



(12) 发明专利

(10) 授权公告号 CN 111448314 B

(45) 授权公告日 2024.10.18

(21) 申请号 201880079095.7

(72) 发明人 J·麦考夫迪 R·佩雷拉

(22) 申请日 2018.12.05

M·R·戴森 K·帕特赫本  
J·L·锡尔加宁

(65) 同一申请的已公布的文献号

申请公布号 CN 111448314 A

(74) 专利代理机构 中科专利商标代理有限责任  
公司 11021

(43) 申请公布日 2020.07.24

专利代理人 张莹

(30) 优先权数据

1720351.4 2017.12.06 GB

(51) Int.CI.

C12N 15/10 (2006.01)

(85) PCT国际申请进入国家阶段日

2020.06.05

(56) 对比文件

US 6300065 B1, 2001.10.09

(86) PCT国际申请的申请数据

PCT/EP2018/083698 2018.12.05

CN 106574264 A, 2017.04.19

审查员 韩沛

(87) PCT国际申请的公布数据

W02019/110691 EN 2019.06.13

权利要求书2页 说明书134页

序列表137页 附图92页

(73) 专利权人 艾恩塔斯有限公司

地址 英国剑桥郡

(54) 发明名称

根据可开发性选择真核细胞展示系统中的多肽药物

(57) 摘要

一种结合剂(例如抗体、受体)在体外培养的高等真核细胞上的表面呈现水平作为所述结合剂的可开发性特征(例如溶解度)的预测指标的用途。高等真核细胞(例如哺乳动物细胞)的展示文库适用于根据可开发性和靶结合亲和力筛选表面展示的结合剂。根据可开发性(包括结合剂溶解度、在高浓度下配制的能力、非特异性结合倾向低和半衰期)进行的内置选择实现了展示文库的高通量筛选。利用可操作地连接到编码结合剂的DNA的诱导型启动子控制结合剂的细胞表面呈现,实现了具有可开发性特征和/或其它质量(例如靶结合和亲和力)的结合剂群体的富集。

1. 一种根据结合剂在溶液中的溶解度和/或自缔合抗性将它们区分或排序和/或富集在溶液中展现较大溶解度和/或较大自缔合抗性的结合剂的方法, 其包括

(i) 提供各含有编码结合剂的DNA的哺乳动物细胞克隆的文库,

(ii) 在用于表达所述结合剂的条件下体外培养所述克隆, 其中所述结合剂呈现于细胞表面上,

(iii) 测定所述结合剂在所述克隆上的表面呈现水平, 任选地通过用包含可检测标记的试剂标记所述结合剂来测定,

(iv) 选出相较于其它克隆展现结合剂存在更高的表面呈现的一种或多种克隆, 以及

(v) 鉴定出由所述一种或多种所选克隆编码的在溶液中具有良好溶解度和/或自缔合抗性的结合剂, 且任选地提供所选克隆用于一个或多个进一步筛选步骤,

其中所述结合剂包含跨膜结构域并且其中所述结合剂是抗体。

2. 根据权利要求1所述的方法, 包括通过用包含可检测标记的试剂标记所述结合剂来测定所述结合剂在所述克隆上的表面呈现水平, 其中所述试剂结合到所述结合剂的恒定区, 任选地其中所述结合剂包括Fc区且所述试剂结合到所述Fc区。

3. 根据权利要求1或2所述的方法, 其中所述可检测标记为荧光标记。

4. 根据权利要求1所述的方法, 包括根据结合剂在细胞上的表面呈现水平将细胞分选到收集部分和丢弃部分中, 其中将展示表面呈现大于预定阈值的细胞分选到所述收集部分中并且将展示表面呈现低于所述预定阈值的细胞分选到所述丢弃部分中。

5. 根据权利要求4所述的方法, 其中所述丢弃部分包括表达临界浓度为至少10mg/ml的对比多肽的细胞, 且其中所述收集部分包括表达临界浓度为所述丢弃部分中的所述对比多肽的至少1.5倍高的结合剂的细胞。

6. 根据权利要求2所述的方法, 其中分选是利用荧光活化细胞分选仪(FACS)进行。

7. 根据权利要求1所述的方法, 其中步骤(ii)包括将所述文库中的所述克隆作为混合物在一个容器中培养。

8. 根据权利要求1所述的方法, 其中步骤(ii)包括在独立容器中培养所述文库中的每种克隆。

9. 根据权利要求1所述的方法, 其中所述结合剂是亲本结合剂的序列变异数。

10. 根据权利要求9所述的方法, 其中所述亲本结合剂已鉴定为需要改善其在溶液中的溶解度或自缔合抗性。

11. 根据权利要求9所述的方法, 其中所述方法包括产生所述亲本结合剂的序列变异数以及将编码所述序列变异数的DNA整合到哺乳动物细胞的细胞DNA中以提供含有编码所述结合剂的DNA的细胞克隆文库,

任选地, 其中所述方法包括分析所述亲本的多肽序列、鉴定出经预测会促进自缔合和/或减小溶解度的一个或多个氨基酸残基, 以及使所述一个或更多个氨基酸残基产生突变。

12. 根据权利要求9所述的方法, 其中所述亲本结合剂在磷酸盐缓冲盐水溶液(PBS)中的临界浓度小于50mg/ml且/或在磷酸盐缓冲盐水溶液(PBS)中的溶解度极限小于50mg/ml,

且/或其中所述方法包括鉴定出临界浓度和/或溶解度极限为所述亲本结合剂的至少1.5倍高的由所述一种或多种所选克隆编码的结合剂。

13. 根据权利要求1所述的方法, 包括基于结合剂在所述细胞克隆上的表面呈现水平来

预测所述结合剂的亲水性和/或鉴定出具有较强亲水性的一种或多种所选克隆的结合剂。

14. 根据权利要求1所述的方法,其中编码所述结合剂的所述DNA是在强启动子的控制下表达。

15. 根据权利要求14所述的方法,其中所述启动子是组成型启动子。

16. 根据权利要求15所述的方法,其中所述启动子是CMV启动子。

17. 根据权利要求14所述的方法,其中所述启动子是诱导型启动子。

18. 根据权利要求1所述的方法,其包括

(i) 同时测定所述结合剂的表面呈现水平和所述结合剂结合靶标的水平,以及协同选择展示同源结合剂的克隆,所述同源结合剂展现的表面呈现高于其它克隆;或

(ii) 同时测定所述结合剂的表面呈现水平和非特异性结合到非靶分子的水平,以及协同选择展示结合剂的克隆,与其它克隆相比,所述结合剂展现更高的表面呈现和更低的非特异性结合;或

(iii) 同时测定所述结合剂的靶标结合水平和非特异性结合到非靶分子的水平,以及协同选择展示同源结合剂的克隆,所述同源结合剂展现的非特异性结合低于其它克隆。

19. 根据权利要求1所述的方法,其包括

测定编码所述结合剂的所述DNA的序列,所述结合剂来自所述一种或多种所选克隆,以及

提供编码所述结合剂的经分离的核酸。

20. 根据权利要求1所述的方法,其进一步包括

测定编码所述结合剂的所述DNA的序列,所述结合剂来自所述一种或多种所选克隆,以及

在体外,在以可溶形式分泌所述结合剂的条件下,在宿主细胞中表达编码所述结合剂的DNA。

21. 根据权利要求20所述的方法,其中分泌的所述结合剂是以至少1mg/ml的产量获得。

22. 根据权利要求20所述的方法,其进一步包括提纯和/或浓缩所述结合剂以获得浓度为至少10mg/ml的所述结合剂的水溶液。

23. 根据权利要求22所述的方法,其中所述浓度是至少50mg/ml。

24. 根据权利要求23所述的方法,其中所述浓度是至少100mg/ml。

25. 根据权利要求20所述的方法,包括将所述结合剂配制成包括医药学上可接受的赋形剂的组合物。

26. 根据权利要求25所述的方法,包括将所述组合物提供于注射用的预装药注射器中,所述预装药注射器任选地存在于试剂盒内,所述试剂盒包括一种或多种额外的组件。

27. 根据权利要求26所述的方法,其中所述一种或多种额外的组件是针和/或包括通过注射施用所述组合物的说明的产品信息说明书。

## 根据可开发性选择真核细胞展示系统中的多肽药物

### 技术领域

[0001] 本发明涉及候选多肽药物的鉴定,所述候选多肽药物具有所期望的可开发性特征,例如溶解度、在高浓度下配制的能力、非特异性结合倾向低,和最优半衰期。本发明进一步涉及通过药物发现(包括抗体发现)来筛选结合剂。

### 背景技术

[0002] 抗体已证实是一类极其成功的药物,迄今为止批准的治疗抗体逾70种,而处于开发流水线中的则更多。然而,在许多方面中,多肽药物(例如抗体)的生产和配制比小分子医药更复杂费劲。许多因素影响开发多肽药物(例如抗体)的可行性,影响着主要候选抗体是否将成功地开发成不仅有效、而且可制造、稳定且安全的药物。此类因素包括化学稳定性(抵抗例如断裂、脱酰胺、氧化和异构化)、物理稳定性(例如构象稳定性、蛋白质展开、聚集和/或沉淀的倾向、胶体稳定性),以及溶液特性(例如溶解度、在溶液中可逆地自缔合的倾向,和高浓度下的粘度)。多肽在细胞培养时以高产量表达的能力也是一个相关考虑因素,因为宿主细胞所分泌的多肽的数量决定了能从培养基中回收的产物的产量。抗体或多肽药物的免疫原性也是一个考虑因素。所有这些因素可以统称为产物的“可开发性”特征。这些因素因为它们影响产品成本和制造可行性、安全特征、给药时程和施药方式而成为重要的考虑因素。在遇到可开发性问题的情况下,它们会影响抗体的效用或商业成功,并且可能导致主要分子在开发过程中失败。抗体药物可开发性和其测量方法的各方面已进行评审<sup>1-3</sup>。

[0003] 希望生产出在高浓度下稳定、可溶的蛋白质以便施用于患者。举例来说,皮下施药寻求50mg/ml以及更大的抗体浓度,其中必须以低体积施以相对较大的剂量。用于治疗用途的理想多肽在缓冲水溶液中具有高溶解度并且因此能够高水平浓缩。成功地生产高浓度的可溶蛋白质至少部分地取决于多肽对自缔合的抗性,否则自缔合会引起粘度增加和/或沉淀。当达到分子的溶解度极限时,所溶浓度不能进一步增加,并且出现非期望的效应,例如分子从溶液中沉淀。当接近溶解度极限时,溶液可能变得粘稠且/或所溶多肽可能在溶液中自缔合(可逆地或以其它方式),此类效应使溶液的操作和配制复杂化。更糟的是,当产品施用于患者时,难溶性和/或不稳定多肽形成聚集体会使免疫原性的风险加强<sup>4,5</sup>。抗药物免疫反应(例如抗药物抗体)可能中和治疗作用,并且超敏反应能导致发病或死亡。

[0004] 多肽高水平表达的能力会影响生产的经济性。在抗体发现的早先步骤期间,通过在哺乳动物细胞培养时优化短暂表达所实现的产量能达到50μg/ml或更大。瞬时表达是一种短暂表达系统,其中将编码所关注产物的质体转染至细胞中并且表达一段时间,通常在2-7天之后因细胞中的质体损耗而下降。为了获得稳定表达型重组细胞用于药物制造,必须将编码DNA稳定地整合于细胞基因组中。建立稳定制造型细胞系之后,可以培养细胞一或两周以获得较高产量,其可能是约1mg/ml或更高。

[0005] 瞬时表达的产量低可能表示候选药物存在潜在的可开发性问题。举例来说,倾向于聚集的抗血管生成素抗体(“Ang2”)据报道只有10μg/ml的产量,而优化的抗体变异体产量为260μg/ml<sup>6</sup>。然而,即使当瞬时表达(例如>50μg/ml)和/或稳定细胞系(例如>1mg/ml)获

得良好产量时,可能出现生物物理学问题,此时多肽浓度大于 $1\text{mg}/\text{ml}$ 。Dobson等人(2016<sup>7</sup>)发现抗NGF抗体MEDI1912与其亲本抗体MEDI576相比,存在显著的生物物理学问题,但这在瞬时培养所产生的表达水平上得不到反映,其据报道为约 $200\mu\text{g}/\text{ml}$ 。细胞培养表达的产量因此只是可开发性的一个方面并且可能预测不了其它重要方面。

[0006] 在过去的几十年间,展示技术已经允许建立抗体和其它蛋白质和肽的多种多样的大型群体,从而能够自其中分离出具有所期望靶标结合特性的个别变异体。所期望的结合特性包括以适当特异性(例如针对直系同源物和旁系同源物)且以适当亲和力结合到靶标上的适当抗原决定基(例如抑制受体-配体相互作用)。展示技术通过允许一些此类方面在选择期间富集而能够有助于找到具有所期望特性的结合剂。举例来说,通过使用有限量的抗原驱动选择能够选出较高亲和力抗体。针对结合到非所需旁系同源物进行取消选择也已有描述。体外结合剂展示平台和其根据期望特性(包括可开发性的一些方面)进行选择的用途已进行评审<sup>8</sup>。还已知体内产生抗体的方法,包括实验动物(例如小鼠)的免疫。通过使用动物免疫系统建立一组抗体已发现多种已批准的抗体药物分子,然后在多孔板中或通过流式分选术、根据期望特性(例如结合肿瘤相关抗原;细胞因子中和)加以筛选。

