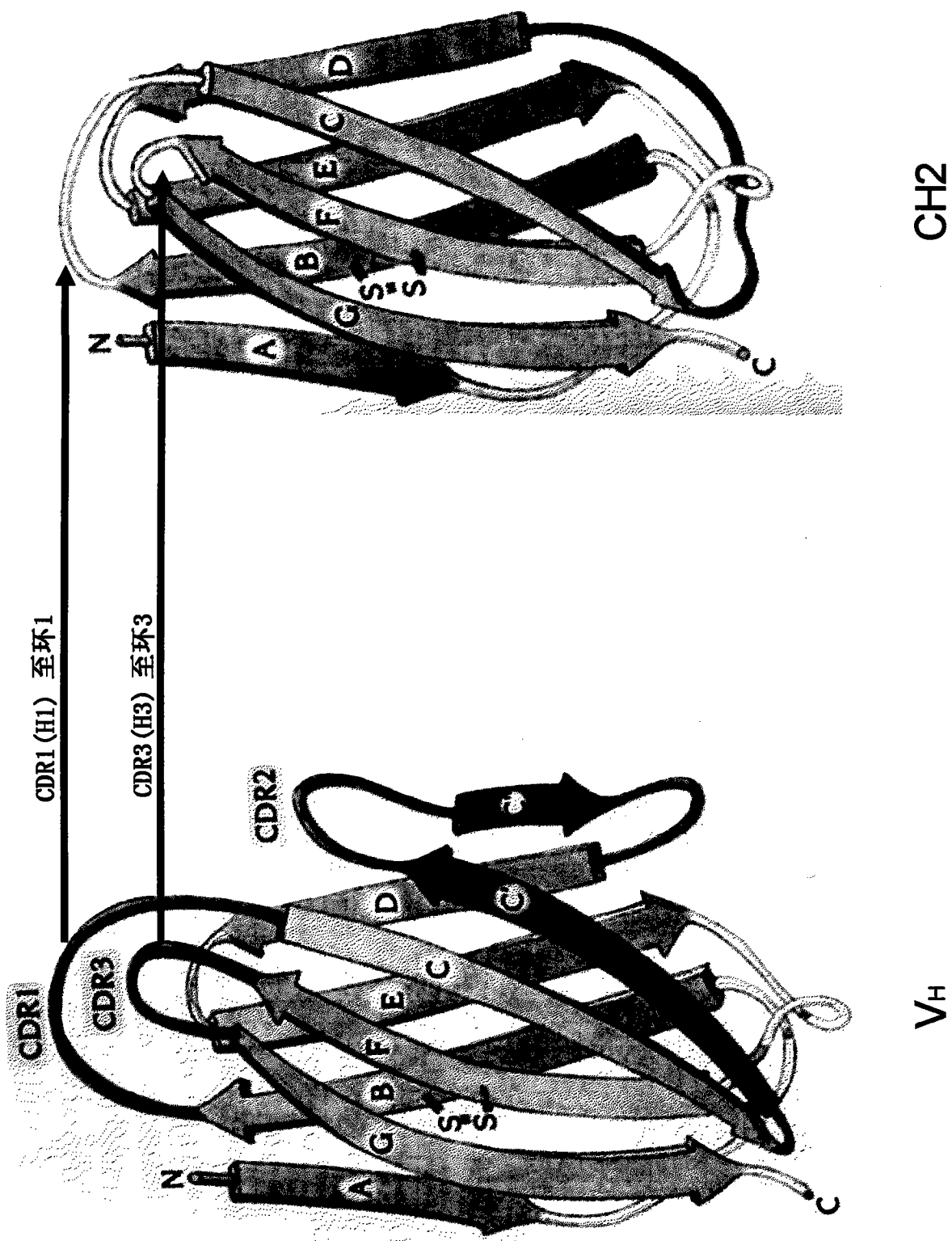


图 3A

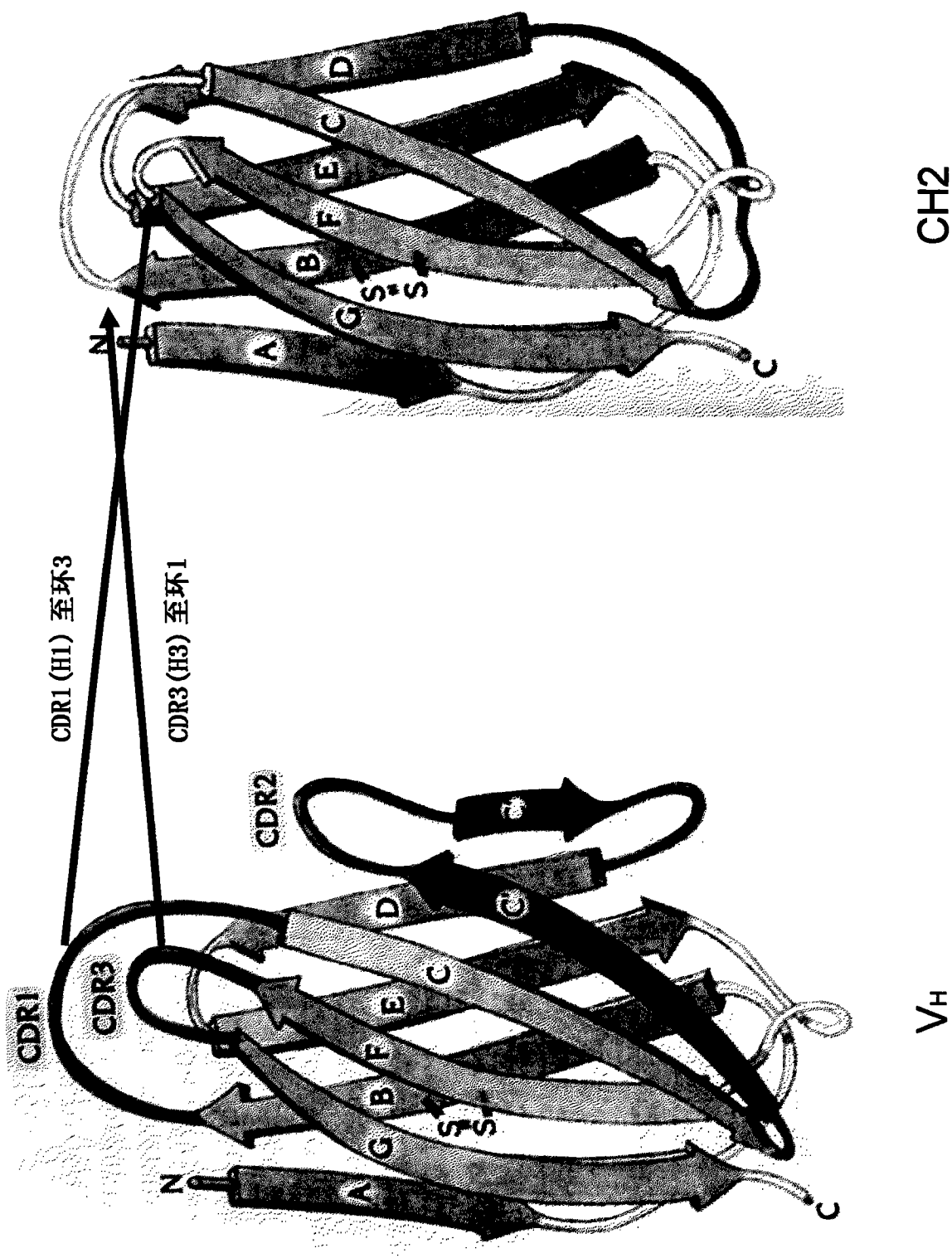


图 3B

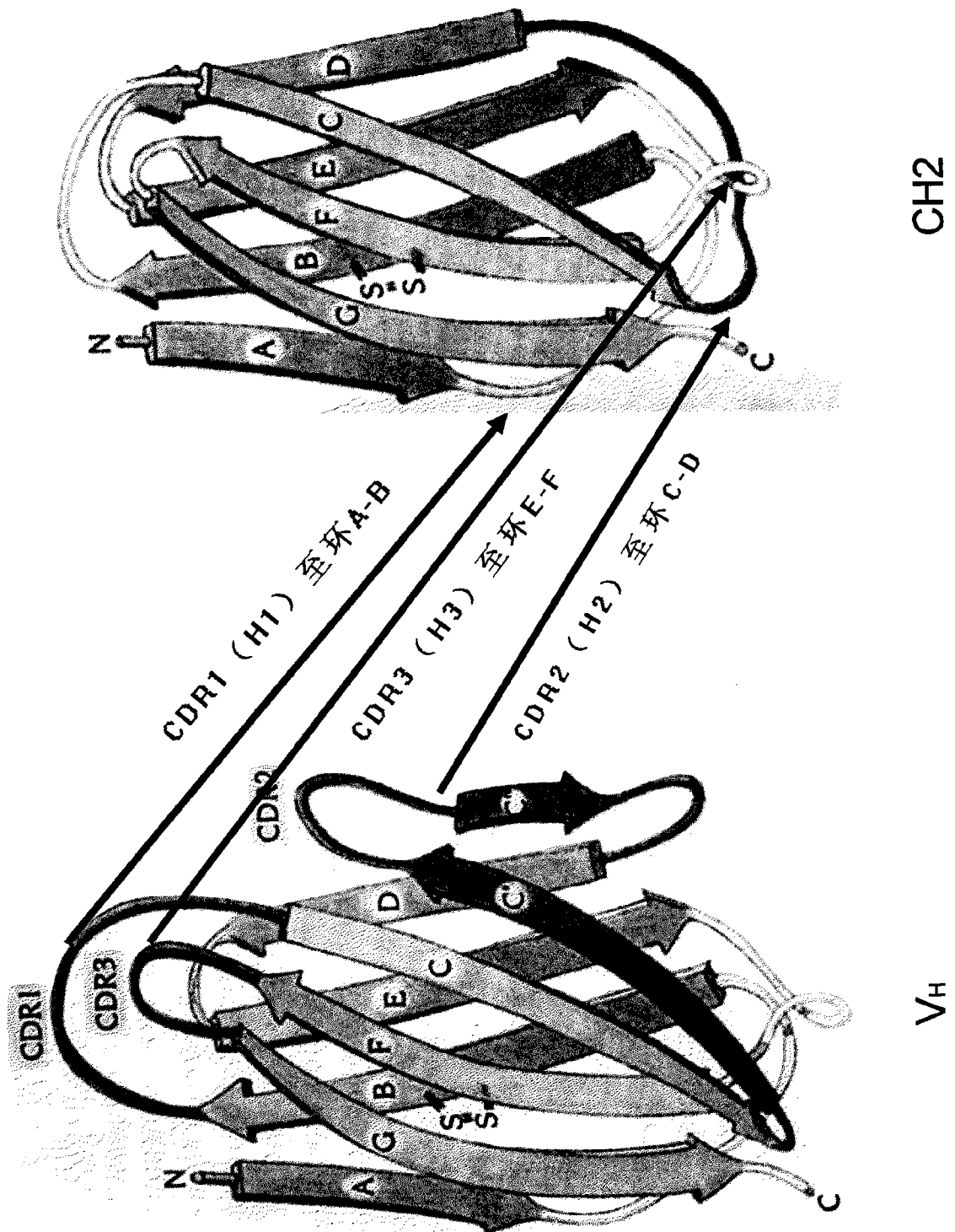


图 3C

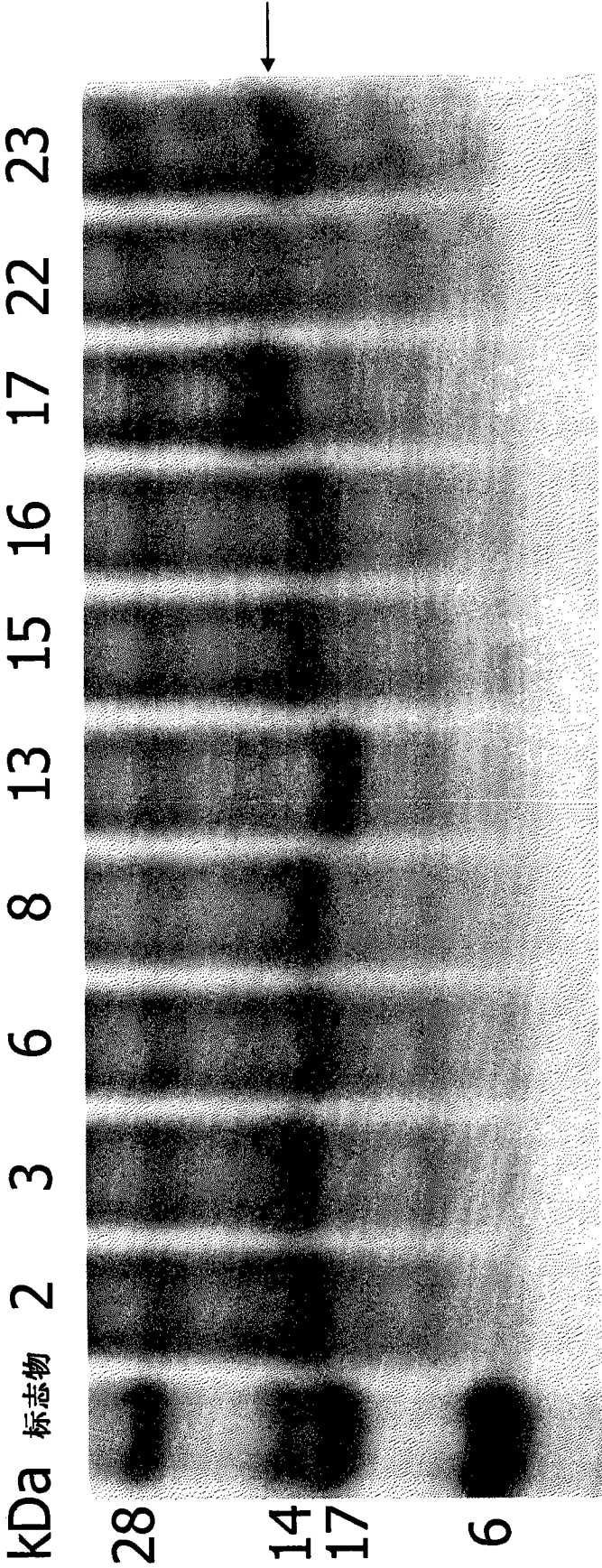