[0007] 不能保证从“天然”来源分离的抗体基因(不论从动物直接获得或内置于组合型展示系统,例如酵母展示或噬菌体展示)将编码展现良好可开发性特征的抗体。这适用于从初始来源获得的抗体基因以及免疫之后活体内产生的那些抗体基因,假定此类抗体将具有良好可开发性特性的先验理由不存在。免疫系统的目标是建立高亲和力抗体,而非建立委身于工业制造或配制成医药产品的抗体。体内或体外产生的抗体不一定能够浓缩到适于医药配制的水平。人类血清中的IgG平均总浓度是 $12\text{mg}/\text{ml}$ ,其中任何个别抗体均以显著较低的浓度表示。个别抗体的比例也存在很大变化,在免疫反应期间,在个别B细胞之间观察到1000倍的表达差异<sup>6</sup>。因此,不论抗体起源(例如免疫的动物/人类供体、合成或半合成来源),无法假定其在高浓度下的溶解度。

[0008] 传统上,药物发现的早期聚焦于鉴定具有所期望的治疗作用方式(例如靶标结合特性)的分子,而对可开发性的评估则推迟到后续阶段,通常在主要分子或一组有限的分子已经选中用于临床前开发之后。由此带来的不幸结果是,可开发性问题可能只在药物开发的相对晚期阶段才显露出来,此时已经耗费了大量的时间和金钱。此类失败代价昂贵。虽然生物医药产业认识到候选药物的失败是药物发现的普遍特点(大部分潜在药物从未到达临床),但是为了减少浪费和使资源能够向更稀有的成功分流,人们希望药物“尽早失败”。

[0009] 已经证明,聚焦于改善亲和力的工程化抗体能够产生亲和力增强的变异体,而其突变不利地影响生物物理学特性,例如倾向于自缔合。抗NGF抗体MEDI1912是其实例。来自亲本抗体MEDI578的已亲和力成熟的MEDI1912据报道对NGF具有pM亲和力<sup>7</sup>。然而,与MEDI578相比,亲和力成熟的MEDI1912抗体展现不良溶解度、胶体不稳定性、低浓度下的聚集和短半衰期。

[0010] 已努力将根据可开发性的一些方面进行的选择或筛选整合到体外选择平台,例如噬菌体展示。对于仅具有VH或VL域的单域抗体或dAb来说,在针对抗原进行选择之前,热量的挑战能够缩小解链温度( $T_m$ )较低的抗体的比例。 $T_m$ 的升高已与生物物理学特征的整体改善有关,例如可逆的展开、抗聚集、溶解度、表达和细菌的提纯产量。然而, $T_m$ 类似的分子能够展现生物物理学特性的差异-参见Dobson等人2016<sup>7</sup>和实例3。scFv分子也已经工程改造

CN 111448314 B 而具有改善的稳定性,以便纳入双特异性平台<sup>9-11</sup>。另外,已经构建噬菌体展示文库(例如Tiller等人(2013)<sup>12</sup>,其试图将多样性引入文库,同时避免纳入任何潜在的翻译后修饰位点(例如脱酰胺位点、异构化位点、蛋白酶裂解位点和氧化位点)。

[0011] 已经创建多种算法用于计算机模拟预测分子行为,此类算法有助于对相对较大数目个潜在候选药物分子的可开发性特征进行比较。这种类型的算法能够通过鉴定可能负责的氨基酸序列的特点而有助于指出可开发性问题背后的可能原因。举例来说,应认识到,序列中的疏水性补丁能够引起分子“粘性”,其展现为非特异性结合和/或自聚集,由于多肽紧密缔合,因此在高浓度下,这更可能发生。蛋白质展开可能使以原生构象内埋于分子内的疏水性残基暴露。

[0012] 在可开发性问题产生的情况下(不论预测到的或只是在开发过程中遇到的),可以尝试利用蛋白质工程解决困难。Dudgeon等人<sup>13</sup>在抗体VH和VL域中鉴定出特定的位置供引入天冬氨酸或谷氨酸残基来改善生物物理学特性。所得抗体据称无聚集、表达充分和能加热再折叠。据报道,所鉴定的突变通过独立于抗体序列的其余部分改变特定位置的局部电荷分布来增强抗聚集性,并且因此充当通用模板用于工程改造人类抗体可变域而具有优良的生物物理学特性。在其它情况下,已经检查出特定的抗体序列存在可能约束可开发性的个别特点。举例来说,从B细胞融合瘤活动中鉴定出的抗血管生成素抗体“Ang2”在瞬时培养时发现具有低水平的表达并且倾向于聚集<sup>6</sup>。其序列分析鉴定出在轻链可变域构架的位置49处存在不成对的半胱氨酸。制备20种个别变异数体,评估聚集情况且鉴定溶解度改善的版本,其中半胱氨酸49变为苏氨酸(C49T)。在这种情况下,亲本克隆的基线表达极低,并且还观察到C49T变异数体的产量改善而结合未受损。

[0013] 另一实例是上述亲和力改善的MEDI1912抗体的“修复”,其停滞于开发中<sup>7</sup>。通过将氢:氘交换与结构建模组合,鉴定出VH域中存在3个非互补位疏水性残基并且将这些残基恢复为亲本抗体中所发现的残基。使位置30处的色氨酸、位置31处的苯丙氨酸和位置56处的亮氨酸分别转化成丝氨酸、苏氨酸和苏氨酸(用W30S、F31T和L56T表示,其中编号表示抗体链内的氨基酸位置,第一个字母表示原始氨基酸且最后一个字母表示替换的氨基酸)。可开发性改善的版本(称为MEDI1912-STT)显示聚集减少并且在许多其它方面有改善,包括非特异性结合减少和半衰期延长。

[0014] Bethea等人(2012)<sup>14</sup>描述了抗IL-13抗体具有不良的生物物理学特性,包括在13mg/ml下自聚集,引起沉淀。重链CDR3中的芳族三联体(由苯丙氨酸、组氨酸和色氨酸组成)经鉴定是潜在的问题。作者描述了引入若干突变,包括改善溶解度和减少非特异性相互作用的单氨基酸变化(在位置100a处用丙氨酸取代色氨酸残基)。在这种情况下,所述变化也减少靶标结合。

[0015] MEDI1912-STT和抗IL-13抗体的实例(其中序列优化能够同时减少非特异性结合和其它非期望的行为,例如自聚集)指出可以将多种可开发性参数互连且多种可开发性参数可以具有共同的潜在原因。然而,虽发现非特异性相互作用但自相互作用证据极少的其它情况已有报道<sup>2,15</sup>。

[0016] 非特异性相互作用由于其能不利地影响药物分子的性能、特异性、体内分布或半衰期而成为药物开发时的重要考虑因素。已经证明,在具有相同Fc区的抗体之间,抗体的体内半衰期能显著不同<sup>16</sup>,表明抗体可变域影响半衰期。在许多情况下,这已归因于非特异性

相互作用。如果施用的药物被吸入与非靶标组分发生的非特异性结合相互作用“漏槽”,则可供结合到其靶分子的利用度更小并且到达或渗透病变靶组织或部位的能力可能减小。甚至低亲和力非特异性相互作用可能是显著的,尤其在靶标丰裕度较高时。举例来说,循环中的抗体暴露于大面积的内皮细胞糖萼(估算面积为 $350\text{m}^2$ ),据报道到达 $0.5\text{mm}$ 或更大的深度。带负电的糖萼是由不同的葡糖胺聚糖构成,例如玻糖醛酸和蛋白多糖,例如肝素硫酸盐,其共同构成主要组分。其也呈现所吸收的血浆蛋白质<sup>17,18</sup>。此基质和所呈现的其它表面对于循环中的抗体的低亲和力结合可能对药物动力学具有显著影响。

[0017] 存在根据非特异性相互作用进行筛选的方法,并且在所关注的个别结合剂已经鉴定之后,通常基于逐个克隆来执行这些方法。Hotzel等人(2012)描述了一种帮助鉴定展现非特异性相互作用的抗体的策略,所述策略是在ELISA中根据对杆状病毒颗粒的结合进行筛选<sup>16</sup>。此外,Xu等人描述了一种多特异性试剂结合分析(PSR MFI),其利用标记的蛋白质混合物与酵母展示平台鉴定展现低特异性的抗体<sup>19</sup>。也已开发出许多色谱方法来帮助鉴定倾向于非特异性相互作用的抗体。这些方法典型地涉及使所关注的相互作用搭配物、化合物或表面固定于基质(例如琼脂糖)上以及然后使测试抗体分子越过基质以及通过固定或改变洗涤条件,比较测试抗体和对照抗体的滞留程度。已经使用例如交叉相互作用色谱(CIC)<sup>15</sup>、疏水性色谱和肝素等方法。

[0018] 关于可开发性的另一个考虑因素是治疗蛋白的体内半衰期。为了使全身性施用的许多药物达成最优功效,必需使其血清浓度在特定水平(或在目标范围内)维持一定的时间段。上文所论述的非特异性结合相互作用对体内多肽半衰期和/或有效浓度能产生一定程度的影响。另外,对于抗体和含有Fc区域的其它分子来说,Fc区对FcRn(内皮细胞中表达的受体)的结合能够强烈地影响半衰期。人类IgG与其它循环分子相比具有相对较长的半衰期,且这归因于其与FcRn的pH依赖性相互作用。胞饮和核内体运输之后,抗体Fc与FcRn受体之间发生了相对较高的亲和力相互作用,从而保护抗体以免被溶酶体降解。在中性pH下,Fc与FcRn之间的亲和力较低且在再循环到细胞表面之后,遇到中性pH。在这些条件下,抗体被释放回到循环中。因此,对FcRn受体的pH依赖性结合是IgG分子的重要特性。FcRn与大鼠IgG2a的复合物的晶体结构<sup>20</sup>显示FcRn结合到CH2和CH3域(在CH 2残基252-254和309-311,和CH 3残基434-436)且有助于解释重要的pH依赖性结合。半衰期改善的Fc变异体已经通过Fc工程改造产生。

[0019] 铃木(Suzuki)等人(2010)已证明中性pH下的低亲和力与体内长半衰期之间存在着正相关<sup>21</sup>,但显然,其它因素也可以促成。抗IL-12抗体布瑞金单抗(briakinumab)和优特金单抗(ustekinumab)尽管具有类似的Fc域,但分别具有8天和22天的半衰期。利用这些抗体和一系列杂交变异体,Schoch等人(2015)<sup>22</sup>证明中性pH下保持的结合与体内半衰期之间良好相关。他们指出布瑞金单抗VL上的带正电大补丁引起与FcRn的静电相互作用增强,从而限制了在胞外pH下发生的FcRn-IgG解离。Kelly等人(2016)描述了类似的pH依赖性相互作用,但他们主张与其它蛋白质的非特异性相互作用也可能导致短半衰期<sup>23</sup>。

[0020] 通常,药物发现中利用展示文库的主要目的是使文库富集表达结合剂的克隆,所述结合剂对于结合到所关注的靶分子具有高亲和力。噬菌体展示文库能够用于相对于非结合剂富集结合剂以及相对于较低亲和力克隆富集较高亲和力克隆。基于亲和力的相对富集已允许利用噬菌体展示进行抗体的亲和力成熟。典型地,生物素化抗原连同相关展示封装

(例如使用抗生蛋白链菌素涂布的磁珠)一起使用可容许回收抗体:抗原复合物。抗原浓度较低时,群体内的较高亲和力抗体比较低亲和力抗体形成复合物的可能性更大,导致这些抗体相对富集。

[0021] 然而,为了结合剂在真核细胞上展示,使用限制抗原联合珠粒固相回收的方案的有效性比富集高亲和力结合剂的噬菌体展示更小。不同于单价噬菌体展示,真核细胞是多价展示封装,其表面上存在相同结合剂的许多拷贝。细胞群内所呈现的结合剂在选择期间的浓度潜在地相对高于设法求解的抗原浓度和/或亲和力。这可能会遮蔽亲和效应并且难以对混合群体中的不同克隆上所展示的结合剂的亲和力的差异进行求解,因此与例如噬菌体展示等方案相比,使得所达成的富集系数减小。

[0022] Boder和Wittrup<sup>24</sup>描述了一种声称可提高严格度且容许在真核细胞展示文库中根据亲和力进行选择的方法,所述方法使用有限浓度的荧光标记单价抗原联合流式细胞术来根据亲和力进行选择,测量每个细胞所结合的抗体的数量以控制表达的差异。在这个系统中,结合的严格度据报道是通过减小靶标的浓度(相对于结合剂)来提高,其中据称是基于从靶标检测到的信号水平来检测较高亲和力结合剂,原因是结合到较高亲和力结合剂的靶分子比结合到较低亲和力结合剂的靶分子多。

[0023] 许多出版物已经描述了其中靶标结合和抗体表达同时予以检查的选择。以此方式能够基于所达成的展示水平将抗原结合程度归一化。US8,771,960 (DKFZ) 描述了使用荧光活化细胞分选(FACS)方法选择较高亲和力单克隆抗体,其中抗体文库或一组不同抗原特异性融合瘤细胞用PE标记的抗原进行染色且用FITC标记的蛋白质G进行对比染色。在FACS中选出PE染色:FITC染色比值最大的细胞,随后扩增各产生抗体的个别细胞,所述抗体对用于选择的靶抗原具有相对较高的亲和力。抗体结合的抗原与抗体未结合的抗原的比率因此用作抗体对其抗原的亲和力的直接度量。