图 4

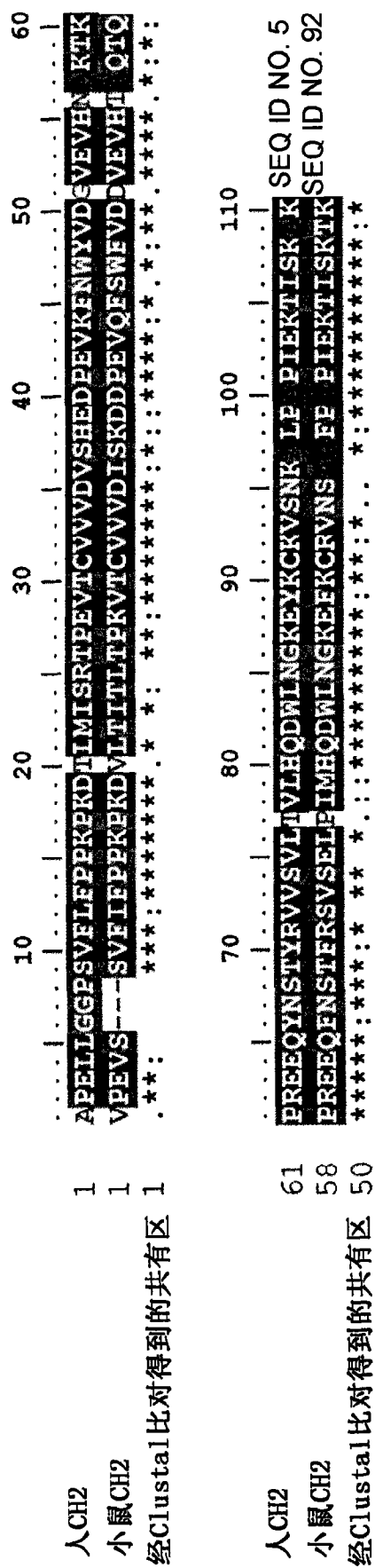


图 5A

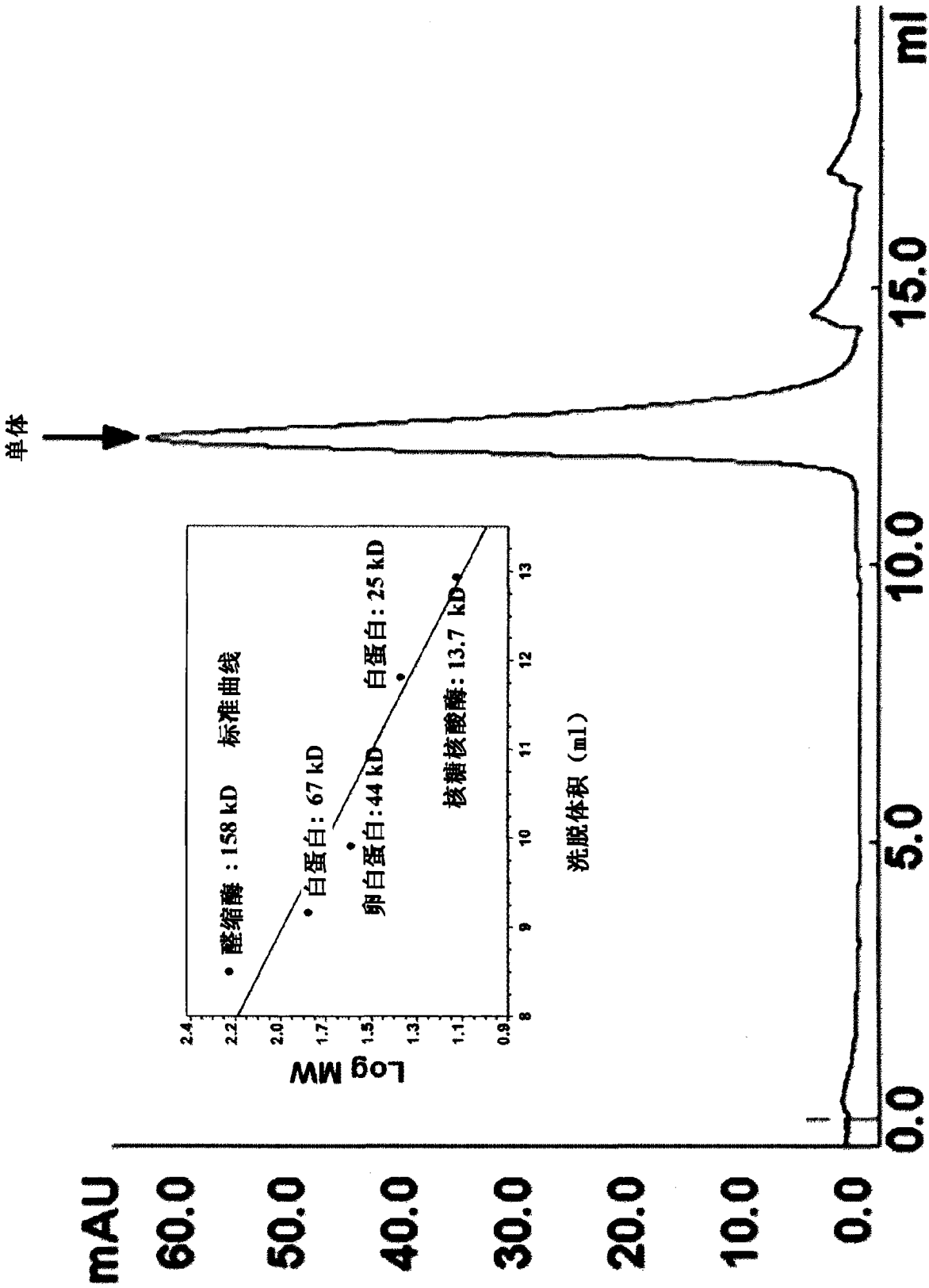


图 5B

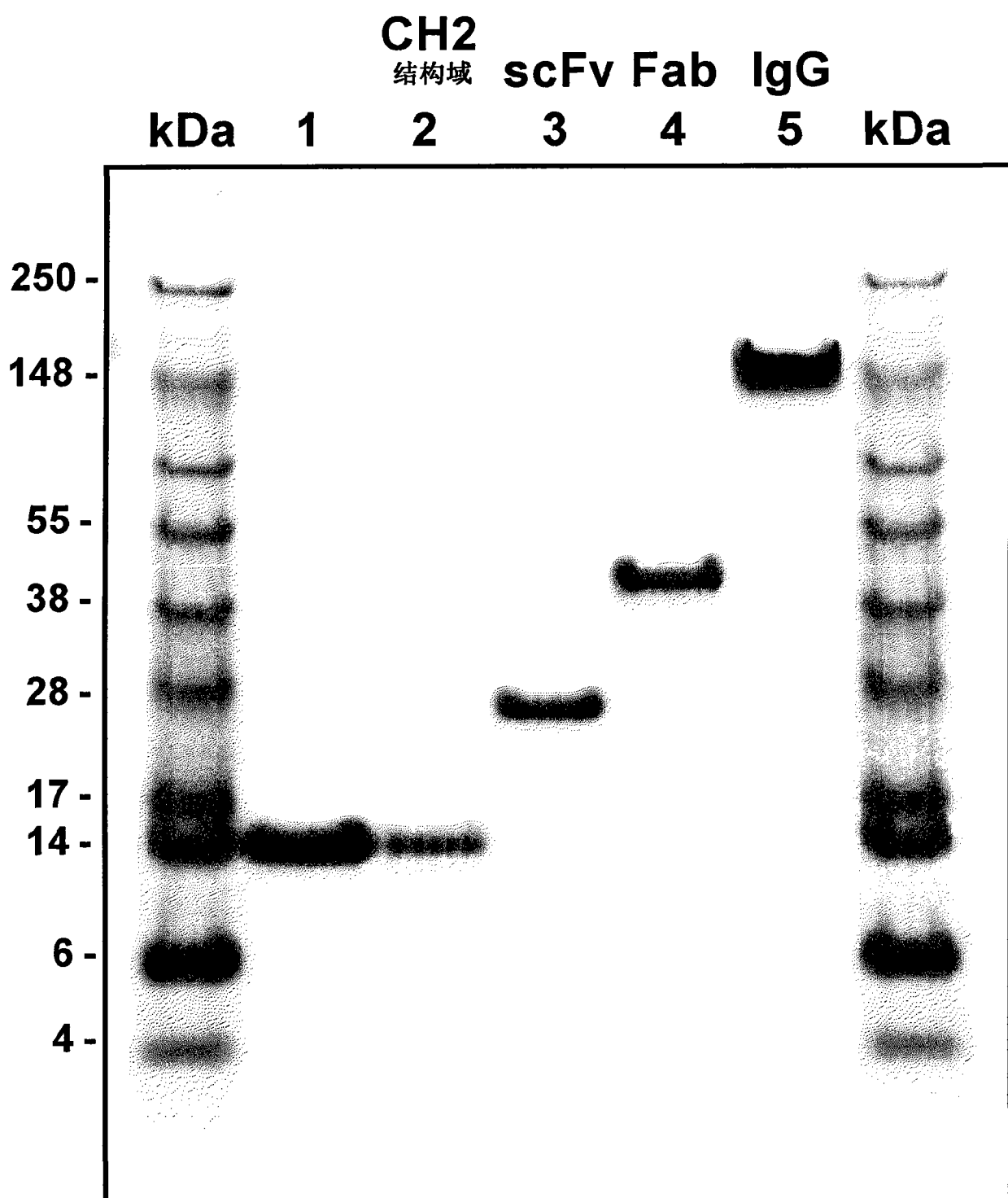


图 5C