[0024] Chao等人(2006)<sup>25</sup>描述了编码scFv谱系的基因,所述scFv以基因方式与酵母凝集素Aga2p亚单元融合。在Aga2p酵母展示系统中,所关注的结合剂(在此为scFv)与Aga2p亚单元融合,所述亚单元经由二硫键连接至存在于酵母细胞壁中的Aga1p亚单元。表达靶标特异性结合分子的酵母细胞能够通过流式细胞术、使用直接或间接标记的靶分子鉴定。举例来说,可以将生物素化靶标添加至细胞中且可以利用抗生蛋白链菌素-藻红蛋白检测对细胞壁内所呈现的scFv的结合。靶分子的有限浓度使得富集表达较高亲和力结合剂的克隆成为可能,原因是那些克隆捕捉的靶分子更多并且因此展现更亮的荧光。为了控制不同细胞中的scFv表达变化,Chao等人(2006)<sup>25</sup>使用荧光标记的抗标签抗体测量每个细胞表面上的抗体表现量,从而允许针对表达水平的变化进行归一化。这种方案因此允许展示高亲和力结合分子的酵母细胞是由高水平表达较低亲和力抗体的那些细胞分化而成。在这种情况下,测量抗体展示的目的因此是针对表达差异进行归一化且借此促进亲和力选择。

[0025] 其中使用展示水平作为分泌形式的相同蛋白质的潜在表达产量的指标的方法也已有描述。为了试图在建立稳定细胞系以便生产抗体期间鉴定出高度表达的细胞克隆,W02015128509(Glenmark Pharmaceuticals)描述了一种方案,其中抗体以分泌形式表达,但对呈现于细胞表面上的比例进行“取样”。此取样是作为剪接事件绕过第一终止密码子的结果出现,并且在抗体的一部分中,mRNA使抗体基因剪接到编码跨膜域的外显子上。在此系统中,抗体的展示水平据报道与所表达的可溶性抗体的数量直接相关。

[0026] 使用酵母进行的许多研究已经报道结合剂在酵母细胞上的表面呈现水平、结合剂的热稳定性和/或通过细胞培养产生的结合剂的产量之间有关。Shusta等人<sup>26</sup>将可溶性单链T细胞受体(scTCR)变异数与Aga2p融合且报道不同突变体的热稳定性与其可溶性分泌水平和以与Aga2p的融合体形式在酵母细胞壁上展示的scTCR的数量强烈相关。他们提出内质网(ER)的质量控制设备使热力学不稳定突变体发生的细胞内蛋白质水解决定了蛋白质表达效率。Kowalski等人<sup>27,28</sup>检查了可溶性多肽III型纤维结合蛋白(FnIII)域的变异数在酵母中的可溶性表达并且还报道了多肽的热力学稳定性与其分泌效率(因此与产量)之间相关。Hackel等人<sup>29</sup>使用酵母展示系统进一步研究了FnIII域的热力学稳定性的作用。Hackel等人培养了表达一系列与Aga2p融合的FnIII变异数的酵母,所述变异数经突变以减少其热稳定性,并且发现热稳定性与表面拷贝数之间正相关。因此,较多的热不稳定性突变体展现降低的表面展示水平。

[0027] WO2012/158739描述了一种用于从基于域的文库中选择多肽的两阶段方法,所述方法涉及从体外核糖体展示文库中进行基于抗原的选择,随后将所选结合剂转变为酵母展示文库用于进一步的基于抗原的选择。体外核糖体展示系统产生的多肽大部分聚集,而酵母展示选择的纳入使高度聚集的多肽在所选群体中的数目减少,但产生单体的克隆的分率仍较低。新出现的克隆的表达行为的适度改善归因于酵母系统不大容许误折叠(例如热不稳定)蛋白质的表达,使得可供酵母文库选择利用的此类蛋白质更少。此外,这项工作涉及Aga1p:Aga2p-FnIII展示系统。

[0028] 另一方面,Julian等人<sup>30</sup>发现根据抗原结合和VH域在酵母上的的表面展示进行的协同选择产生亲和力较高、但稳定性较低的抗体。最高亲和力VH域是高度不稳定的。此研究发现在热稳定性的整个范围内,酵母上的展示水平只发生适度的变化(1.6倍),并且作者报道了展示水平不能引导选出使亲和力和稳定性共同改善的突变集合。

[0029] 抗体亲和力、特异性、稳定性和溶解度之间关系的最新评论描述了一种特性(例如亲和力)的改善如何能导致其它特性(例如稳定性)的不足以及如何能平衡这些得失以协同优化抗体的多种特性<sup>31</sup>。

[0030] 因此,在多个前沿做出的广泛工作已寻求鉴定和理解影响可开发性的多肽的不同特征之间的潜在关系。尽管如此,但缺乏将可开发性筛选方便地整合到候选药物选择早期的多肽药物发现方法,尤其是例如药物溶解度和避免非特异性结合等方面。

## 发明内容

[0031] 本发明提供在从展示文库发现阶段期间有助于检测多肽的可开发性问题的方法和产品,从而能够避免分子出现可开发性的可靠性问题并且能够从候选药物分子池中富集具有更佳可开发性特征的那些分子。

[0032] 本发明人已惊人地发现,多肽呈现于真核宿主细胞表面上的水平与多肽的特定可开发性特征有关,包括其在溶液中的特性,例如其在溶液中的溶解度和其对自缔合的抗性。可以使用从重组基因表达多肽且将多肽展示于其表面上的宿主细胞、通过测定多肽在细胞表面上的呈现水平来根据此类可开发性特征评估或筛选多肽。这适用于高通量并行筛选多种宿主细胞克隆,从而允许对相对的表面呈现进行比较和选出展示具有更佳可开发性特征的多肽的克隆。因此,多肽的表面呈现水平代表着多肽结合剂筛选方法(例如抗体发现)的

可开发性预测指标。另外,表面呈现水平与生物物理学特性(例如自缔合)之间的这种关联允许从结合剂文库中选出或在此类文库内富集具有此类最优生物物理学特性的结合剂。反之,可以根据结合剂文库选出或从结合剂文库中排除生物物理学特性较差(例如溶解度较低)的结合剂。

[0033] 本发明人还已经设计出根据与多肽药物体内特性(例如在血清中或在靶标器官和组织中的非特异性结合、半衰期和有效浓度)有关的可开发性方面在体外筛选高等真核细胞所表达的多肽的方法。

[0034] 根据本发明的方法和用途在早期阶段筛选期间具有特定的优势,包括筛选结合剂(例如抗体)的大型多样化文库。通过在药物发现的最早阶段整合候选多肽药物的可开发性筛选,本发明使代价高的晚期失败风险降低。根据本发明的可开发性筛选还可以用于对候选多肽药物的较晚期池(任选地,享有共同谱系的抗体“家族”)进行选择,以从所述池中富集具有更佳可开发性的多肽且/或告知关于选择主要分子用于药物开发的决定。另外,本发明的技术可以用于鉴定现有候选多肽药物(例如尚未满足一个或多个可开发性标准或期望改善一个或多个可开发性特征的候选药物)的改善变异体。因此,本文描述了产生和快速筛选衍生物序列的方法,所述衍生物序列展现改善的可开发性特征。

[0035] 哺乳动物细胞中的多肽表达路径始于内质网上的翻译机器(核糖体等),随后新生多肽经由高尔基体复合体行进且输送到质膜,在质膜中,他们可以被细胞分泌或滞留于细胞表面上(例如作为膜蛋白)。当重组产生的多肽被分泌时,其立即在大体积的培养基中稀释且在若干天/周积聚之后,以典型地在1-100 $\mu\text{g}/\text{ml}$ 之间的浓度存在。相较于多肽药物在被配制成施用于患者的药剂中的期望浓度,这是较低的。与分泌的多肽相比,滞留于细胞表面上的所表达多肽能够在细胞表面上形成高局部浓度。所表达多肽在细胞表面上的滞留大量减少可供多肽使用的体积且因此提供实现高浓度的机会。当从强启动子(例如细胞巨大病毒(CMV)启动子)表达所展示的多肽时,浓度可能特别高。因此,多肽结合剂在重组细胞文库中的哺乳动物细胞和其它真核细胞表面上的滞留使得结合剂在质膜表面上以局部高密度浓缩,尤其当使用强烈表达重组基因的宿主细胞时,其中所编码的多肽代表总体多肽合成中的大部分。应用于表达多肽谱系的一组克隆时,不同多肽的表面呈现水平在克隆间的变化可以反映多肽特征,例如其对自缔合的抗性。自缔合倾向较低的多肽在细胞表面上能浓缩到较高水平,其当紧密接近时能够抵抗聚集有助于此。培养的真核细胞展示文库因此可以给展现良好可开发性特征的结合剂充当体外选择环境。这由本文提出的实例中的证据证明。

[0036] 根据本发明的第一方面,多肽结合剂(例如抗体)在所培养真核细胞克隆表面上的表面呈现水平用作结合剂的可开发性特征的预测指标,例如其在溶液中的溶解度、对自缔合的抗性和/或在溶液中浓缩的能力。不希望受理论束缚,与结合剂溶解度有关的一种要素可以是其亲水性,其中较大亲水性(较低疏水性)与较高溶解度有关;在水溶液中对自缔合的较大抗性;和在溶液中达到较高浓度的能力。因此,本发明的方法可以用于基于在如本文所述的真核细胞展示系统中的表面呈现程度将较大亲水性结合剂(候选多肽药物)与较小亲水性结合剂区分开来。

[0037] 本发明提供一种方法,其包括

[0038] 提供各含有编码结合剂的DNA的高等真核(例如哺乳动物)细胞克隆的文库,

[0039] 在用于表达所述结合剂的条件下体外培养所述克隆,其中所述结合剂呈现于细胞表面上,

[0040] 测定所述结合剂在多种克隆上的表面呈现水平,

[0041] 选出相较于其它克隆展现结合剂存在更高的表面呈现的一种或多种克隆,以及

[0042] 鉴定出具有良好可开发性特征的由一种或多种所选克隆编码的结合剂。

[0043] 本发明可以用于根据结合剂可开发性特征区分结合剂或将结合剂排序和/或选出具有良好可开发性特征的一种或多种结合剂。此类方法评估的可开发性特征可以是结合剂的溶液特性,例如溶解度、对自缔合的抗性,和/或在水溶液中浓缩的能力,如本文在别处详细论述。选出展现较高表面呈现的克隆得到所选细胞群,所选细胞群富集展现结合剂存在较高表面呈现的克隆,其然后能够任选地用于一种或多种其它方法中,例如额外多轮筛选。

[0044] 测定表面呈现水平的方法在本文中别处详细地描述且任选地包括用包含可检测(例如荧光)标记的试剂标记结合剂。抗体和其它结合剂包括Fc区时,用结合Fc区的试剂标记是很方便的,例如检测剂可以是经标记的抗IgG抗体。方法可以包括测定表面呈现水平和观察所述文库中的细胞上的结合剂呈现水平范围。拷贝数范围的实例见于本文中别处。

[0045] 本发明的另一个方面涉及在从展示文库中发现结合剂期间评估非特异性结合。在初始结合剂发现期间,鉴定与非靶分子的相互作用是有利的。本发明的方法可以用于筛选高等真核细胞上所展示的具有最优生物物理学特性且与非靶分子的体内非特异性相互作用倾向低的结合剂(例如抗体)群体。虽然常规做法是选择令人期望地结合到靶分子(例如多肽药物的分子靶标,多肽药物在体内结合其且对其发挥生物效应)的候选多肽药物,但是如果还注意对非期望结合到非靶分子(例如结合剂非特异性结合的组分或一类分子,例如结合到带负电的聚合物,如核酸)的候选多肽药物进行负向选择,则可以减少可开发性的问题。在选择结合剂的方法中,可以用非靶分子取代靶分子,其例外之处在于然后丢弃而非保留识别非靶分子的结合剂,借此富集不识别非靶分子的结合剂。

[0046] 本发明提供了选择非特异性结合倾向较低的结合剂的方法,以从候选药物池中富集展现较小非特异性结合的那些药物,以及用于比较不同候选药物产品的预测药物动力学性能。此类方法可以在药物发现期间使用,任选地在筛选早期使用,或告知关于主要分子选择用于药物开发的决定。

[0047] 本发明提供了一种体外筛选方法,其包括;

[0048] (i) 提供各含有编码结合剂的DNA的真核细胞克隆文库,

[0049] (ii) 在用于表达所述结合剂的条件下体外培养所述克隆,其中所述结合剂呈现于细胞表面上,

[0050] (iii) 使所述结合剂暴露于非靶分子的基质,允许结合,

[0051] (iv) 丢弃对所述基质展现较大结合水平的细胞,

[0052] (v) 选择对所述基质展现较低结合水平的细胞,以提供所选细胞群,所选细胞群富集表达对结合非靶分子倾向低的结合剂的克隆。结合剂和由此展示其的细胞从而根据其对基质的相对结合加以分离。丢弃结合到基质的细胞,同时收集非结合细胞。

[0053] 细胞表面上所展示的抗体价数高将有助于低亲和力交叉反应性的检测。当展示结合剂的细胞群越过基质时,可以根据延迟的通过来体现对基质的结合。与随后出现的展示展现非特异性相互作用的结合剂的克隆相比,展示相互作用潜力低的结合剂的细胞更容易

前行通过基质且能够加以收集。使克隆文库越过基质从而实现克隆随时间的分离,原因是展示对结合基质的一种或多种组分的倾向较大的结合剂的那些克隆将花费更长的时间越过基质或甚至可能根本不从基质中出现。所述方法可以包括收集较快越过基质的细胞以及丢弃较慢通过和/或保持结合到基质的细胞。基质通常将包括其上固定一种或多种非靶标组分的固体或半固体底物(例如珠粒,任选地堆积于柱内)。能够测试多种非靶分子,例如肝素硫酸盐蛋白多糖和血流中遇到的其它丰裕组分。基质可以包括糖萼的一种或多种组分,例如玻糖醛酸、肝素硫酸盐。或者,可以对非靶分子使用流式分选或磁珠分选,其中收集参数是基于对所关注的相互作用搭配物、化合物或表面的结合程度。所述方法因此可以用于富集表达结合剂的细胞,在所述结合剂所施用的哺乳动物受试者体内,所述结合剂展现较低的与非靶分子结合的倾向。

- [0054] 可以在包括以下的方法中筛选出结合到一种或多种非靶分子的结合剂
- [0055] (i) 提供各含有编码结合剂的DNA的真核细胞克隆文库,
- [0056] (ii) 在用于表达所述结合剂的条件下体外培养所述克隆,其中所述结合剂呈现于细胞表面上,
- [0057] (iii) 使所述结合剂暴露于一种或多种非靶分子,允许结合,
- [0058] (iv) 丢弃对一种或多种非靶分子展现较大结合水平的细胞,以及
- [0059] (v) 选择展现较低结合水平的细胞,以提供所选细胞群,所选细胞群富集表达对结合非靶分子倾向低的结合剂的克隆。
- [0060] 此类方法可以应用于展示结合剂的高等真核细胞文库中的细胞(或细胞样品),以富集表达结合剂的细胞,所述结合剂对结合一种或多种非靶分子展现低倾向。
- [0061] 为了有助于鉴定和/或分离出表达识别非靶分子的结合剂的细胞,可以可检测方式对非靶分子进行标记,例如用荧光标记进行标记。荧光的使用允许在FACS中通过流式分选来分离细胞,其中未染色(未标记)的细胞能与染色(荧光标记)的细胞区分开来且能相应地引入收集或丢弃部分中。这可以方便地与标记组合以检测FcRn结合和/或检测结合剂的存在和表面展示水平(例如使用抗Fc抗体结合含有Fc区的结合剂)和/或检测靶标结合(例如使用经标记的抗原)。可以使用多种不同标记(例如不同波长的荧光团)同时进行标记和根据组合特性的分选。
- [0062] 本发明的其它方面涉及筛选与FcRn受体发生pH依赖性相互作用的多肽。抗体(或其它含Fc药物)可以与延长或缩短其半衰期的FcRn发生相互作用。如已经指出,包含Fc域的结合剂可以与内皮细胞上的FcRn受体相互作用并且免于降解。FcRn再循环路径的运转具有pH依赖性,需要在低pH的核内体代谢区内发生较强的结合以保持多肽安全地对接于受体上,且需要在较高pH的胞外环境中发生较低亲和力结合以便使多肽释放回至血流中。在一些情况下,出于控制半衰期之外的原因,希望增加对FcRn的结合。举例来说,使用“扫描抗体”方法<sup>32</sup>是希望所施用的抗体与FcRn在中性pH下充分接合以确保优先相互作用(与血清中的其它天然抗体相比)。抗体继而能将所结合的靶分子递送到核内体代谢区,在代谢区中,降低的pH使所结合的靶标释放且随后降解,例如在溶酶体内降解。

- [0063] 人们可以希望将根据可开发性的多个方面(例如根据本文所述的本发明的任一此类方面)的选择加以组合或与根据对靶标的结合的选择组合。为此目的,可以通过将结合剂暴露于靶标、允许同源结合剂识别靶标来执行选择,其中展示同源结合剂的克隆结合到靶

标。然后选择展示同源结合剂的一种或多种克隆。靶标可以携带可检测(例如荧光)标记以有助于选择展示同源结合剂的克隆。所述方法可以方便地包括同时测定结合剂的表面呈现水平和结合剂的靶标结合水平,以及协同选择展现较高表面呈现的同源结合剂的克隆。荧光活化细胞分选仪(FACS)的使用能够根据细胞发出的荧光分选细胞,并且表面呈现和靶标结合的并行选择可以通过使用检测表面呈现(相对于所结合的靶标)的不同荧光标记来进行。从而可以选择展现表面呈现高于所选阈值并且还展现靶标结合的克隆。

[0064] 本发明的另一个方面涉及结合剂选择的改善,所述结合剂对所关注的靶标产生高亲和力结合。本发明人注意到虽然结合剂高水平呈现于细胞表面上能够在展示文库中具有优势,但其可能会非期望地限制基于亲和力的针对靶标结合的选择的灵敏度或严格度。本发明人意识到限制结合剂在展示文库中的呈现水平可能有助于选择较高亲和力结合剂。

[0065] 根据这个方面,本发明提供一种方法,其包括

[0066] (a) 提供各含有编码结合剂的DNA的高等真核(例如哺乳动物)细胞克隆的体外文库,其中所编码的结合剂从弱活性启动子表达且/或在细胞表面上以每个细胞100-60,000范围内的拷贝数表达,

[0067] (b) 使所述文库暴露于靶标且允许同源结合剂识别所述靶标,其中展示同源结合剂的细胞结合到所述靶标,以及

[0068] (c) 分离出结合到所述靶标的细胞以提供展示同源结合剂的所选细胞群。

[0069] 从而获得克隆池,所述克隆池富集编码对靶标具有较高亲和力的结合剂的克隆。所述方法可以进一步包括

[0070] (d) 使所选细胞群暴露于针对所述靶标的额外一轮或多轮选择,任选地其中逐渐地减小靶标浓度以提高选择严格度,和/或

[0071] (e) 选出展示同源结合剂的一种或多种克隆,所述同源结合剂对所述靶标具有期望的结合水平。

[0072] 所用靶标的浓度可以预先测定或可以基于使用靶标浓度范围和选择抗原浓度,凭经验判断,其中细胞结合程度高于对照细胞(例如不表达结合剂或表达不识别靶标的结合剂的细胞)上所发现的结合程度。

[0073] 流式分选可以用于在靶标浓度和/或结合剂展示水平的不同条件下鉴定文库内具有所期望结合水平的克隆群体。所选克隆可以基于结合到细胞的荧光程度鉴定。当所结合靶分子的数目减少时(根据较低的靶标浓度和/或降低的结合剂展示水平),将经标记的细胞与未标记的细胞区分开来的能力减小,特别是未标记的细胞展现显著的自体荧光基线时。在这种情况下,使用可回收的靶分子(例如生物素化靶标)和磁珠(例如抗生素链菌素涂布的珠粒)分离经标记的细胞可以允许将标记程度显著减少的细胞与未标记的细胞分离,从而允许提高用于选择的严格度。

[0074] 所述方法可以用于鉴定以所期望的亲和力识别靶分子的结合剂,以根据亲和力选择结合剂和/或从克隆池中富集表达针对靶标的较高亲和力结合剂的那些克隆。靶标可以是所关注的任何分子,例如人类受体、配体、酶或其它多肽。

[0075] 本发明提供如上文在(a)下定义的文库,以及其用于选择对靶标具有期望亲和力的结合剂的用途。本文进一步描述此类文库和产生其的方法。

[0076] 如上文所概述,可以利用细胞文库、根据结合剂在细胞表面上表达的水平、基于不

同特征(可开发性和亲和力)区分结合剂。在药物发现期间,人们可能自然地希望根据良好亲和力和良好可开发性进行选择,在此情况下可以将本发明的多个方面进行组合。举例来说,人们可以使用以相对较低水平展示结合剂的细胞选择针对所关注靶标的结合剂且借此获得富集展示高亲和力结合剂的克隆的细胞群。然后可以在以较高水平表达结合剂的克隆中提供它们的编码DNA,以选择具有所期望的可开发性性状的结合剂。这些选择可以按照任一次序执行(根据亲和力进行选择,随后根据可开发性或类似的其它方式进行选择)且可以包括多轮选择(例如初始进行亲和力选择,然后进行可开发性选择,然后再进行多轮亲和力选择)。

[0077] 通过将结合剂的表面表达置于外部控制的开关或可调节元件下能够构建适用于根据亲和力和可开发性进行选择的双用途展示文库。举例来说,编码结合剂的DNA可以可操作地连接到启动子,所述启动子的表达能够通过向细胞培养基中添加诱导剂或抑制剂来调节。然后从细胞的外部控制表达,从而让方法的操作人员能够按照意愿上调或下调表达水平。因此,在诱导型启动子的情况下,外部操作人员添加诱导剂使得启动子活化,因此这可以视为外部诱导型,但是诱导最终施加于细胞内。多种诱导型启动子系统已有描述,经典的实例是四环素诱导型启动子<sup>33</sup>。结合剂DNA表达处于外部控制(例如处于诱导型启动子控制)下的双用途文库具有快速直接改变结合剂基因表达水平而无需将编码DNA再克隆到新细胞群的优点。

[0078] 涉及此类文库的结合剂展示方法形成本发明的一部分。一种鉴定识别靶标的结合剂的方法可以包括:

[0079] (i) 提供各含有编码结合剂的DNA的真核(例如哺乳动物)细胞克隆的文库,其中结合剂在细胞表面上的表达在外部能调节(例如其中结合剂的表面表达处于外部能调节的启动子的控制下),且其中结合剂呈现于细胞表面上,

[0080] (ii) 在低呈现于细胞表面上的条件下(例如其中启动子活性弱)培养所述文库中的细胞,

[0081] (iii) 使所述文库暴露于所述靶标,允许同源结合剂识别所述靶标,其中展示同源结合剂的细胞结合到所述靶标,

[0082] (iv) 选择展示同源结合剂的细胞,借此提供所选细胞群,

[0083] (v) 在使细胞表面上的呈现增强的条件下(例如其中启动子活性更强,任选地活性最大),培养所选细胞群,

[0084] (vi) 测定所述结合剂在所述多种克隆上的表面呈现水平,任选地通过用包含可检测(例如荧光)标记的试剂标记所述结合剂来测定,以及

[0085] (vii) 选出相较于其它克隆展现结合剂存在更高的表面呈现的一种或多种克隆。

[0086] 从而鉴定出具有良好可开发性特征的结合剂(和表达其的克隆),通过选择/富集展现较高表面呈现的结合剂(和因此克隆)加以选择和/或富集。所述方法因此能够提供克隆池,所述克隆池富集表达具有良好可开发性特征的结合剂的克隆。

[0087] 此类方法将在本文别处详细描述的本发明的多个方面有效地组合,即,靶标结合剂的选择与基于表面呈现结合剂的细胞的选择。如在本文别处所述的这些方面的特点可以用于组合方法中,包括选用于严格选择的靶标浓度、结合剂的表面呈现水平,和选择细胞的方法。

[0088] 能调节的启动子可以是诱导型启动子。根据所用诱导型启动子系统的类型,可以在培养基缺乏诱导剂(例如四环素,例如四环素或多西环素(doxycycline))的情况下获得低水平表达或基础表达。或者,可以添加低浓度的诱导剂。启动子的表达可以通过添加诱导剂或通过增加诱导剂在培养基中的浓度来滴定,以优选获得最大启动子活性。诱导型启动子的替代形式是可抑制型启动子,其中启动子的默认状态是活性,其活性是通过添加抑制剂到培养基中来衰减或阻断。

[0089] 通常方便的是,文库在其基础表达状态下开始,且在上调启动子活性以增强结合剂的细胞表面呈现且根据可开发性选择之前,对靶标执行初始多轮选择。然而,在一些情况下,可能期望根据可开发性选择,首先活化启动子,然后抑制启动子活性或使其活性衰减(例如通过从培养基中去除诱导剂),随后根据亲和力进行选择。如果使用可抑制型启动子,则此可为较方便的。因此,参照上文阐述的编号的方法步骤,可以执行步骤(i i) - (iv),随后执行步骤(v) - (vii),或可以执行步骤(v) - (vii),随后执行步骤(ii) - (iv)。

[0090] 本发明进一步提供了用于上述方法中的文库,其实例是含有编码结合剂谱系的真核(例如哺乳动物)细胞克隆的体外展示文库,其中结合剂的表达(和因此其在细胞表面上的呈现)处于四环素诱导型启动子的控制下。优选地,将编码DNA整合于细胞DNA中的固定基因座。此类文库可以通过包括以下的方法产生:

[0091] 提供编码所述结合剂的供体DNA分子,和真核(例如哺乳动物)细胞,

[0092] 将所述供体DNA引入所述细胞中,借此建立含有整合于所述细胞DNA中的供体DNA的重组细胞,

[0093] 其中所述供体DNA的表达置于四环素诱导型启动子的控制下以便呈现于细胞表面上,以及

[0094] 培养所述重组细胞以产生克隆,

[0095] 借此提供含有编码所述结合剂谱系的供体DNA的细胞克隆的文库。

[0096] 任选地,供体DNA的整合是通过将位点特异性核酸酶提供于细胞内来实现,其中所述核酸酶使细胞DNA中的识别序列裂解以建立整合位点,所述供体DNA在所述整合位点整合于细胞DNA中,整合是通过细胞内源的DNA修复机制发生。

[0097] 四环素诱导型启动子优选地位于编码结合剂的供体DNA分子上,但其在必要时可以分开整合。一般来说,重组细胞的供体DNA和/或细胞DNA将含有编码结合剂的DNA,所述DNA位于用于表达的启动子的下游。文库构建之后,供体DNA的表达能够从启动子诱导并且在使结合剂呈现于细胞表面的条件下培养细胞。