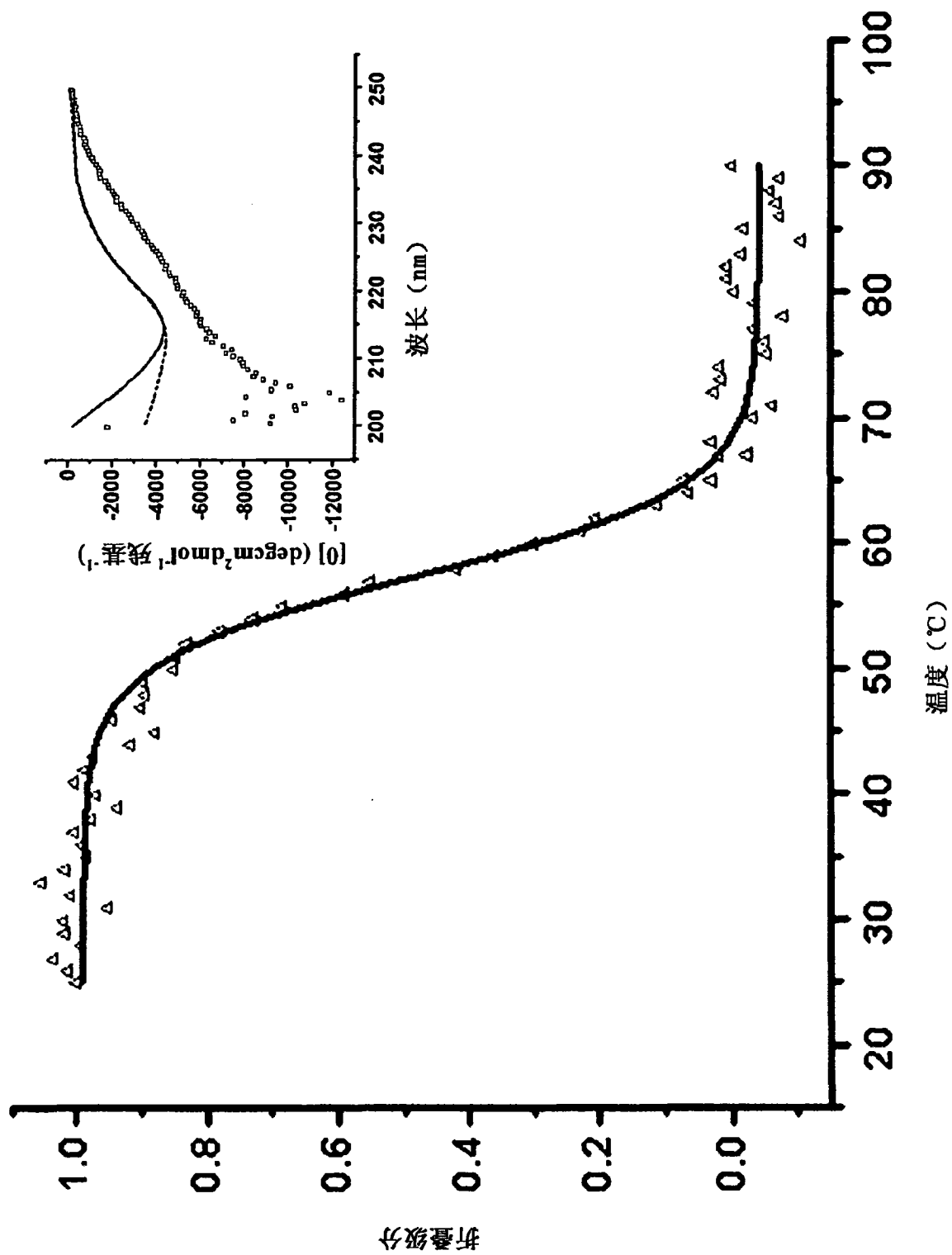


图 6A

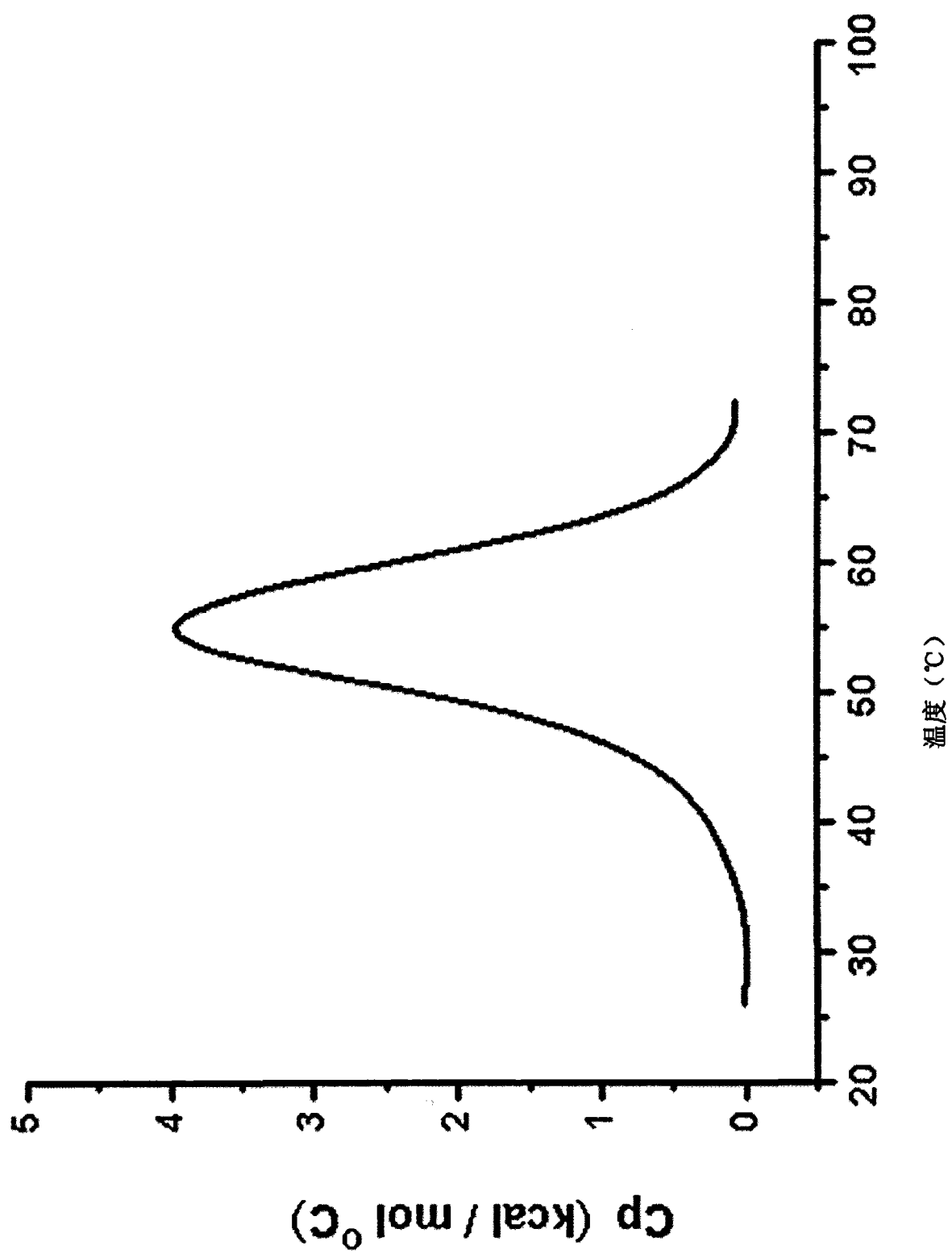


图 6B

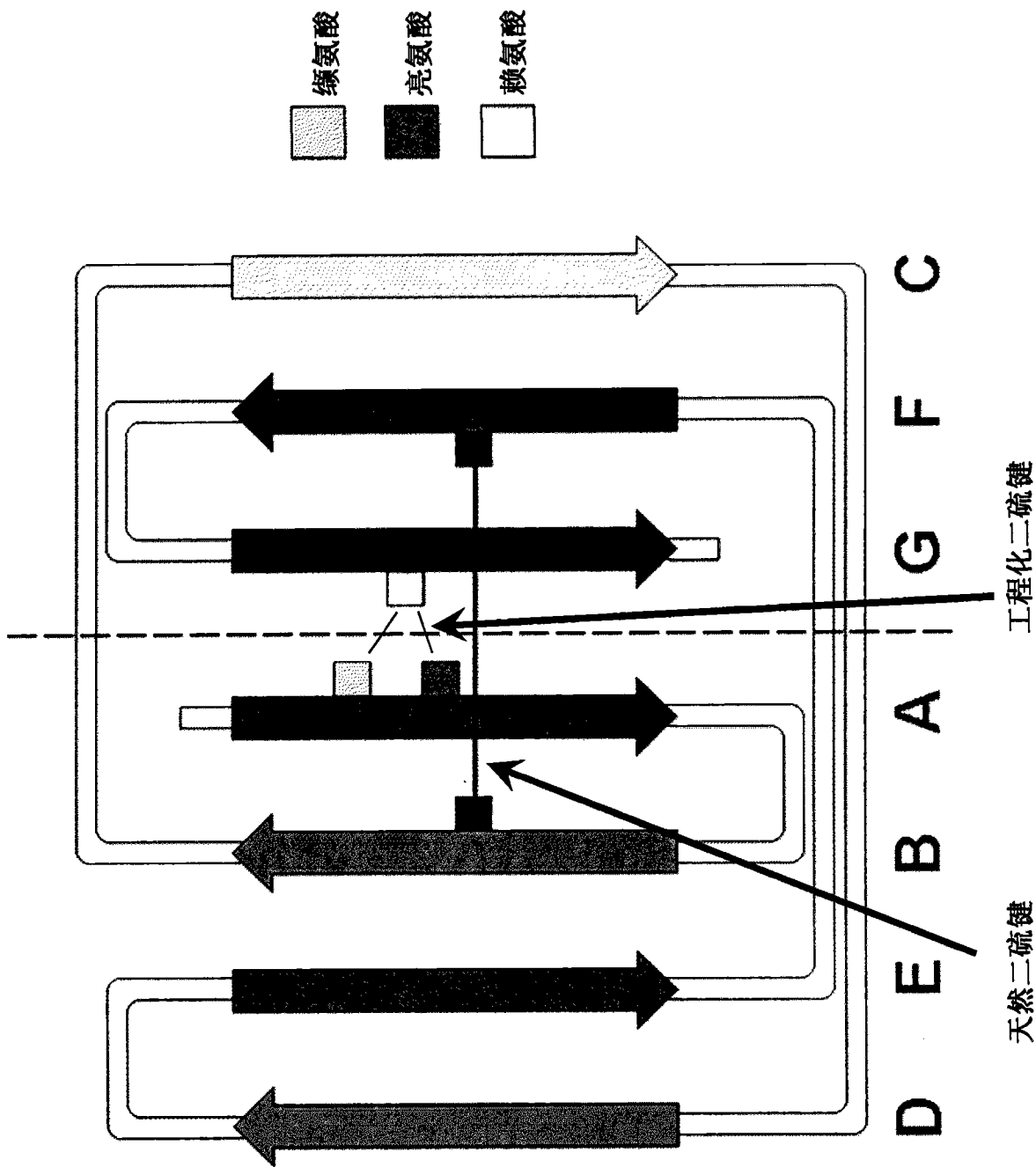


图 7

m01 m02 CH2 M