[0098] 结合剂DNA的表达置于诱导型启动子控制下的细胞还给评估结合剂在表达型克隆上的细胞表面呈现达到特定水平的速率以及通过比较短期诱导之后的呈现水平来比较多种克隆的速率提供了机会。这可以预先测定(例如4小时或8小时或12小时)或可以凭经验通过观察诱导之后的结合剂呈现外观得到。能够将细胞表面呈现相对于时间作图,起始时间点为从诱导型启动子开始表达,以观察表达和表面呈现的增加速率。多肽表达典型地随时间增加,直至其在细胞表面上达到稳定或平衡的水平。通过在早期测定和在克隆之间比较细胞表面呈现,在最终表达水平已达到之前,可能已经发现可开发性方面的初始迹象。这些可开发性方面可以包括多肽在溶液中浓缩的能力、其在溶液中的溶解度、对自缔合的抗性,和/或本文所论述的其它可开发性特征,此外为可从表达回收的产量。

[0099] 高等真核细胞上所展示的结合剂的转换、降解或内化的速率也可以用作此类可开发性特征的指标。这可以任选地在所展示的结合剂未经新表达的结合剂连续补充(因此表现为结合剂从细胞表面上耗竭,即,表面呈现水平降低)的条件下评估。可以标记结合剂,洗掉未结合的标记并且观察经标记的结合剂因降解和/或内化而随时间从细胞表面耗竭的速率。可以观察到结合剂在一些克隆上的耗竭水平高于其它克隆。选择表面呈现水平较高的克隆将选出可开发性特征更佳的那些克隆(例如在溶液中的溶解度较高、对自缔合的抗性更佳、非特异性结合较低)。

[0100] 在细胞文库内,可以观察到一些克隆的表面呈现增加速率大于其它克隆。这可以用作多肽的可开发性潜力的早期指标,其允许从文库中选择具有最优特性的克隆。

[0101] 根据本发明的方法,一般来说,一旦人们已经选中具有期望特性的克隆或选中富集具有期望特征的克隆的细胞群,则可以将所选克隆或群体用于一种或多种进一步筛选方法中,所述进一步筛选方法的实例提供于本文中。所选克隆可以一起或个别地培养。方法可以加以组合,以便针对多个特征筛选克隆,包括靶标结合和本文所述的多个可开发性特征。

[0102] 本发明的方法可以用于在抗体发现的最早阶段评估抗体在高浓度下的行为。在优选实施例中,通过使用文库选择技术(例如流式分选、基于珠粒的选择或色谱)筛选克隆文库来检测抗体或其它多肽结合剂在高浓度下发生自身相互作用的倾向。

[0103] 在本发明的任何方面中选择一种或多种含有编码所期望结合剂的DNA的克隆之后,可以回收编码所述结合剂的核酸且/或可以测定编码所述结合剂的核酸序列。编码结合剂的核酸(例如DNA)可以分离的形式提供,例如以重组载体提供。

[0104] 编码所关注结合剂的DNA可以在体外的宿主细胞中表达,任选地在用于表达可溶形式的结合剂的条件下表达。因此,宿主细胞可以分泌结合剂,从而促进其从细胞培养基中的回收。结合剂(例如来自稳定转染的宿主细胞)的产量可以是例如至少0.5mg/ml、至少1mg/ml、至少2mg/ml或至少5mg/ml。可以提纯且/或浓缩结合剂以提供所述结合剂的水溶液。有利的是,结合剂可以浓度为至少1mg/ml、任选地至少10mg/ml、至少50mg/ml或至少100mg/ml的溶液提供。在一些实施例中,结合剂是以浓度在50mg/ml与200mg/ml之间(例如在50mg/ml与100mg/ml之间)的溶液提供。

[0105] 使用本发明方法鉴定或选出的结合剂本身也作为本发明的方面提供于本文中,包括实例中所述的那些(包括(但不限于)包括所公开的序列的结合剂,例如本文阐述的抗体VH和/或VL域序列)以及作为执行本发明方法的结果而获得的其它结合剂。可以将所关注的结合剂配制成包括医药学上可接受的赋形剂的组合物。本发明扩展到此类组合物和其临床用途,包括用于通过疗法治疗人体或动物体的方法的结合剂。所述方法可以包括通过皮下施用来施用包含结合剂的组合物。包括结合剂于溶液中的组合物可以提供于注射用的预装药注射器中,任选地提供于试剂盒中,所述试剂盒包括一种或多种额外组件,例如针和/或产品信息说明书,包括通过注射(例如皮下注射)施用组合物的说明。

[0106] 在本发明的情形下,真核细胞优选高等真核细胞,例如哺乳动物细胞。哺乳动物细胞通常用于旨在临床用途的多肽药物产品(例如抗体)的大规模表达。因此,当在药物发现中评估候选多肽的特征时,使用哺乳动物细胞是有利的。在发现的早期期间,多肽结合剂可以其预定的最终分子形式在哺乳动物细胞中表达。举例来说,在所期望临床产品是全长抗体(例如IgG)的情况下,本发明的方法中可以使用表达全长抗体(例如IgG)的哺乳动物细

胞。

[0107] 本发明用于细胞群可为最有利,其中每个基因组存在恒定数目个(优选一个)所整合的结合剂基因以避免拷贝数影响或不均一性,所述拷贝数影响或不均一性起因于群体中的不同克隆中所表达的结合剂基因的数目或一致性。另外,编码结合剂的DNA优选地整合于所有克隆的细胞DNA中的相同基因座,以避免对表达的外在影响复杂化使编码型DNA的整合位置发生变化,所述相同基因座可以另外来源于转录活性不同的基因组区域。这具有使结合剂表达的转录归一化的好处。因此,优选地,本发明使用各含有编码不同结合剂序列的DNA的多种(例如大型文库)哺乳动物细胞克隆,其中编码型DNA处于细胞DNA中的固定基因座且其中编码的结合剂呈现于细胞表面上。然而,仍可以使用其中编码型DNA随机整合于细胞DNA中或提供于用于瞬时表达的质体上的克隆获得本发明的好处(例如实例11)。编码结合剂的DNA整合于单个或有限数目个基因座在必要时还能够更好地控制表达,例如使用诱导型启动子,并且优选地,编码结合剂的DNA在每个细胞中存在一次或在二倍体基因组的每个染色体拷贝中存在一次。因此,文库的克隆各自优选地表达结合剂谱系中的仅一或两个成员。

[0108] 本发明的各种特点进一步描述如下。整篇本说明书中所用的标题仅为了帮助导览,而不应解释为限定的。不同章节中所述的实施例在适当时可以组合。

## 具体实施方式

[0109] 本文内的标题和子标题仅仅为了帮助导览而纳入,而不应解释为限制性的。本发明的多个方面和实施例可以组合,并且基于可开发性选择多肽药物的方法宜涵盖本文所述的个别特点和步骤的依序和/或并行组合。

### 可开发性

[0111] 希望将根据多个可开发性特征进行的筛选整合到药物发现过程中,以在早期提供对可开发性的洞察和降低可开发性的风险。结合剂(和它们编码的细胞克隆)可以根据在此详述的一种或多种所期望的可开发性特征鉴定且选择。应理解,在本发明中,针对可开发性的筛选或选择通常是指预测的可开发性,其中所评估的内容是指示可开发性性状的替代标志,从而在早期发现中能够进行高通量选择和可开发性选择的整合。本发明允许通过富集具有较有利可开发性性状的结合剂鉴定和降低可开发性风险。然后可以通过例如下文所述的方法直接证实个别多肽的可开发性特征性。鉴定结合剂具有某种特征因此可以包括基于本发明方法获得的数据得出结论:所述结合剂经预测具有那种特征。

[0112] 鉴定结合剂是用于开发的候选结合剂或鉴定结合剂具有良好可开发性特征可以包括生成鉴定所述结合剂具有良好可开发性特征的报告。鉴定结合剂池富集具有良好可开发性特征的结合剂可以包括生成鉴定所述池富集具有良好可开发性特征的结合剂的报告。类似地,鉴定所选细胞表达具有良好可开发性特征的结合剂(或富集表达具有良好可开发性特征的结合剂的克隆)可以包括生成报告,所述报告鉴定所述细胞(例如所选细胞群)表达具有良好可开发性特征的结合剂(或针对具有良好可开发性特征的结合剂的表达加以富集)。报告可以是书面报告,其可以电子形式和/或印刷形式提供。报告可以提供关于可开发性的定量和/或定性信息,包括例如来自本文所述的一种或多种方法的数据。

[0113] 溶液中的溶解度、浓度和自缔合

[0114] 重组表达的多肽通常需要从细胞培养基中提纯且以较高浓度配制,例如用作药剂。举例来说,蛋白质可以100 $\mu\text{g}/\text{mL}$ 在瞬时细胞培养基中表达或以1 $\text{mg}/\text{mL}$ 在稳定的细胞培养基中表达,对于皮下施药而言,蛋白质需要以较高浓度(例如50 $\text{mg}/\text{mL}$ 或更大)提供给患者。

[0115] 用于可溶性结合剂(例如抗体)表达、提纯和定量的许多直接技术已为所属领域的技术人员所知,例如Walker等人(2008)<sup>34</sup>,Janson等人(2012)<sup>35</sup>。提纯的多肽结合剂在溶液中的浓度能够用许多方式测定。举例来说,能够测量溶液对280nm波长的光的吸光度且利用数值、基于多肽的消光系数测定浓度。或者,可以将测试样品与已知浓度的标准蛋白质进行比较,用于其的多种比色分析和荧光分析已为所属领域的技术人员所知。

[0116] 所期望的可开发性特征是水溶液中的高溶解度。多肽在10 $\text{mg}/\text{mL}$ 、20 $\text{mg}/\text{mL}$ 、50 $\text{mg}/\text{mL}$ 、75 $\text{mg}/\text{mL}$ 、100 $\text{mg}/\text{mL}$ 、150 $\text{mg}/\text{mL}$ 、200 $\text{mg}/\text{mL}$ 或更大时宜为可溶的。溶液可以是水性缓冲溶液,例如PBS。多肽的溶解度极限可以大于10 $\text{mg}/\text{mL}$ 、>20 $\text{mg}/\text{mL}$ 、>50 $\text{mg}/\text{mL}$ 、>75 $\text{mg}/\text{mL}$ 、>100 $\text{mg}/\text{mL}$ 、>150 $\text{mg}/\text{mL}$ 或>200 $\text{mg}/\text{mL}$ 。有利的是,含有多肽的溶液以大幅度低于其溶解度极限的浓度提供,以最小化分子的自身相互作用和当更紧密接近溶解度极限时可能发生的其它非期望影响。在产品质量和稳定性方面,多肽的自缔合能导致非期望的结果,包括增加的粘度或相分离。当达到分子的溶解度极限时,所溶浓度不能进一步增加,并且出现非期望的效应,例如分子将发生沉淀。在一些实施例中,所选多肽结合剂(例如改善的变异数体)的溶解度极限是在10 $\text{mg}/\text{mL}$ 与200 $\text{mg}/\text{mL}$ 之间,例如在50 $\text{mg}/\text{mL}$ 与100 $\text{mg}/\text{mL}$ 之间。

[0117] 可以将多肽溶液驱动到沉淀点且然后在过滤或离心之后,测定剩余可溶性材料的浓度,确定其溶解度极限。这也被称为蛋白质的最大溶解度。然而,测量自缔合的开始存在更灵敏的方式,其可以在低于沉淀所需浓度的浓度下发生。举例来说,在临界浓度下,单体分子开始形成二聚体和高阶可溶性聚集体。此类聚集体能够以不同方式检测。“临界浓度”定义为自身相互作用明显(使用这些方法中的一种或多种)的浓度。结合剂的所期望可开发性特征在于其临界浓度高,以便结合剂溶液的医药配制舒适地在低于临界浓度的情况下进行。临界浓度优选地大于10 $\text{mg}/\text{mL}$ 、>20 $\text{mg}/\text{mL}$ 、>50 $\text{mg}/\text{mL}$ 、>75 $\text{mg}/\text{mL}$ 、>100 $\text{mg}/\text{mL}$ 、>150 $\text{mg}/\text{mL}$ 或>200 $\text{mg}/\text{mL}$ 。在一些实施例中,所选多肽结合剂(例如改善的变异数体)的临界浓度在10 $\text{mg}/\text{mL}$ 与200 $\text{mg}/\text{mL}$ 之间,例如在50 $\text{mg}/\text{mL}$ 与100 $\text{mg}/\text{mL}$ 之间。

[0118] 自身相互作用的一些结果(例如沉淀)可以在延长的时间段内发生并且这与产品的存放期有关。缓冲液组成、pH和温度也可能有影响。例如浓度等参数因此可以在一组参考条件下测定,例如在4°C使用标准缓冲液,例如PBS pH 7.4。延长的时间段之后,可以证实多肽对自身相互作用的抗性,例如在这些条件下储存4周时间段之后(和任选地在所述时间段期间的其它时点)通过测试证实。多肽可以任选地在此类条件下抵抗自身相互作用至少六个月,证实溶液低于其临界浓度。

[0119] 就此而言,具有不良可开发性特征的多肽是自聚集迹象明显的多肽,即使在溶液中的浓度较低的情况下,例如在标准缓冲液(例如PBS)中在1 $\text{mg}/\text{mL}$ 或更小的情况下,例如在4°C。理想的多肽具有100 $\text{mg}/\text{mL}$ 或更大的临界浓度,即,其将抵抗此类自身相互作用,从而允许其浓缩到100 $\text{mg}/\text{mL}$ 。在这些极值之间,可能存在如下情况:使用本文概述的方法选出的改善的多肽展现比起始多肽改善1.5倍或更多的临界浓度。因此,“改善”能够结合起始多肽定义,例如在本文中别处所述的方法,其中将变异数体与亲本结合剂比较。差异也可以是在来自