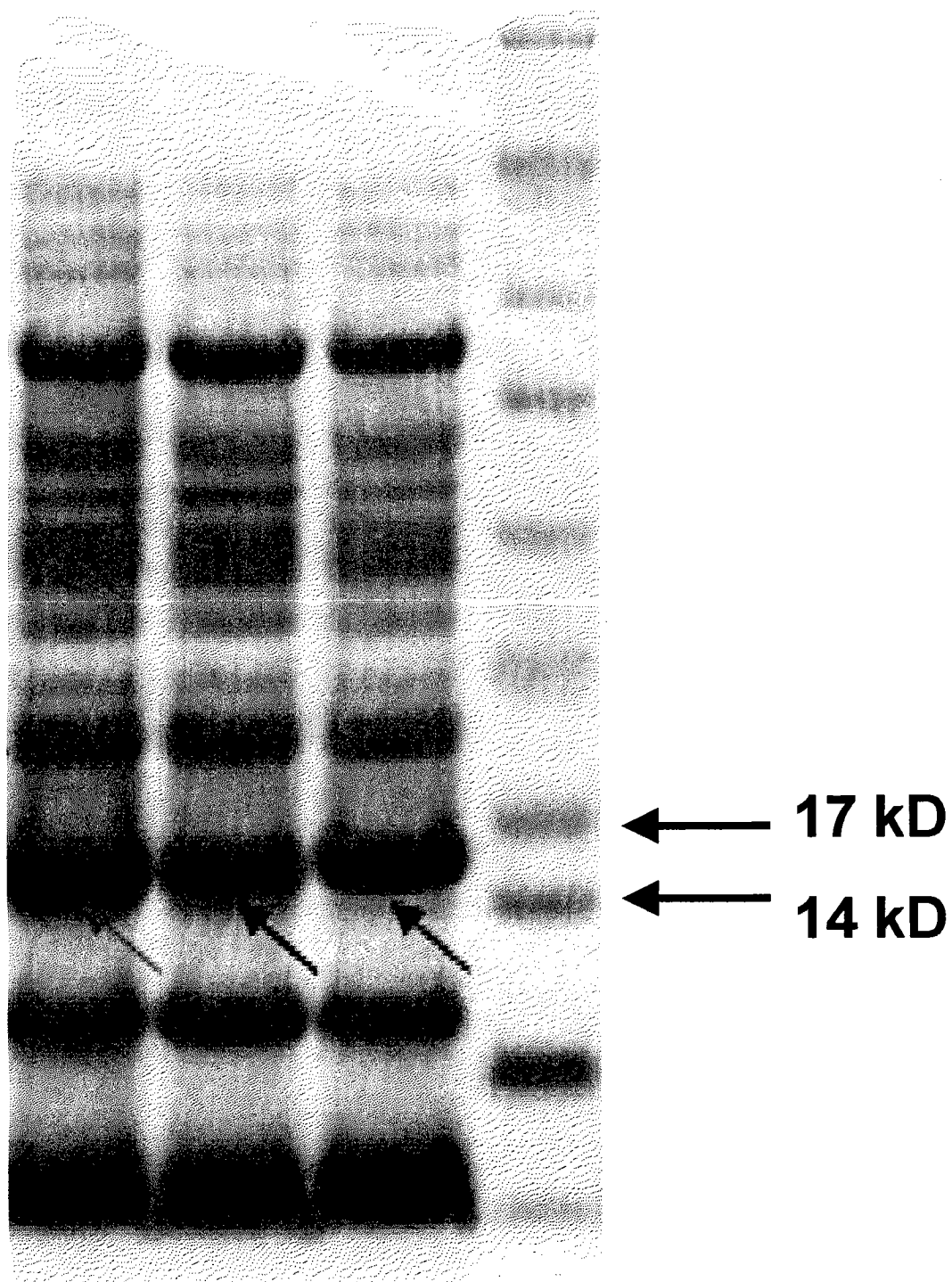


图 8

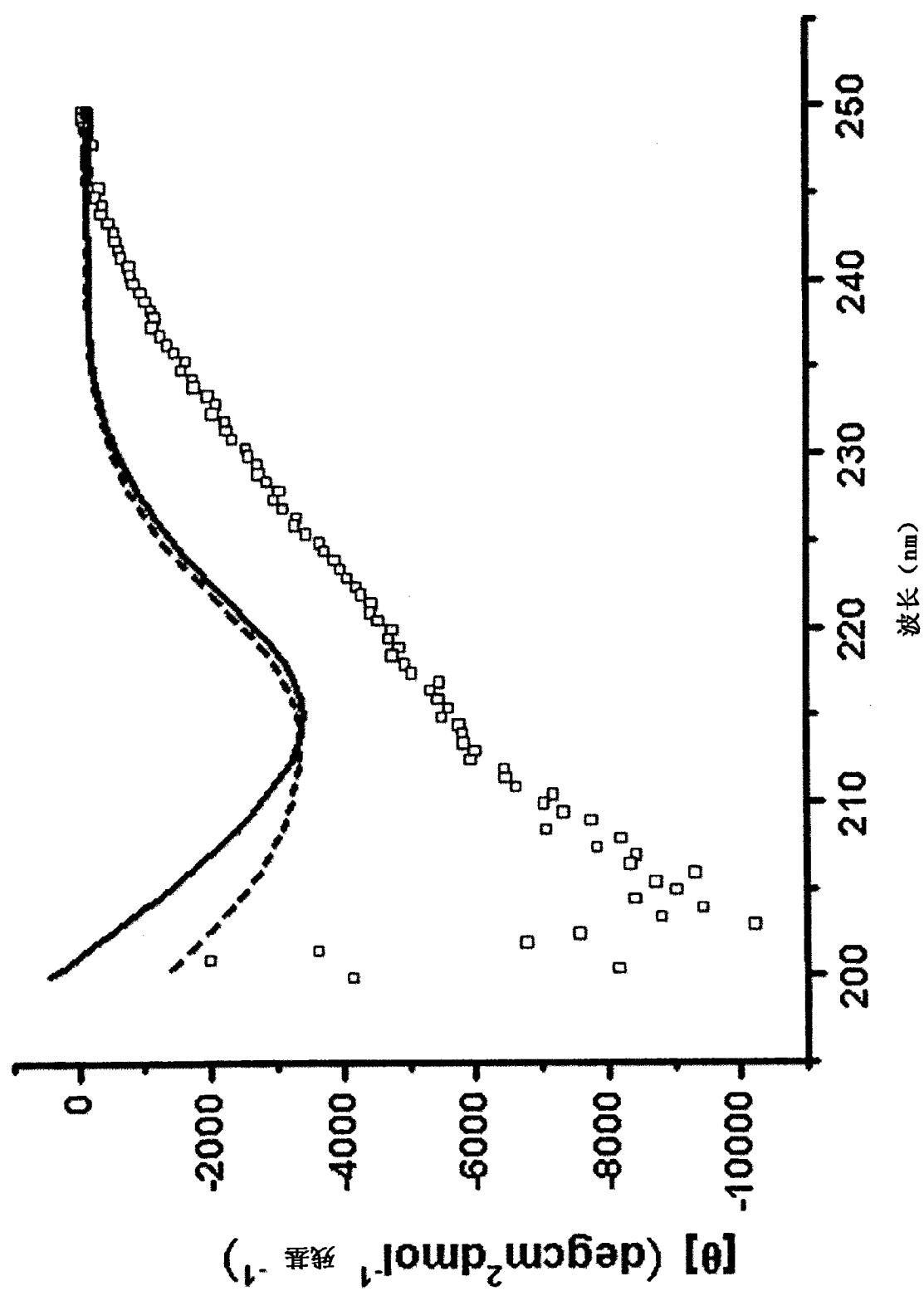


图 9A

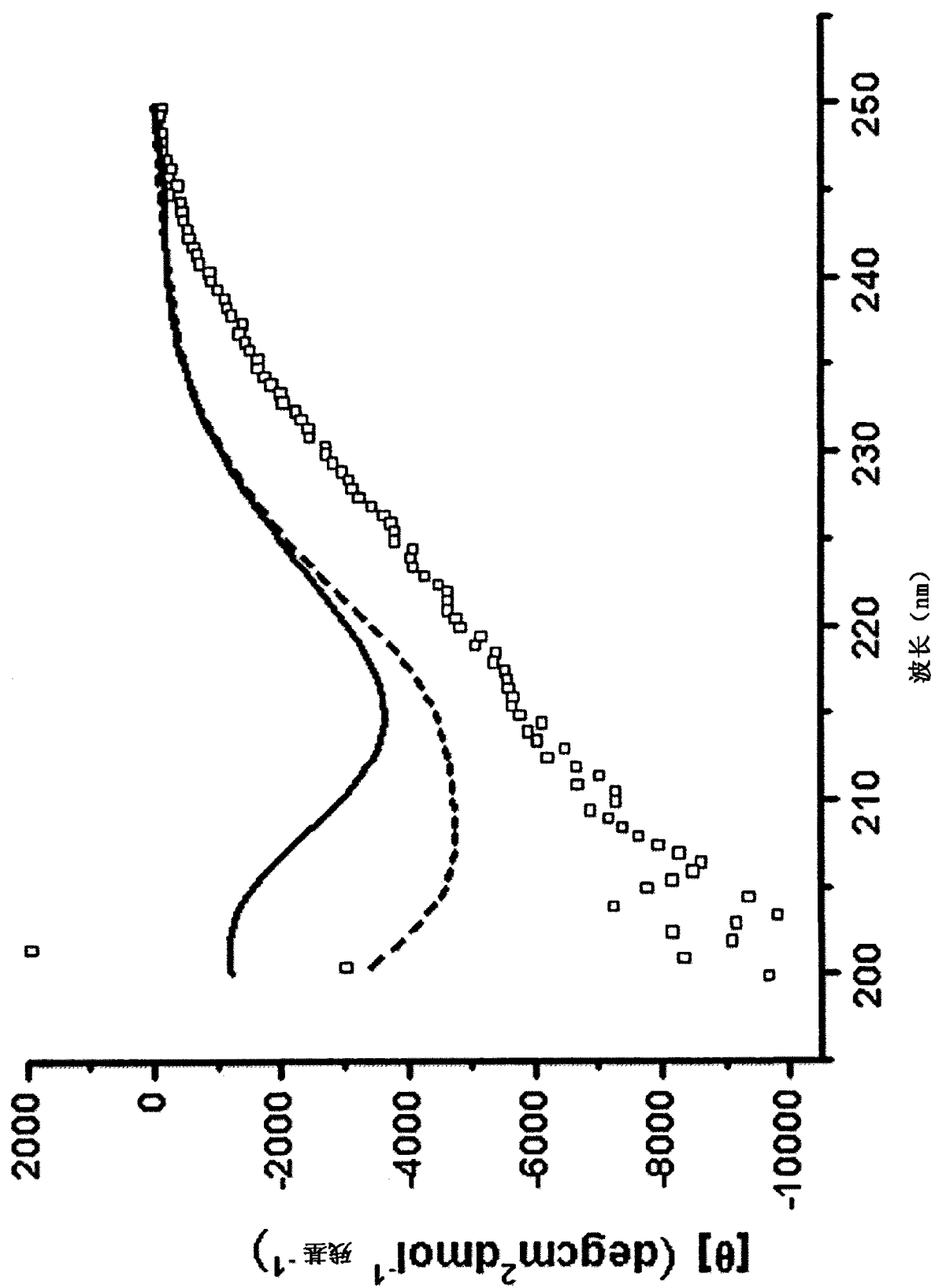


图 9B

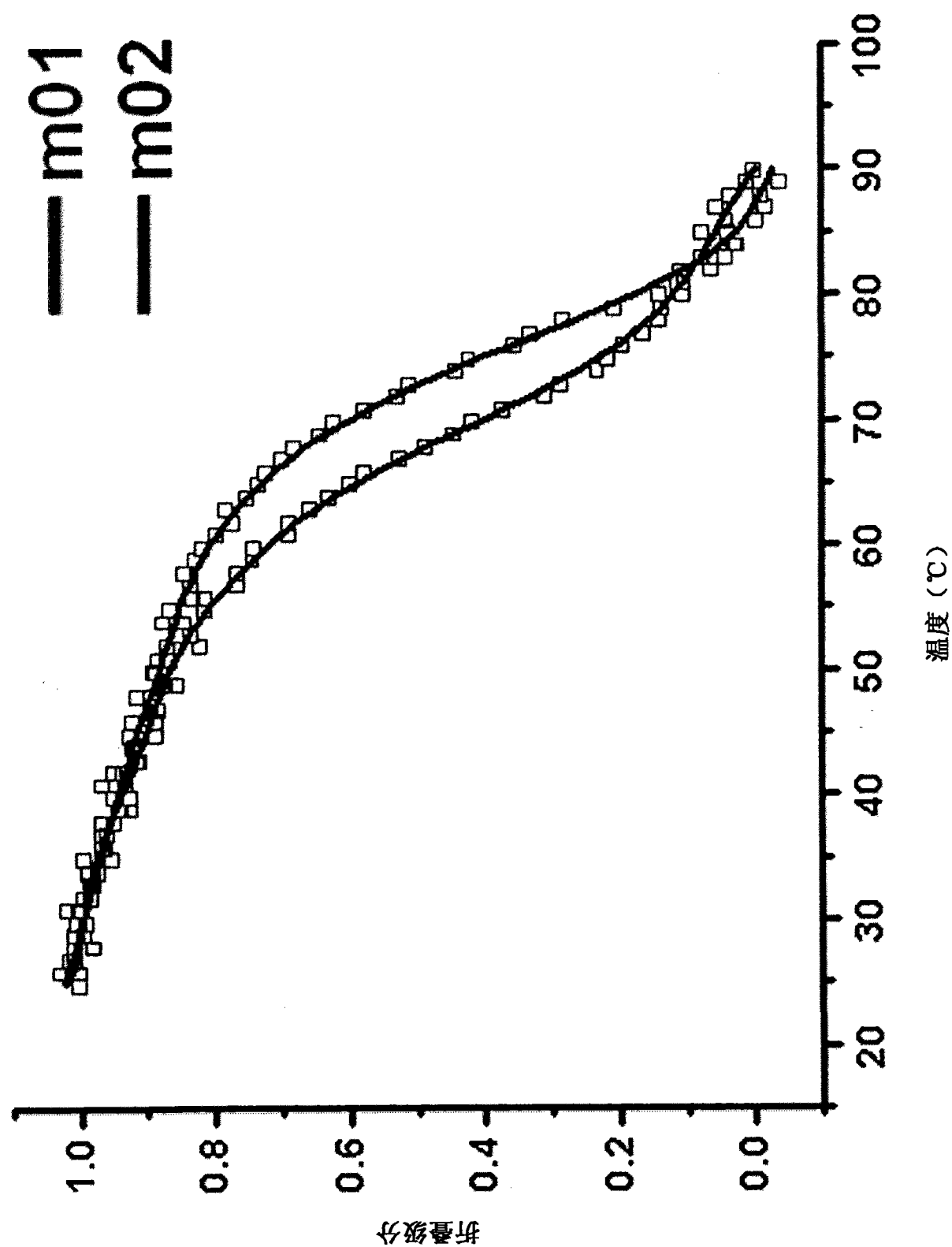


图 9C

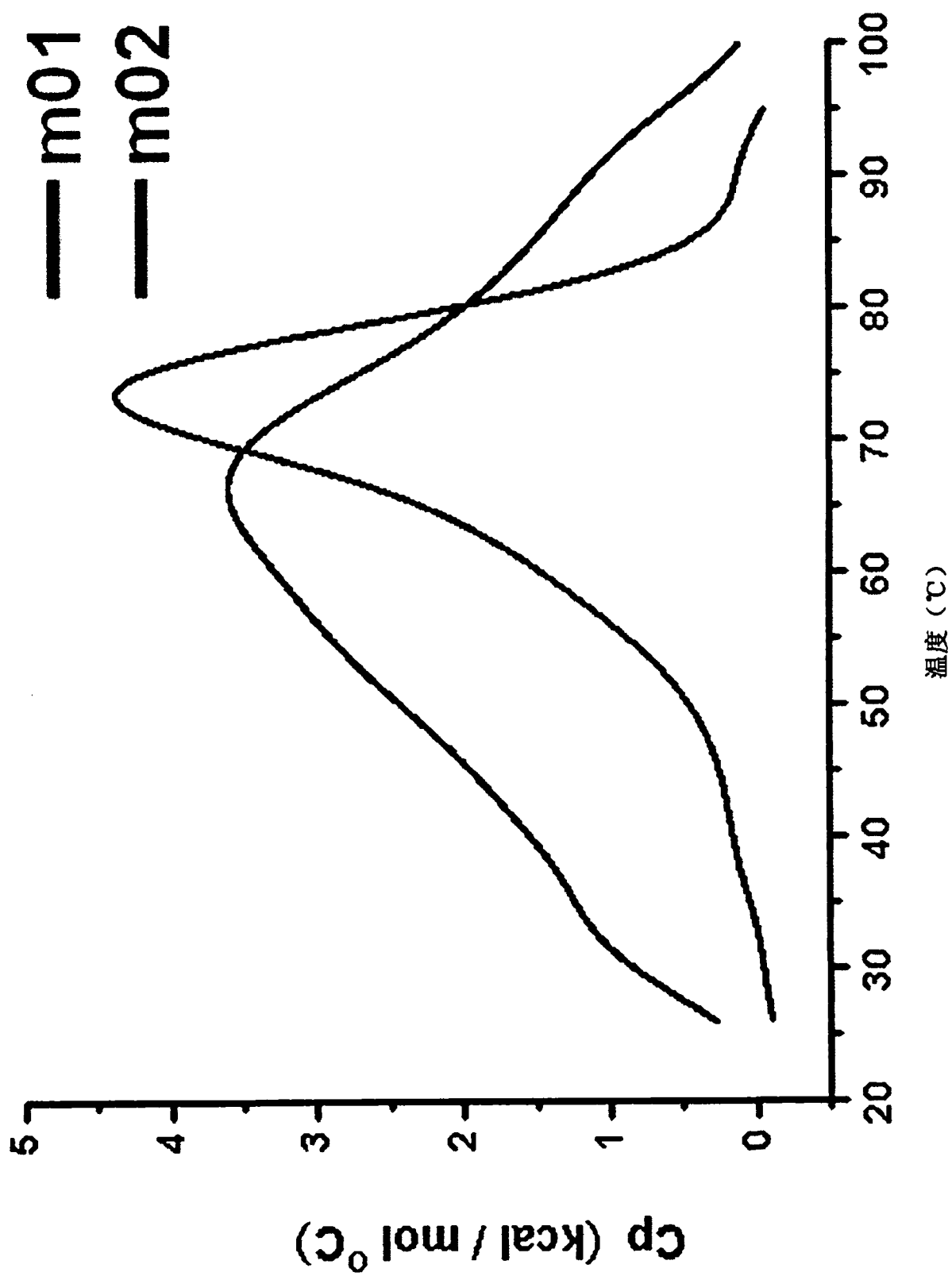


图 9D

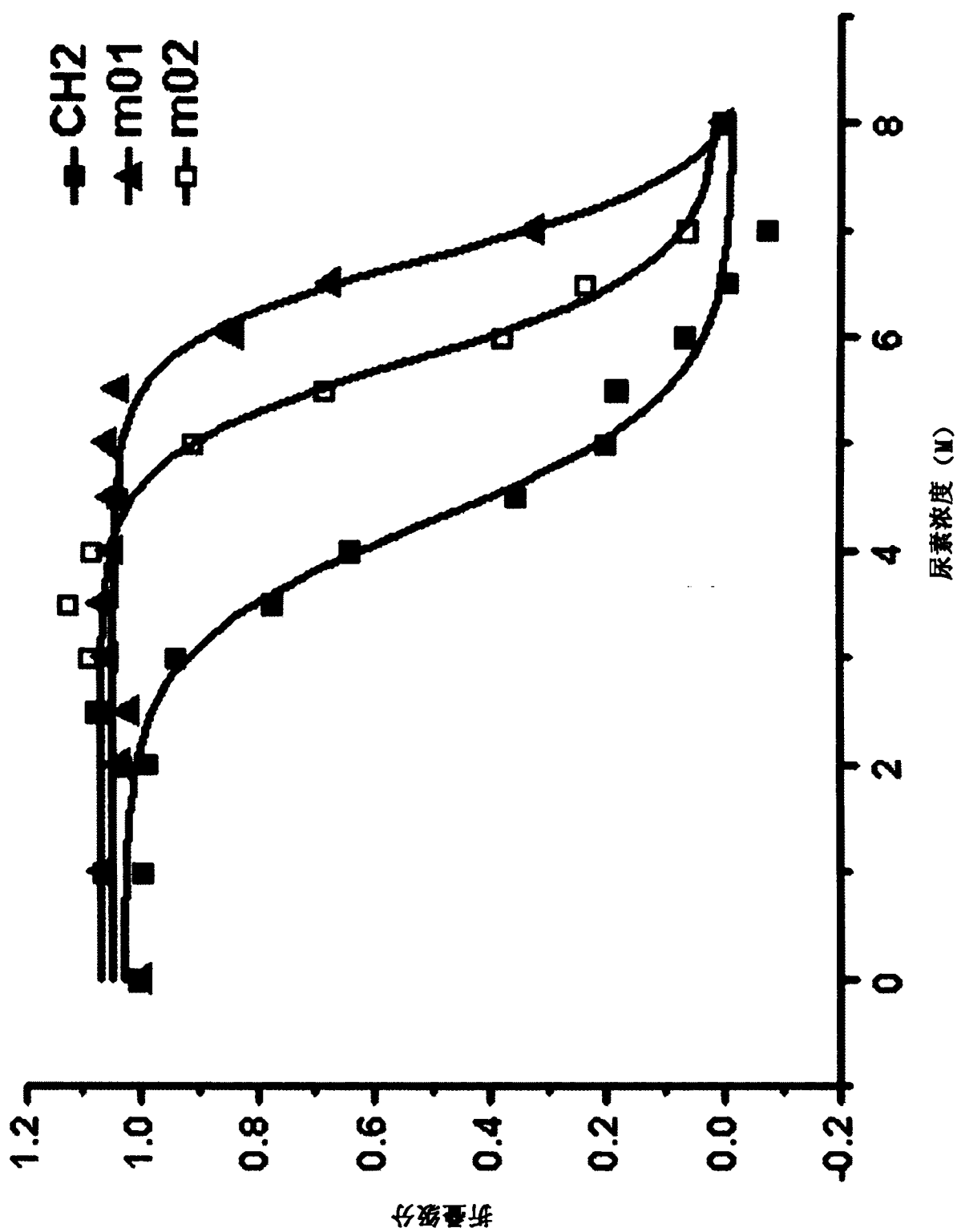


图 9E

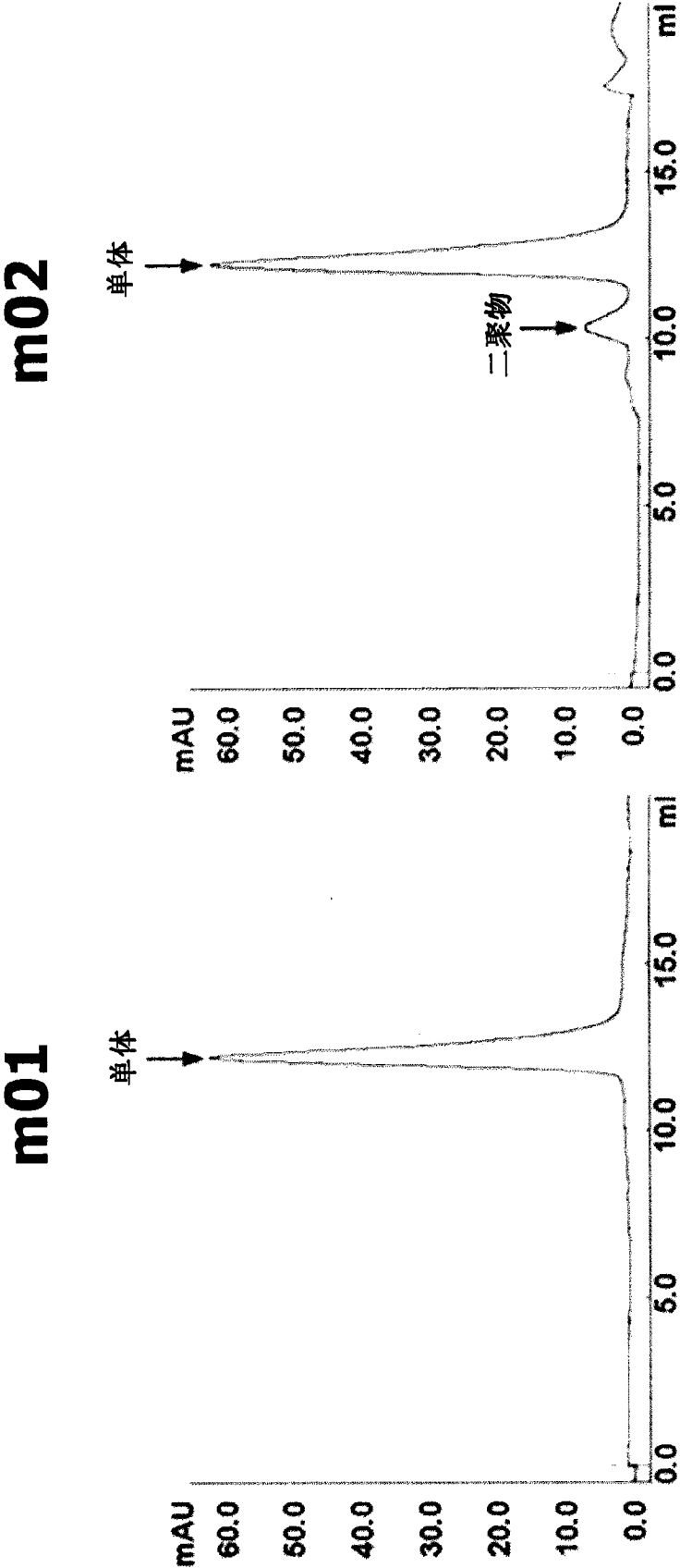


图 10A