不同克隆的结合剂之间比较,更通常例如在亲本结合剂的不同衍生物或变异体之间比较,或在文库中的整个克隆群体间比较,不论初始文库或从另一种方法得到的富集群体(任选地,从噬菌体展示选择的输出获得的克隆群体或从免疫获得的抗体群体)。可以将编码使用本发明以较高水平呈现于细胞表面上的结合剂的克隆与已经取消选择的同一群体内的编码具有较低表面呈现水平的多肽的克隆进行比较。能够将来自所选克隆的多肽的生物物理学行为与同一群体中的展现较低表面呈现水平的已取消选择的克隆所产生的多肽进行比较。受益于使用本发明发现的抗体或其它结合剂是如下抗体或结合剂:任何方法均能表明较高浓度水平能够在自身相互作用的开始之前达成,或在指定的分析内,自身相互作用的程度与对比多肽相比减少。举例来说,具有在10mg/ml自聚集证据的起始多肽可以是提高到15mg/ml或更大。或者,选自文库的最优多肽在10mg/ml可以抵抗自聚集,而由同一群体中的丢弃克隆所产生的其它多肽在10mg/ml或更小可以展现自身相互作用。在各种情况下,我们将亲本多肽或由丢弃克隆所产生的多肽称为“对比多肽”。

[0120] 作为选用于本发明中的自缔合和其它生物物理学特性对照,可以使用期望特性已知的抗体且将亲本抗体和改善的其衍生物(或所选文库成员相对于取消选择的成员)的相对性能与此进行比较。美国国家标准与技术研究所(The National Institute of Standards and Technology)已经使用抗体NIST RM 8671作为“黄金标准”参考抗体,其已被逾100位共同研究者广泛表征,所述比较作为美国化学学会(American Chemical Society)所出版的三册为一卷中的一册出版<sup>36</sup>。NIST RM 8671抗体显示具有最小的自身相互作用(Saro D等人,《对所提出的NIST单克隆抗体的可开发性评估(Developability Assessment of a proposed NIST monoclonal antibody)》<sup>37</sup>)。或者,抗体阿达木单抗(adalimumab)已表明展现最小的自身相互作用和交叉相互作用且已在一些研究中用作对照,例如Jain等人(2017)<sup>2</sup>和Sun等人(2013)<sup>38</sup>。出于定标目的,可以将如本文所述的结合剂(尤其是抗体)的临界浓度或溶解度与一种或多种此类参考抗体进行比较。

[0121] 自身相互作用能够用多种方式测定,从而允许在选用于高水平表面呈现的抗体与基于较低展示水平而取消选择的抗体之间进行直接且可定量的比较。

[0122] 尺寸排阻色谱(SEC)(例如高压液相色谱(HPLC))允许分离多肽的单体和多聚体形式(包括二聚体和高阶形式)。材料在二聚体和/或多聚体形式中的比例能够加以定量并且能够在结合剂之间(例如在亲本结合剂与改善的变异体之间)比较多聚体形成的程度,所述程度可以根据通过柱之后的延长的滞留时间和/或较宽的溶离峰检测。HPLC-SEC特征可以在单一浓度下测定且对多聚体比例和/或滞留时间进行比较。滞留时间或峰宽的任何可检测、可再现的缩短将视为改善。多聚体峰减小到亲本克隆的其数值的60%将视为改善。或者,多聚化阈值存在时的临界浓度对于每种克隆来说能够通过测试浓度范围并且测定这种多聚化阈值(例如5%)存在时的浓度来确定。或者,可以将多聚化的增加相对于时间作图且对多聚体积聚速率进行比较<sup>2</sup>。在各种情况下,临界浓度的增加代表改善。临界浓度的改善量级可以是例如至少1.5倍或至少2倍。

[0123] 溶解度极限的改善或临界浓度的改善任选地在1.5倍与50倍之间,例如1.5倍与15倍之间、1.5倍与10倍之间或2倍与10倍之间。

[0124] 测定多肽的自缔合的方法包括自身相互作用色谱(SIC)。在此技术中,使结合剂(例如抗体)固定在基质上且使相同结合剂的溶液越过基质。滞留时间相较于对照抗体延长

显示自身相互作用<sup>38</sup>(或与基质的相互作用)。或者,可以使用高通量方案,其中使不同结合剂固定在生物层干涉测量芯片(BLI)上并且通过浸没其可溶形式的溶液来测试自身相互作用,从而产生与结合的可溶性结合剂的数量有关的信号。“使用生物层干涉测量抗体克隆自身相互作用”(CSI-BLI)方案显示与较费力的方案(例如SIC)良好相关并且需要的材料较少。在各种情况下,可以将测试结合剂的滞留时间(在SIC的情况下)和在(CSI-BLI)情况下所获得的信号与亲本结合剂和已知抵抗自聚集的对照结合剂(例如上述基准抗体)进行比较。

[0125] 在亲和力捕捉式自身相互作用纳米颗粒光谱法(AC-SINS)中,测试抗体呈现于细胞表面上并且评估其与呈现于其它珠粒上的抗体发生亲合性自身相互作用的潜力<sup>39-41</sup>。相互作用引起的颗粒间距离的减少能够根据金胶体溶液的等离子体波长的增加检测到。AC-SINS是用于测定临界浓度的优选技术,因为这是灵敏且直接的分析。在优选实施例中,结合剂可开发性的改善能根据多肽在溶液中的临界浓度的增加检测到,其中临界浓度能通过在AC-SINS分析中测定等离子体波长的变化来测量。

[0126] 自缔合还能够利用静态和动态光散射(DLS)测量<sup>42,43</sup>,其中能够计算出样品中的分子的流体动力学半径和多分散性百分比,从而提供关于分子在溶液中的尺寸和形状的信息。利用DLS,评估与抗体浓度有关的互扩散系数作为自缔合的度量。可以使用其它方法测量自身相互作用,例如分析型超离心<sup>44</sup>膜渗透压测定法<sup>45</sup>和中子散射<sup>46</sup>。

[0127] 本发明的方法可以包括选择一种或多种克隆,所述克隆与其它克隆相比展现结合剂存在更高的表面呈现;以及鉴定由一种或多种所选克隆编码的结合剂在溶液中具有良好溶解度和/或对自缔合的抗性,其中所述一种或多种结合剂的临界浓度相较于由所述群体中的一种或多种其它克隆编码的结合剂或相较于亲本结合剂(其所选结合剂是变异体)增加至少10%、至少25%或至少50%,和/或其中所述一种或多种结合剂的临界浓度相较于由所述群体中的一种或多种其它克隆编码的结合剂或相较于亲本结合剂(其所选结合剂是变异体)增加至少1.5倍、至少2倍、至少3倍、至少5倍、至少10倍、至少20倍或至少100倍。

[0128] 如所论述,本文中的方法能够用于排除可开发性特征不良的结合剂(或减少可开发性特征不良的结合剂的频率或发生率),且有利于选择可开发性特征更佳的那些结合剂(增加具有更佳可开发性特征的结合剂的频率或发生率,富集具有更佳可开发性特征的结合剂)。此类方法不限于只鉴定和避免“问题”结合剂。本文所述技术的灵敏度允许其用于在展现良好溶解度的多种结合剂之间区分出最佳(例如最大可溶性)候选物。本发明还能够用于施加有利于在基于亲和力的选择期间维持或增强可开发性的选择性压力。

[0129] 对比多肽可以是来自未选定的克隆群体(或展现结合剂呈现水平小于测定的阈值水平的克隆)的任何结合剂。对比多肽可以具有至少5mg/ml、至少10mg/ml、至少20mg/ml、至少30mg/ml、至少40mg/ml或至少50mg/ml的溶解度极限,例如如本文所述的任何方法所测定。对比多肽可以任选地具有至少5mg/ml、至少10mg/ml、至少20mg/ml、至少30mg/ml、至少40mg/ml或至少50mg/ml的临界浓度,例如如本文所述的任何方法所测定。与对比多肽(例如来自未选定的群体的克隆,或亲本结合剂)相比,来自所选克隆的结合剂可以展现可开发性特征(例如溶解度极限或临界浓度)的至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%或至少100%改善。与对比多肽相比,来自所选克隆的结合剂可以展现可开发性特征的至少1.5倍改善,例如增加的溶解度极限或增加的临

界浓度对比多肽可以任选地具有至少10mg/ml、至少20mg/ml、至少30mg/ml、至少40mg/ml或至少50mg/ml的临界浓度。应理解,当将多肽与对比多肽的特性进行比较时,所述比较是在与所属领域中使用对照的标准条件相同的条件下进行。

[0130] 所述方法可以包括证实临界浓度的改善,例如通过在AC-SINS分析中测定等离子体波长的变化而在实验上测量临界浓度。用于测定参数(例如溶解度极限和临界浓度)的标准缓冲液条件包括PBS pH 7.4, 4°C。在将新鲜合成的多肽配制成所述水性缓冲溶液之后可以立即测量参数。或者,可以在标准缓冲液条件下储存一段时间(例如1小时、1周、2周、28天或6个月的储存时间)之后测量参数。

[0131] 人们可能希望在溶液恢复到4°C以便测量之前的储存期间暴露于较高温度(在此情况下,温度可以增加到例如25°C、37°C或50°C)之后测定参数。

[0132] 非特异性结合

[0133] 非特异性结合、“多反应性”、“多特异性”或“低特异性”可以指结合剂与除其同源靶标之外的非靶分子的相互作用,例如多肽可以非特异性结合到疏水性、带负电或带正电的表面。多反应性多肽可以被描述为显示“粘性”,这反映了除结合剂多肽与其靶标之间的特异性识别之外,其还通过这种类型的相互作用结合到非靶分子。在许多情况下,非特异性结合表现为低亲和力相互作用,但其仍存在问题,其中非靶分子是丰裕的。多反应性可归因于结合剂多肽表面上的簇聚的疏水性或带电荷氨基酸残基,在抗体的情况下,这可以位于重链和/或轻链可变域上,从而与其它(非靶标)分子产生一类非特异性相互作用。举例来说,多肽可能展现结合到疏水性表面,如果多肽在其表面上显示一个或多个疏水性补丁,或如果疏水性补丁通过多肽在溶液中的展开而暴露,则结合到疏水性表面可能会发生。或者,多肽可能展现结合到带负电的表面或携带净负电荷的分子(在中性pH下),例如DNA的带负电主链,或其它带负电的聚合物,例如肝素或乙酰肝素硫酸盐。不论负责这些非特异性相互作用的潜在分子基序,多反应性通常是候选多肽药物的非期望特点。多肽药物的多反应性与不良的药物动力学有关,例如体内短半衰期和/或药物从循环到组织的不良吸收。

[0134] 在本发明的各方面中,使表达结合剂的细胞(例如文库,或来自文库的细胞的样品)暴露于一种或多种非靶分子或“多反应性探针”。这能够用于区分呈现结合剂的群体内的细胞克隆,所述结合剂与来自那些细胞克隆的所述多反应性探针相互作用而不结合到此类探针或对此类探针的结合较低。多反应性探针可以包括以下中的任一种或多种:核酸(例如DNA)、抗生蛋白链菌素、肝素、乙酰肝素硫酸盐、软骨素硫酸盐、羧基聚葡萄糖、其它硫酸化蛋白多糖、胰岛素、脂多糖、杆状病毒、KLH、FcRn、层粘连蛋白、胶原蛋白、触发因子、Hsp70、Hsp90或其它热休克蛋白或伴随蛋白、玻糖醛酸或其它糖萼组分。非靶分子可以分离的形式呈现或以混合的制剂呈现,例如来自哺乳动物细胞(例如CHO细胞)的膜制剂可以用于测试多反应性。非靶分子可以是体内(例如哺乳动物中,例如人体中)发现的分子。其可以发现于细胞外基质或血流中,和/或细胞表面上。也可以制备非靶分子的合成替代物且用作多反应性探针。

[0135] 选择多反应性探针是为了根据非特异性结合的特定方式进行筛选,例如检测对带负电表面的非特异性结合,人们可以选择在中性pH携带净负电荷的多反应性探针,例如肝素硫酸盐、DNA或抗生蛋白链菌素,和/或具有例如延伸的带负电表面区域的分子。FcRn.适合的多反应性探针可以基于计算出的其等电点(pI)鉴定。pI是蛋白质电荷的度量且定义为

蛋白质不携带净电荷的pH。可以选择pI小于6的多反应性探针以便检测对带负电表面显示非特异性结合的结合剂。实例已为熟练的技术人员所知且包括抗生蛋白链菌素、核酸和硫酸化蛋白多糖，例如肝素硫酸盐或羧基聚葡萄糖。反之，为了检测到对带正电表面的非特异性结合，可以选择具有带正电表面区域和/或碱性pI（例如pI大于8）的多反应性探针。此类探针还可用以检测具有带负电表面补丁的结合剂，带负电的细胞表面可能在体内产生非期望的静电排斥力。为了通过疏水性吸引力探测多反应性，可以选择在其表面上具有一个或多个疏水性区域的多反应性探针，例如Hsp70或Hsp90。