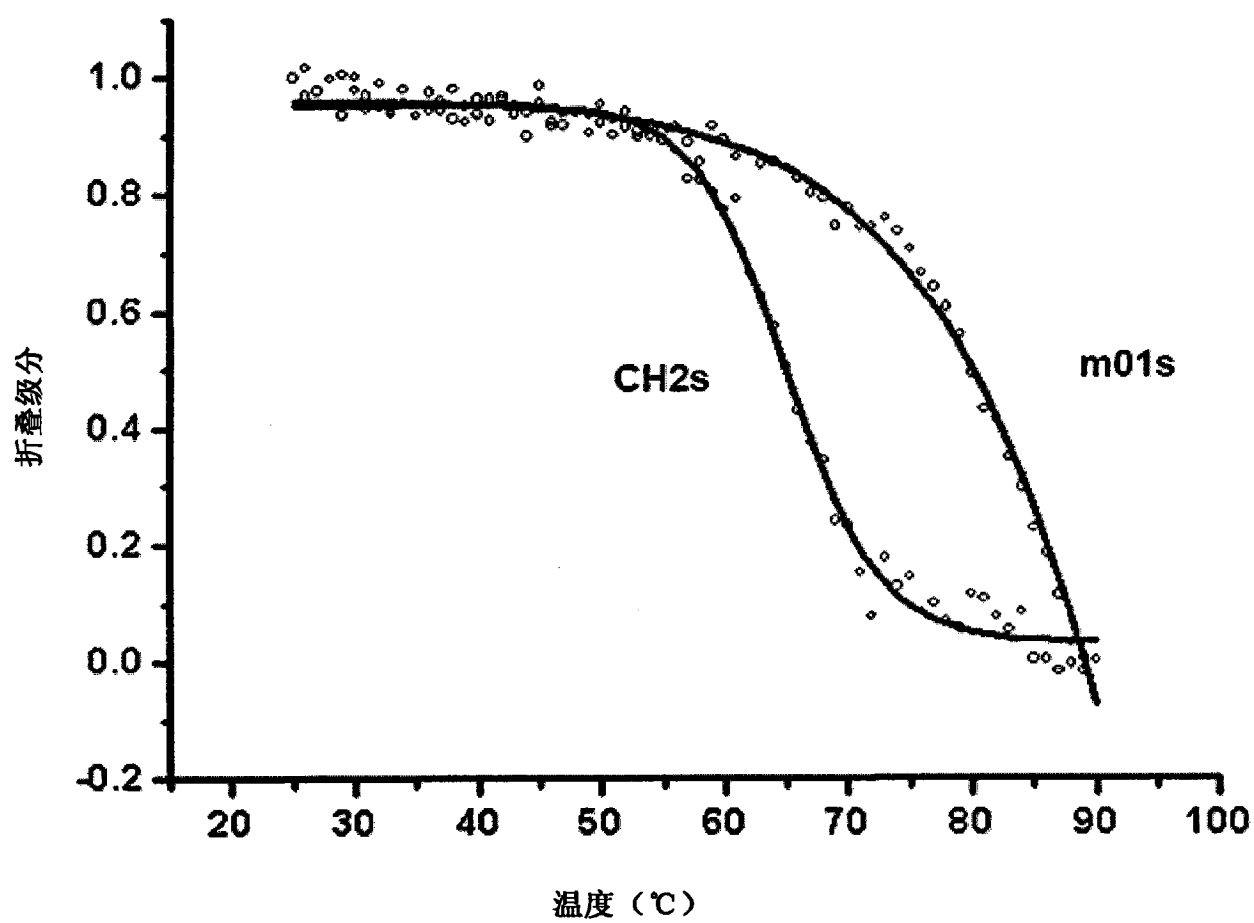


图 10B

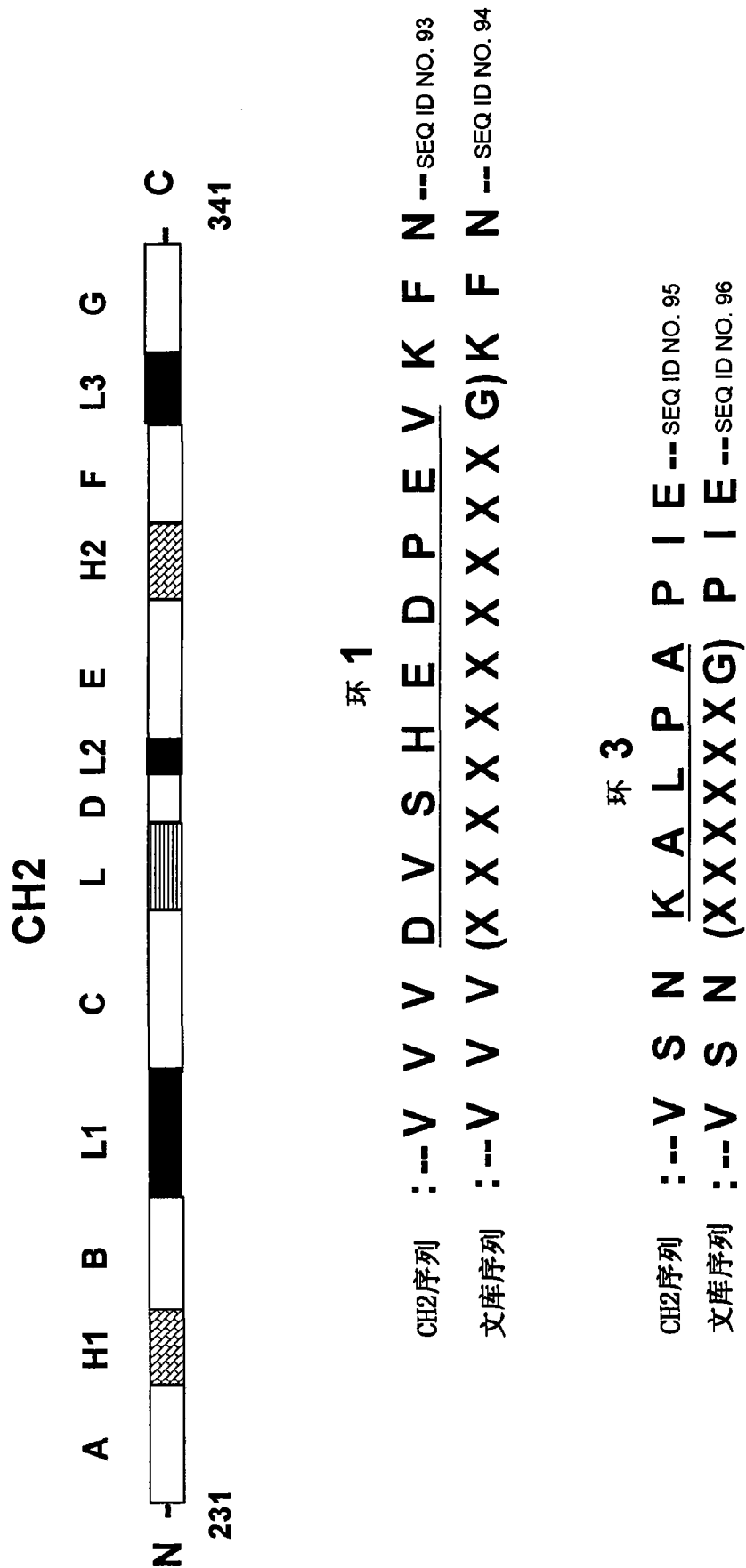

$$X=Y/A/D/S$$

图 11

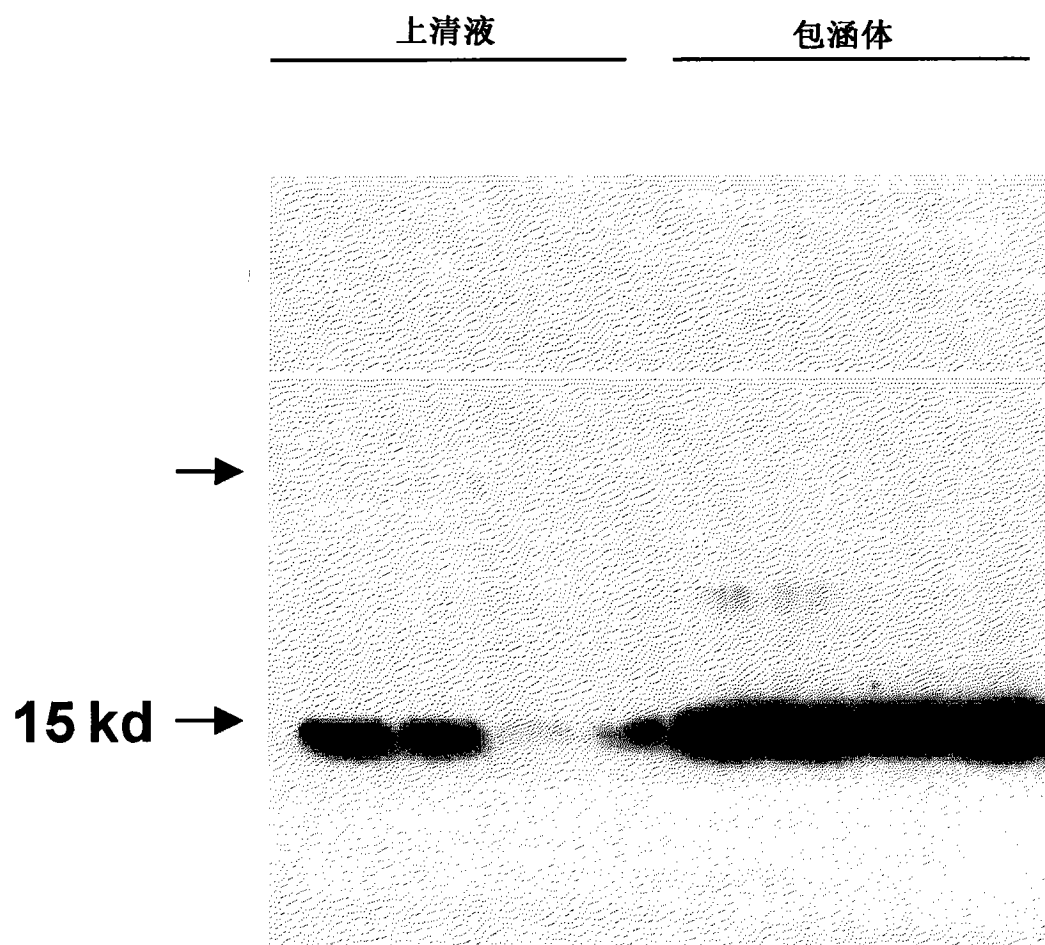


图 12A

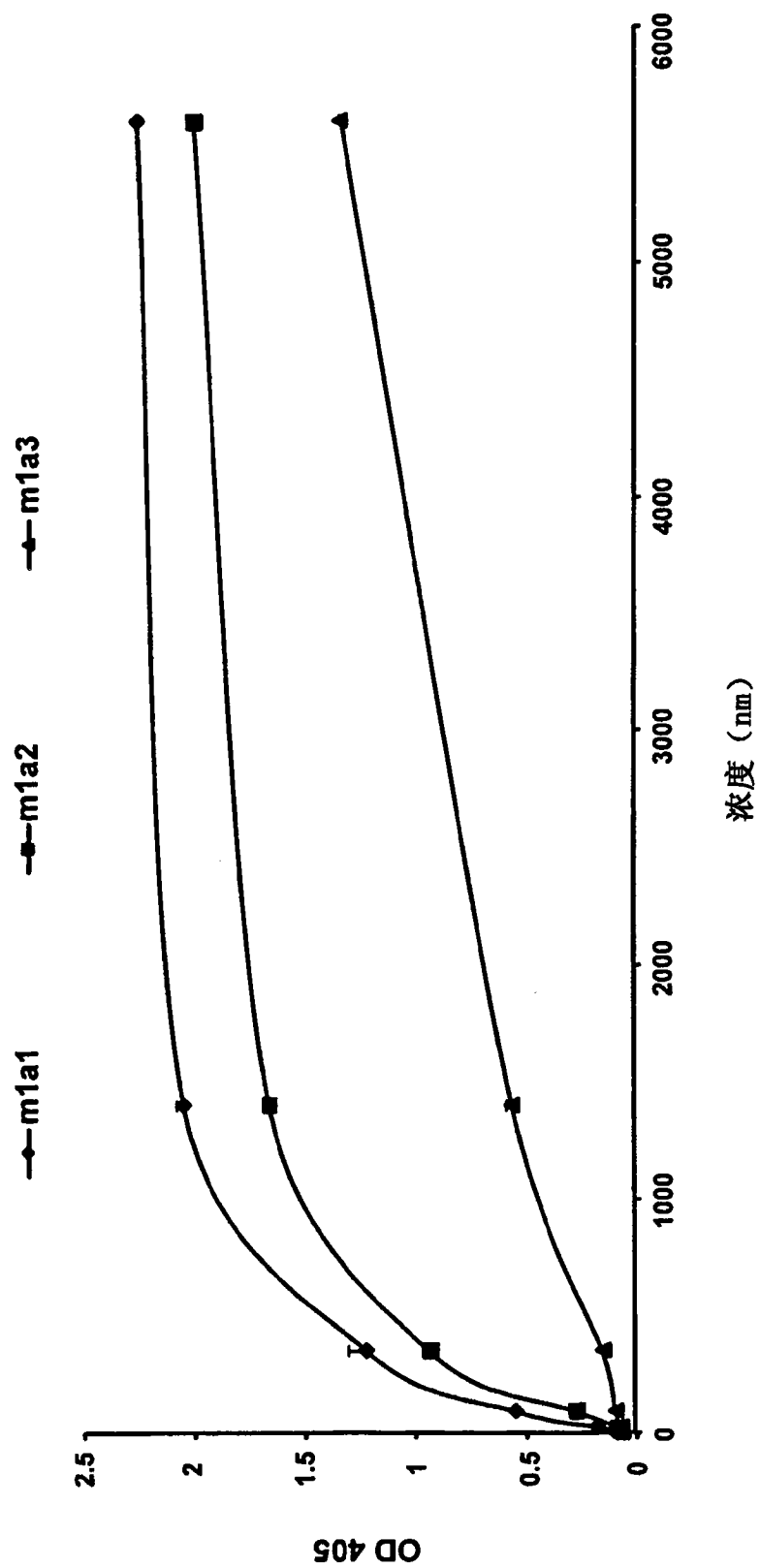


图 12B

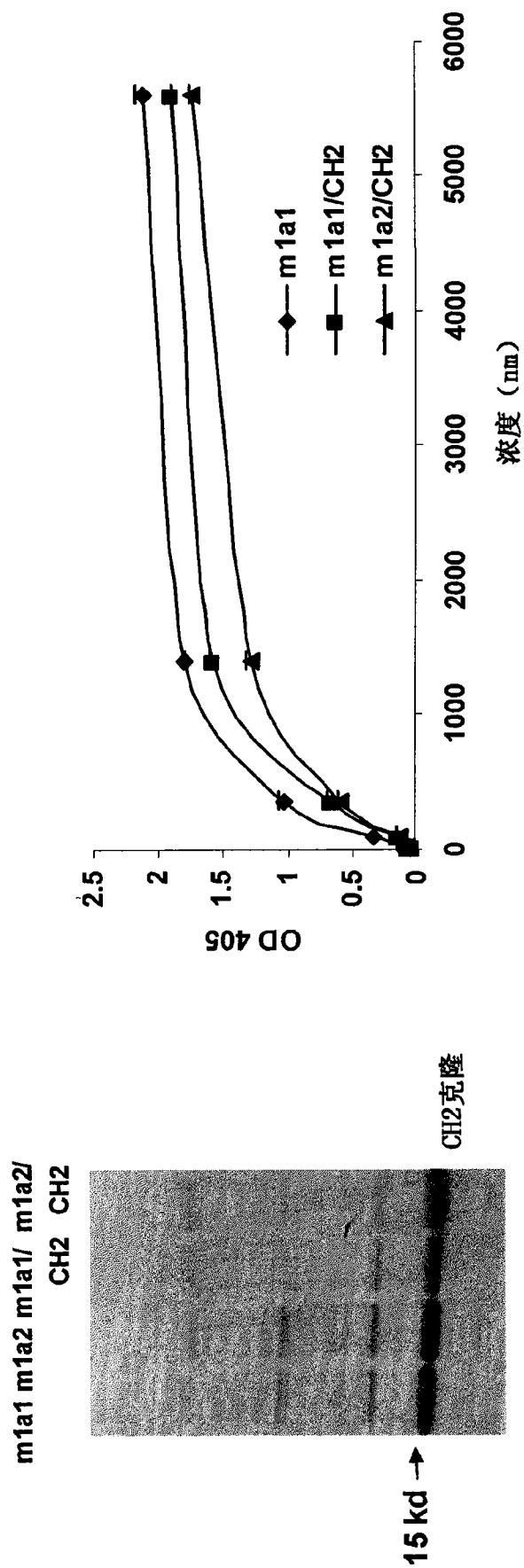


图 13A

m1a1m1a1s

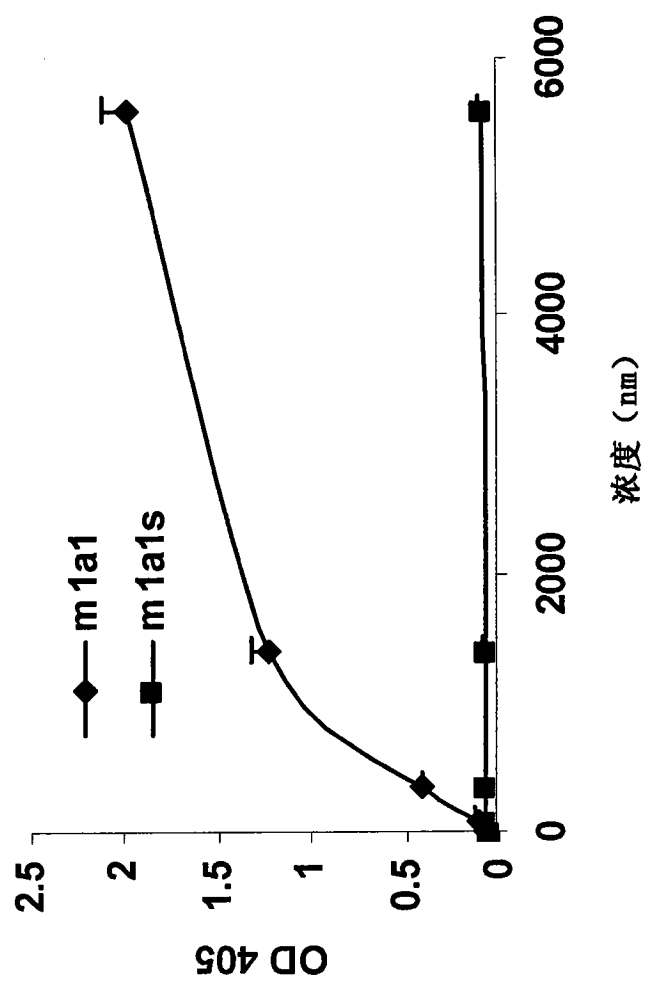
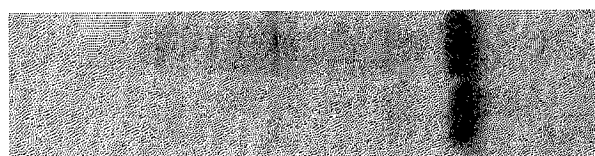