[0136] 有多种方法可用于筛选特异性低的个别克隆。交叉相互作用色谱 (CIC)<sup>47</sup> 是一种方法，其中使靶分子或混合物（通常为多株抗体制剂）固定于色谱基质上并且测量各种抗体的滞留时间。通过与固定的分子或树脂自身的相互作用发生的滞留延迟是低特异性的指示。这种方案已经在许多研究中用于表征例如重组抗体<sup>2,15,38</sup>。如上文所述的尺寸排阻色谱还可以用于测定结合剂与基质的相互作用。在本文中，应理解，在这些方法中，所关注的非靶分子占据“靶标”的位置，其为据以测试结合剂相互作用的分子组分。

[0137] 与替代基质或靶分子分相互作用也已经用于抗体和其它多肽的表征。举例来说，可以根据对存在于糖萼中的丰裕分子（例如乙酰肝素硫酸盐蛋白多糖 (HSPG)，其由核心蛋白与所连接的乙酰肝素硫酸盐 (HS) 葡糖胺聚糖 (GAG) 链构成）的结合来筛选细胞。鉴于此类分子在糖萼中的较大表面区域和数量，因此这是一类特别适于测试的分子。肝素亲和层析涉及使样品越过含有固定化肝素的基质<sup>48</sup>。预制备的树脂获自许多来源（肝素琼脂糖 (Pharmacia)、生物凝胶肝素 (Bio-Rad, 奥地利维也纳)、Eupergit肝素 (Riihm Pharma, 德国威特塔得 (Weiterstadt, Germany)) 和 Toyopearl肝素650M (TosoHaas, 德国斯图加特 (Stuttgart))）。结合物质将在其通过时被滞留或延迟。未结合物质将通过或在洗涤后去除。利用改变的缓冲液条件（例如增加盐浓度）去除所结合的材料。这种方法能够用于筛选可溶性抗体并且测定他们与肝素相互作用的程度。疏水相互作用色谱还能够用于根据抗体与疏水性基质发生相互作用的倾向表征抗体<sup>49,50</sup>。

[0138] 在一种或多种非靶分子呈现于基质的情况下，结合剂对基质的结合可以体现为细胞对基质的结合增加，例如细胞越过基质时被阻滞，或结合到经非靶分子涂布的珠粒（例如磁珠）。显示较大结合到一种或多种非靶分子的结合剂从而能够通过去除（丢弃或不选择）显示较大结合的细胞部分来分离，而显示较小结合的那些细胞可以回收且任选地适用于其它步骤。如上文所论述，对比多肽和对照多肽能够用于证实本发明引起的改善。改善可以通过鉴定非特异性相互作用明显发生时的浓度来确定，所述浓度任选地在标准条件下定义。或者，非特异性相互作用的程度可以在固定的浓度下测定。分析非特异性结合的方法可以向结合剂产生“粘性度量”，从而对可以与其它结合剂比较的非特异性结合提供定量度量。因此，任何指定分析的“粘性度量”将是测试多肽的数值与已知具有极低非特异性相互作用的对照多肽之间的差异。举例来说，已批准的抗体阿达木单抗<sup>2,38</sup>或抗体NIST RM 8671 具有最小的自身相互作用或与其它多克隆 IgG 分子缔合，如通过自身相互作用色谱 (SIC) 和交叉相互作用色谱 (CIC) 方法所估算。（Saro D等人，《所提出的NIST单克隆抗体的可开发性评估 (Developability Assessment of a proposed NIST monoclonal antibody)》<sup>37</sup>）。根据本发明的已批准的多肽是观察到“粘性度量”相较于对比克隆改善的多肽。对比克隆可以是正改善的起始克隆，或可以是通过使用本发明取消选择的克隆。

[0139] 许多方法(包括这些色谱方法)典型地用于表征个别抗体和产生“粘性度量”。色谱基质能够用于基于其与所固定分子的相互作用来分离细胞。举例来说,固定于溴化氰活化琼脂糖上的凝集素已被用于分离T细胞群<sup>51</sup>。此类系统可加以修改以基于其与所固定靶标(例如多克隆抗体、肝素硫酸盐或其它测试分子)的相互作用分离表达抗体的细胞。在与所固定的靶标或载体基质存在相互作用的情况下,带有此类抗体的细胞相对于无相互作用的细胞将被滞留或延迟。当用于分离展示具有不同结合倾向的抗体的细胞时,可以修改上样缓冲液或洗涤缓冲液以实现所期望的严格度。

[0140] 也可以使用非色谱方法鉴定和定量个别抗体内的低特异性。举例来说,Hotzel等人(2012)在ELISA中利用抗体对杆状病毒颗粒的结合来鉴定展现非特异性相互作用的抗体<sup>16</sup>。以类似方式可以使其它测试分子(例如肝素硫酸盐)固定于或呈现于珠粒上,以测试与个别抗体的相互作用。非色谱方法(例如这些方法)可以适于从高等真核细胞上所展示的文库中鉴定出展现低特异性的克隆。

[0141] 已经制备洗涤剂溶解型膜蛋白的混合物,生物素化且用于鉴定展示具有低特异性的抗体的酵母文库内的克隆<sup>19</sup>。使用酵母文库可以耐受洗涤剂的存在,但不大可能适合于基于高等真核细胞(例如哺乳动物细胞)的展示系统。用于鉴定低特异性相互作用的测试分子或混合物可以用荧光团或分子(例如生物素)标记,所述荧光团或分子有助于标记或回收经抗生蛋白链菌素涂布的表面上的分子和其复合物。举例来说,例如荧光团标记的软骨素硫酸盐或肝素硫酸盐等分子(例如得自AMS目录号AMS.CSR.FACS-A1、C1或D1或E1或AMS.CSR.FAHS-P1)可以用于在流式分选中根据与经标记的测试分子相互作用的程度分离克隆。多肽在细胞表面上的高亲合力表达增加了方案的灵敏度,尤其当使用多价靶分子时。文库内的克隆可以基于荧光分子的结合、通过流式细胞术分离。这些测试分子可以联合其它经标记的分子使用,以预先、同时或随后根据其它期望特性进行选择,例如对靶标的结合,或所关注的其它分子(例如Fc受体)的其它结合/回避。其它方法包括如上文所提及的AC-SINS。在这种技术中,使测试结合剂呈现于细胞表面上并且评估其与其它细胞或珠粒上所呈现的一种或多种非靶标组分相互作用的潜力<sup>39</sup>。根据金胶体溶液的等离子体波长的增加能够检测到相互作用后发生的颗粒间距离的减少。

[0142] 可以使表达结合剂的细胞暴露于一种或多种非靶分子,其中克隆混合物(例如文库,或来自其的样品)一同存在于一个容器中,对于例如FACS等方法或对于色谱技术来说,这是方便的。在其它情况下,可以在独立的容器中使表达结合剂的细胞暴露于一种或多种非靶分子,例如每个容器一种克隆,并且然后能够个别地测量且比较所产生的相互作用或结合水平,对于测量颗粒间距离的方法(例如AC-SINS)来说或当正比较的表达结合剂的克隆数目相对较小时,这可以是更方便的。

[0143] 因此,能够将荧光标记的多反应性探针与细胞群混合并且使用检测或分离方法将表达多反应性结合剂(如根据对多反应性探针的结合所鉴定)的细胞克隆与不表达多反应性结合剂的结合剂表达克隆区分开来。未能结合经标记的探针的细胞克隆能够通过流式分选、磁珠分离和其它分离方法分离出来,以实现针对表达多反应性抗体的克隆的富集。实例6表明其不仅能在群体内的多反应性克隆与非多反应性克隆之间实现充分的区分,而且此能够使用允许其分离的直接实用步骤进行。

[0144] 鉴定候选药物的可开发变异体

[0145] 虽然本发明可以在药物发现的所有阶段使用(包括结合剂的早期选择(例如从初始文库或免疫或其它展示方案所产生的所选群体),和稍后比较一组入围候选分子的质量),其也适用于其中所关注的结合剂已经得到鉴定、但然后发现一种或多种可开发性特征需要改善的情形。可以发现来自任何来源的已鉴定结合剂展现不大理想的可开发性特征,并且在此类情况下,优选的是可以改进现有分子的序列,而非再次从头开始新的药物发现程序来找到替代分子。

[0146] 本发明的方法可以用于鉴定结合剂的变异数体,其中已鉴定所述结合剂的一种或多种可开发性特征(例如所论述的自缔合、溶解度、非特异性结合等)需要改善,且其中本发明用于预测一种或多种变异数体是否展现改善的可开发性。因此,本发明的选择方法可以对展示“亲本”结合剂的变异数体的细胞群执行。虽然这些可以称为文库,并且在一些情况下将展示变异数体结合剂的大型多样化群体,但是用于比较的克隆数目在一些情况下可以是相对较小的,例如最多10种。本发明的方法因此可以包括提供克隆文库或多种克隆,其中结合剂呈现于细胞表面上,其中克隆是通过产生亲本结合剂序列的变异数体序列且将编码变异数体的DNA引入细胞中以便将DNA整合于细胞DNA中来产生。适合的方法和技术在本文的其它章节中详述。

[0147] 产生变异数体的“亲本”结合剂可以是已经在一种或多种可开发性分析中鉴定出因性能不良而需要改善的结合剂。或者,可能只是希望研究其可开发性是否能通过序列变异改善。本文论述了可开发性的各个方面并且可以鉴定亲本分子的这些可开发性中的任一种需要改善。可以发现亲本分子具有例如小于50mg/ml、小于20mg/ml、小于10mg/ml、小于5mg/ml或小于1mg/ml的溶解度极限(最大溶解度)。可以发现亲本分子具有例如小于50mg/ml、小于20mg/ml、小于10mg/ml、小于5mg/ml或小于1mg/ml的临界浓度。亲本分子可能会在溶液中展现非期望的聚集且/或可能在溶液中不能浓缩到大于1mg/ml而不聚集和/或沉淀。亲本分子可以对一种或多种非靶分子显示非特异性结合。可以鉴定亲本分子对FcRn的结合且/或与FcRn的解离需要改善。

[0148] 任选地,对亲本多肽序列执行生物信息学评估以鉴定序列的潜在特点,其中预测突变会减少经鉴定的可开发性问题且从而改善性能。从而可以鉴定出一个或多个经预测与可开发性(例如溶解度、自缔合、非特异性结合)有关的氨基酸位置。

[0149] 此类生物信息学评估能够用于告知突变策略。因此,能够产生亲本多肽序列的变异数体,任选地包括在生物信息学评估中鉴定的一个或多个氨基酸位置的突变。从而可以使亲本结合剂的多肽序列的一个或多个氨基酸残基发生突变,所述突变经预测会促进自缔合、聚集及/或非特异性结合,且/或减小溶解度。突变产生编码亲本序列的一种或多种变异数体的DNA,所述DNA可以引入高等真核细胞中以产生编码变异型结合剂的细胞群(其方法描述于本文中)。可以在用于表达结合剂的条件下培养细胞,其中结合剂呈现于细胞表面上。表达结合剂的多个细胞可以用作如本文所述的文库且可以执行选择以鉴定结合剂发生较高表面呈现的克隆,作为改善的可开发性特征的指标。

[0150] 在如本文所述的不同实例中,利用序列分析鉴定潜在的问题残基。或者,可以使用随机序列变异。产生变异数体和衍生物文库的方法描述于本文中别处。可以产生个别变异数体并且评估改善的生物物理学特征。能够建立许多此类变异数体的大型文库和直接根据改善的特征(例如对自聚集的抗性)进行选择将大大有助于发现溶解度特性最优的抗体和其它结

合剂。由于有益于溶解度的变化能够同时减少靶标结合,尤其是涉及互补位残基时<sup>14</sup>,因此本发明的方法可以将针对可开发性进行的选择与根据所滞留的靶标结合进行的选择加以组合。此类选择任选地同时执行,并且根据亲和力和溶解度同时或依序筛选的方法已描述。

[0151] 包括基于结合剂的表面呈现水平进行选择的本发明方法可以用于鉴定溶液特性改善的变异数,如所论述。举例来说,改善“亲本”结合剂(例如抗体)的可开发性特征的方法可以包括

[0152] 将突变引入结合剂的氨基酸序列以产生变异数,

[0153] 将编码变异数的DNA引入真核(例如哺乳动物)细胞以提供各含有编码衍生物抗体的DNA的多种细胞克隆,

[0154] 将编码亲本结合剂的DNA引入真核(例如哺乳动物)细胞以提供含有编码亲本结合剂的DNA的细胞克隆,

[0155] 在体外,在使所述结合剂呈现于细胞表面上的条件下培养所述克隆,

[0156] 测定所述结合剂在多种克隆上的表面呈现水平,

[0157] 选择衍生物抗体的表面呈现高于表达亲本结合剂的克隆的一种或多种克隆,以及

[0158] 鉴定由一种或多种所选克隆编码的与亲本抗体相比具有改善的可开发性特征的一种或多种变异型结合剂。

[0159] 自身相互作用倾向与非靶标相互作用倾向之间有关。已发现少数氨基酸变化(1-3)对两个方面均能产生有益的作用<sup>6,7,14</sup>。在其它情况下,低特异性的证据表明可能存在有限的自身相互作用。在任一情况下,也可以使用本发明的方法,包括选择展现较低非特异性结合的结合剂。因此,如本文中别处所论述,所述系统可以用于鉴定与其它分子的非期望非特异性相互作用,亦称为“低特异性”以及“多特异性”。本发明有益于在高等真核细胞(例如哺乳动物细胞)上的表达的背景下执行此类筛选,其中例如糖基化等修饰更密切地反映了在生产型细胞系中所发现的,所述细胞系典型地用于生产临床中所用的产品。多肽在细胞上的高呈现水平<sup>23</sup>将增加任何相互作用的亲合力,借此用于增加系统在检测低亲和力非期望相互作用时的灵敏度。另外,高等真核生物自身(或内质网和高尔基体的环境)的表面能够充当基质,其中结合剂可以在相对较高浓度下暴露于多种多肽,从而允许特异性低的结合剂与非靶分子在相同或相邻细胞上相互作用。这继而将降低呈现水平。这还可以引起呈现细胞与群体中的其它细胞聚集。所产生的细胞聚集体能够加以去除(例如通过过滤或沉积),从而使群体中的此类克隆耗竭。所聚集细胞的去除还可以用于减少自身相互作用克隆的呈现。