图 13B

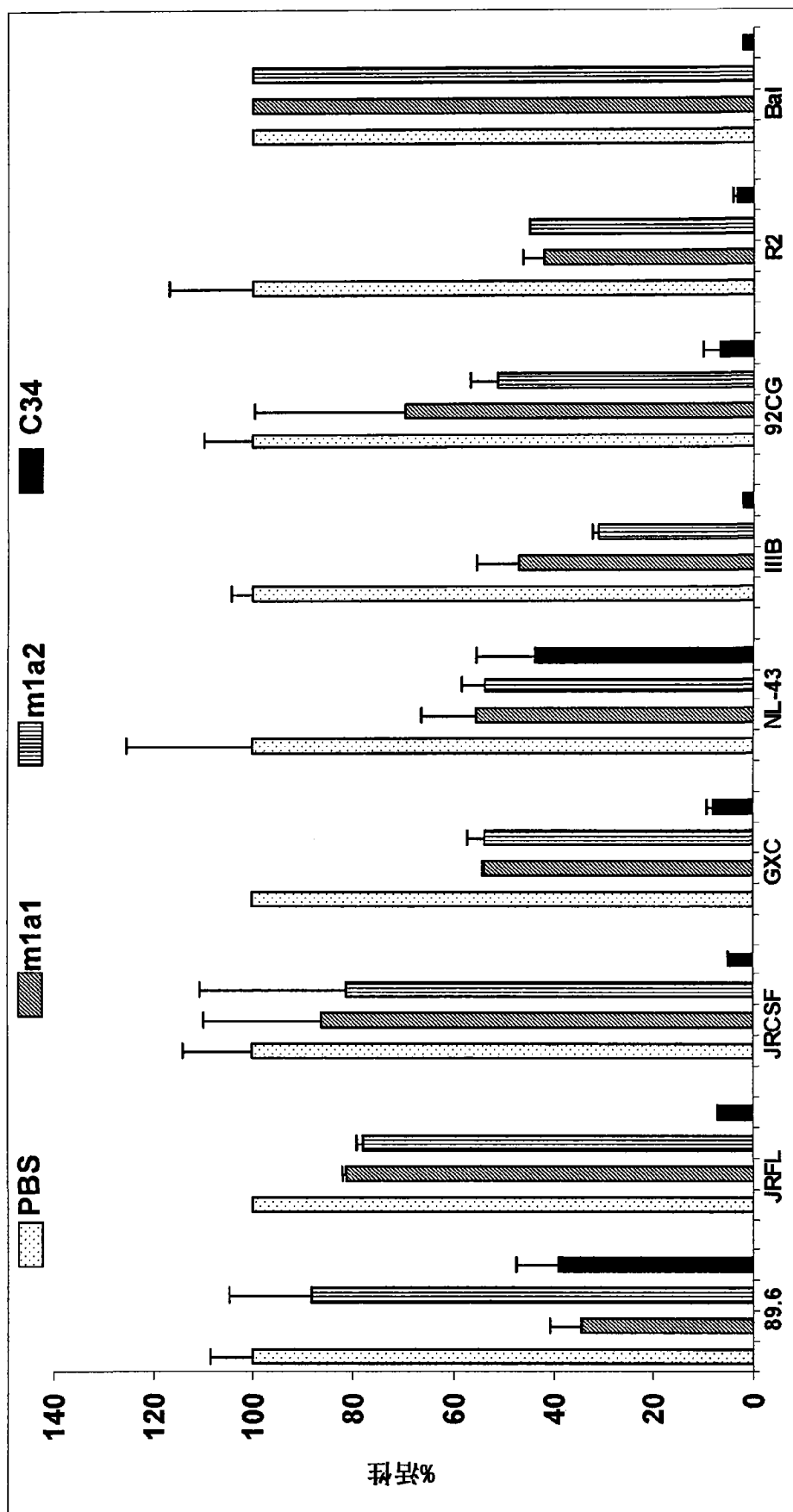


图 14

环2

CH2序列 :--Q Y N S T Y R--SEQ ID NO. 97

文库序列 :--Q Y (X X X) Y R--SEQ ID NO. 98

环3

CH2序列 :--S N K A L P A P I--SEQ ID NO. 99

文库序列 :--S N (X X X P X) P I--SEQ ID NO. 100

X = N / T / I / D / A / V / S / Y / F

图 15

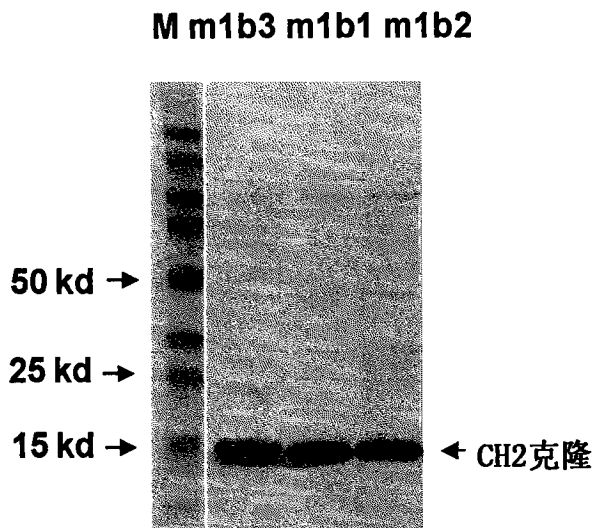


图 16A

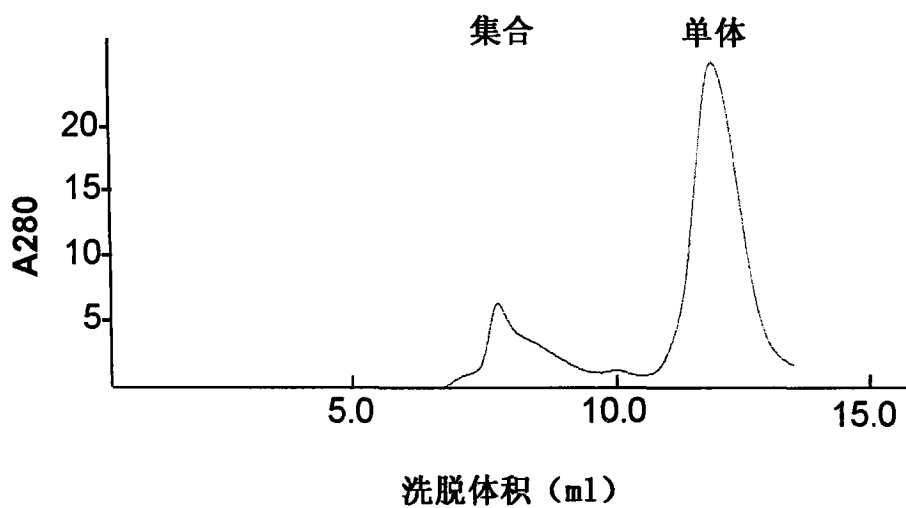


图 16B

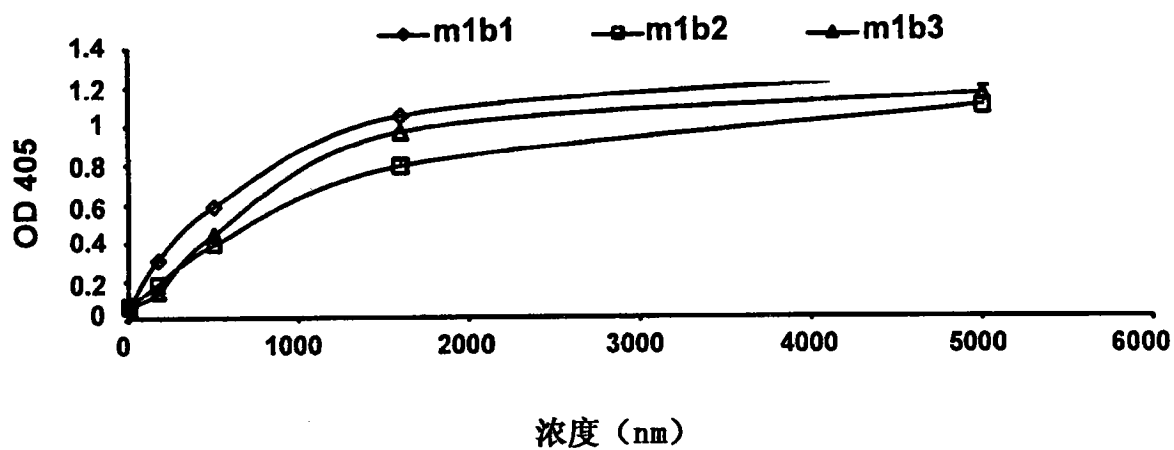


图 16C

L2: CH2 ---NST---
m1b3 ---ATD---

L3: CH2 ---KALPA---
m1b3 ---TTAPT---

图 16D

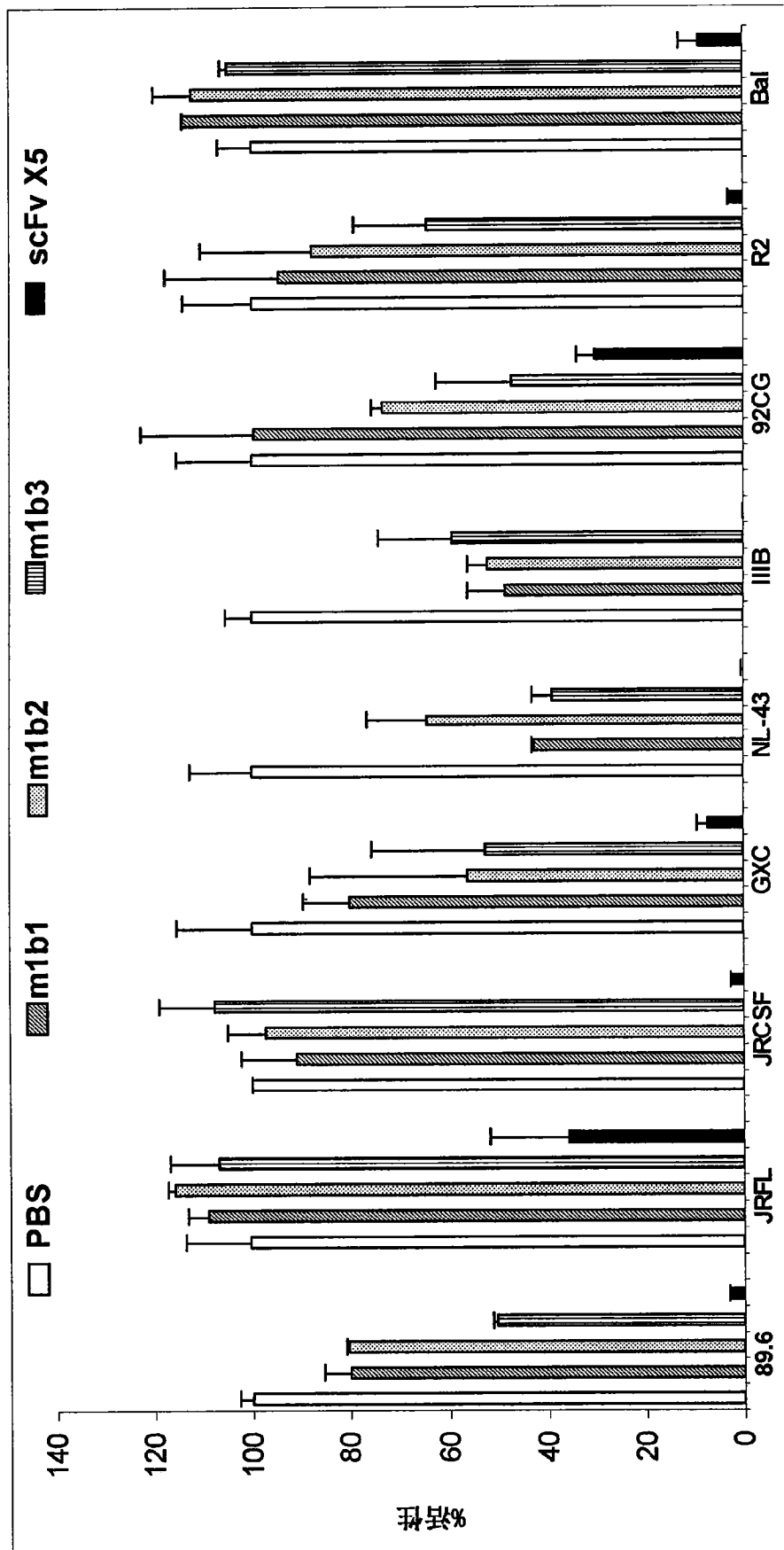


图 17

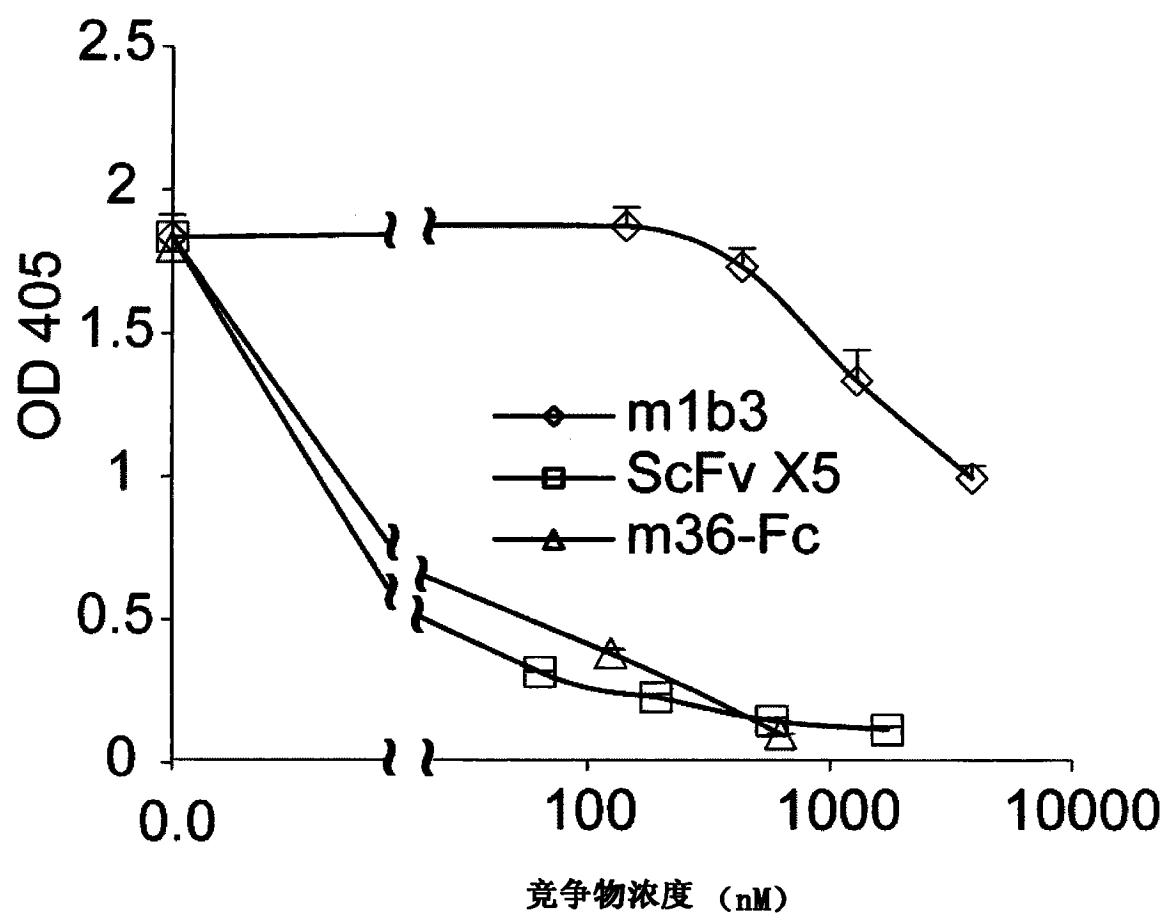


图 18