[0160] 在另一应用中,表达来自内源或外源基因的非所需靶标的宿主细胞可用于耗竭交叉反应性克隆。举例来说,表达糖萼组分的内皮细胞或哺乳动物细胞用所关注的基因转染。结合到靶标的克隆可以基于低表面表达或细胞聚集来耗竭。

[0161] 可以在起始克隆与使用本发明产生的改善克隆之间比较在多肽的单一浓度下进行生物物理学测量的数值。或者,可以比较测量非所需生物物理学参数时的浓度。方法可以包括选择展现一种或多种所期望可开发性特征改善的变异数,例如较大溶解度、在溶液中较低的自缔合倾向、对非靶分子的非特异性结合较低、较高临界浓度等。差异倍数或差异%是可量测的且实例数值在本文中别处提供。比较通常针对亲本结合剂进行。然而,当本发明用于根据本文所述的其它生物物理学特性选择时,由本发明所选的克隆产生的变异型多肽

与被本发明取消选择的对比多肽之间的比较还可以用于证实和定量改善。

[0162] 产生变异体和衍生物文库的方法

[0163] 依循本文所述的任何选择方法,能够从一个或多个所选细胞中回收编码所展示的结合剂的DNA,所述DNA任选地经突变而产生变异体,且/或对其进行亚克隆,例如允许在第二展示系统内进行额外多轮选择。这可以是另一种真核展示系统,其引起不同的表面呈现水平,或可以是完全不同的展示系统,例如噬菌体展示。第二系统可以使用分泌式表达且/或可以允许如此前所述的直接功能选择,参见参考文献<sup>52-54</sup>和W02015/166272。或者,输入DNA编码型结合剂可以是来自未选定文库的结合剂群体或免疫或另一种展示技术(例如酵母或噬菌体展示)所产生的群体。

[0164] 相应地,在利用本发明方法产生文库之后,可以选择且使用一种或多种文库克隆产生另一个第二代文库。当已通过将DNA引入如本文所述的真核细胞来产生文库时,可以培养文库以表达结合剂,并且可以回收表达所关注的结合剂的一种或多种克隆,例如通过选择针对如本文在别处所述的靶标的结合剂来回收。这些克隆随后可以用于产生含有编码第二谱系的结合剂的DNA的衍生物文库。

[0165] 在其它情况下,希望产生亲本结合剂的变异体,以便提供多种变异体,从中选出展现改善的可开发性特征的变异体。

[0166] 为了产生衍生物文库,使一种或多种所回收的克隆的供体DNA发生突变以提供第二谱系的结合剂。类似地,为了产生变异体,使编码亲本结合剂的DNA发生突变。突变可以是一个或多个核苷酸的添加、取代或缺失。通过一个或多个氨基酸的添加、取代或缺失,突变将改变所编码的结合剂的序列。突变可以聚焦于一个或多个区域,例如抗体分子的一个或多个CDR,提供结构类别共同的结合剂谱系,所述结合剂谱系在一个或多个多样性区域不同,如在本文别处所述。

[0167] 一般来说,可以在DNA或RNA层面进行核酸序列的操纵和/或修饰。除非上下文另外要求,否则本文中提及DNA可以因此推广到包括其它等效核酸(例如RNA)。因此,提供、分离或使编码结合剂的RNA突变是提供、分离或使编码结合剂的DNA突变的替代方式。RNA任选地用于产生cDNA。

[0168] 产生衍生物文库可以包括:从一种或多种回收的克隆中分离出编码结合剂的核酸(例如分离出供体DNA,或其编码的RNA)、将突变引入核酸(例如DNA)以提供编码结合剂的第二谱系的供体DNA分子衍生物群体,以及将供体DNA分子的衍生物群体引入细胞以建立含有编码结合剂的第二谱系的DNA的细胞衍生物文库。

[0169] 分离编码结合剂的核酸(例如供体DNA)可以涉及从克隆中获得且/或鉴定DNA或RNA。此类方法可以涵盖扩增编码来自所回收克隆的结合剂的DNA(例如通过PCR来扩增)和引入突变。可以对DNA测序且合成所突变的DNA。

[0170] 可以替代地通过诱导克隆内的DNA突变而将突变引入一种或多种所回收克隆的供体DNA中。从而可以从一种或多种克隆建立衍生物文库而不需要分离DNA,例如通过使禽鸟DT40细胞发生内源突变。或者,编码结合剂的基因可以存在于基因组内且通过引入寡核苷酸和短同源臂来执行诱变。已经证明,使用80bp的单链寡核苷酸修复有缺陷的GFP基因已经实现高达45%的转染效率(Igoucheva, O., Alexeev, V., Yoon, K., 2001. 在哺乳动物细胞中利用小单链寡核苷酸进行靶基因校正(Targeted gene correction by small single-

stranded oligonucleotides in mammalian cells).《基因疗法(Gene Ther.)》8(5),391-399<sup>55</sup>;Liang等人,(2017).通过改善gRNA、Cas9核酸酶和供体DNA的设计和递送来增强CRISPR/Cas9介导的精确基因组编辑(Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA,Cas9 nuclease, and donor DNA).《生物技术杂志(J.Biotechnology)》241,136-146<sup>56</sup>)。

[0171] 抗体展示本身特别适合于建立衍生物文库。一旦分离出抗体基因,则可以使用多种诱变方案(例如易出错PCR、寡核苷酸定向诱变、链改组)建立相关克隆的展示文库,从中能够选出改善的变异体。举例来说,利用链改组能够将编码所选VH克隆群体、寡克隆混合物或群体的DNA亚克隆到载体中,所述载体编码适合的抗体形式并且编码适当格式化的VL链谱系<sup>57</sup>。或者且再次使用VH的实例,可以将VH克隆、寡克隆混合物或群体引入真核细胞群体中,所述真核细胞群体编码且表达适当格式化轻链搭配物的群体(例如与IgG或Fab格式化重链结合的VL-CL链)。VH群体可以源自上文所论述的任何来源,包括免疫动物的B细胞或来自所选噬菌体群体的scFv基因。在后一实例中,将所选VH克隆到轻链谱系中可以将链改组与重新格式化(例如格式化成IgG格式)组合成一个步骤。

[0172] 结合剂的细胞表面呈现

[0173] 结合剂滞留于细胞表面是展示文库的特点,原因是其提供结合剂与编码型DNA之间的物理性结合,从而有助于在物理分离表达具有期望特性的结合剂的细胞之后获取DNA。结合剂的表面呈现(或简称“呈现”)也可以称为结合剂的展示或表面表达。结合剂的表面呈现水平反映了其表达水平和其在高浓度延长暴露于细胞表面上之后保持的呈现水平。在本文所述的方法中,在表达不同结合剂的克隆之间比较结合剂的表面呈现相对水平且用于鉴定结合剂,所述结合剂具有较低的自结合倾向、较大溶解度且能够在适于医药用途的浓度下配制成水性缓冲溶液。因此,呈现水平可以用于选择具有期望特性的克隆。结合剂在展示文库上的表面呈现因此代表能够用于选择可开发性特征的内置特点。

[0174] 能够使用不同方式固定于细胞表面上。结合剂可以包括或连接到膜锚,例如跨膜域,用于结合剂的胞外展示。这可以涉及使结合剂与膜定位信号(例如GPI识别序列)或与跨膜域(例如PDGF受体的跨膜域)直接融合。<sup>58</sup>

[0175] 实现结合剂滞留于细胞表面上的其它方法包括结合剂与同一细胞内所表达的另一细胞表面滞留分子的间接结合。这种结合的分子自身可以是异源二聚体结合剂的一部分,例如系拴的抗体重链与未直接系拴的轻链搭配物的结合。WO2015/166272(并入本文中供参考)描述了用于使所表达的结合剂保留于其宿主细胞上的多种技术,包括将分泌式表达与膜展示组合的方法。从而可以保持所表达结合剂在细胞表面上的比例,同时使相同结合剂的其它拷贝以可溶形式从同一细胞中分泌。细胞可以保持大部分结合剂(例如80%或更多,或90%或更多)呈现于细胞表面上,少数分泌到培养基中。任选地,结合剂仅保持于细胞表面上而不以可溶形式分泌。

[0176] 如实例1中所说明,可以使抗体重链与PDGFR TM域融合并且连同其同源轻链一起表达以便表面呈现全长免疫球蛋白(例如IgG)。编码细胞内的结合剂或其多肽亚单元(例如抗体重链)的基因可以包括编码前导序列的DNA,以便经由内质网(ER)分泌到细胞表面。编码结合剂的基因可以包括编码膜锚(例如跨膜域,例如哺乳动物(例如人类)蛋白质的TM域)的DNA。或者,其可以包括编码翻译后膜锚(例如糖基化磷脂酰肌醇(GPI)锚)的附接信号的

DNA。多肽结合剂的C端区域通常选用于TM域的膜锚附接或融合。

[0177] 影响细胞表面表达水平的方法包括控制结合剂从其编码型DNA表达的水平,例如使用启动子控制,且此方法描述于本文中别处。表面表达水平能够可替代地通过影响编码表达结合剂的DNA剪接到编码跨膜域的外显子的程度来控制,例如如WO2015128509 (Glenmark Pharmaceuticals) 中所述。这种方案也在本文中举例说明,参见实例8a和实例8b,其描述了可以有效地用于本发明方法中的表达系统,其中希望结合剂存在不同的表面呈现水平。

[0178] 在本发明的实施例中,结合剂在细胞中表达且输送到质膜,作为膜蛋白滞留于细胞表面,例如结合剂可以包括一种或多种具有至少一个跨膜域或膜锚的多肽。举例来说,在多肽结合剂包括抗体重链(或其一部分)和抗体轻链(或其一部分)并且包括连接到重链和/或轻链的跨膜域或膜锚的情况下,结合剂在ER内合成且萌芽到高尔基体中,通过细胞内小泡输送到细胞表面,与质膜融合,之后通过其跨膜域或膜锚滞留结合剂并且向细胞外呈现。结合剂整合到膜内因此在细胞内发生,随后输送到细胞表面。在适用的情况下,多亚单元结合剂(例如包括各别重链和轻链的抗体,或其一部分)的组装典型地也在细胞内发生。

[0179] 细胞表面呈现水平可以按照拷贝数(每个细胞所展示的结合剂数目)测量。有多种方法可供利用,包括与校准珠粒进行比较以及配体浓度和受体占有率的史卡查图(Scatchard plots)<sup>59-61</sup>。通过了解细胞半径和关于可利用体积或表面面积的某些假设,这可以用于分别评估浓度或密度(实例3)。在涉及本发明的一般层面,拷贝数与在呈现水平增加时发现的较高浓度下实现的浓度有关。细胞表面上的拷贝数与浓度之间的关系也受到细胞尺寸的影响。由于抗体在较大细胞上将占据较大体积,因此相同数量的结合剂呈现于小细胞和大细胞上将产生不同浓度(比较细胞尺寸和所实现的浓度,参见实例3)。

[0180] 大量重组表达系统可获得并且在不同系统之间,在细胞表面上展示的结合剂的绝对数量将不同。当使用不同细胞(例如不同尺寸的细胞)、不同启动子、不同诱导机制、不同剪接机制、输送效率等时,熟练的技术人员在适当时能够针对呈现水平的范围来校正本发明的方法。然而通过举例提供以下指导。

[0181] 利用弱活性启动子,可以使结合剂以每个细胞100-100,000范围内的拷贝数呈现于细胞表面上。每个细胞的结合剂数目可以是至少100、至少1,000或至少10,000。每个细胞的结合剂数目可以是多达1,000、多达10,000或多达100,000。在优选实施例中,拷贝数是每个细胞约80,000或更少、每个细胞约60,000或更少、每个细胞约50,000或更少,或每个细胞约40,000或更少。为了促进检测,拷贝数可以是每个细胞至少100、至少1,000或至少10,000。因此,拷贝数可以在例如每个细胞1,000-60,000范围内。其可以是约10,000、约50,000或约60,000。

[0182] 利用强活性启动子,可以使结合剂以每个细胞100,000-10,000,000范围内的拷贝数呈现于细胞表面上。每个细胞的结合剂数目可以是至少100,000或至少1,000,000。每个细胞的结合剂数目可以是多达1,000,000或多达10,000,000。其可以是约1,000,000。

[0183] 当然,不同细胞上的确切的结合剂数目将是变化的,即使是单一克隆也如此,但一种克隆的细胞之间的此类拷贝数变化与本文所述的选择和富集方法所利用的克隆间拷贝数变化相比是微小的。实例数字和范围表示近似平均值(均值)。拷贝数可以是群体中的细胞的平均(均值)拷贝数。根据亲和力进行选择时,为了增强高亲和力克隆的结合和富集严

格度,优选较低拷贝数;而根据可开发性(例如结合剂的溶液特性)进行选择时,优选较高拷贝数。

[0184] 表达结合剂的细胞文库内的拷贝数可以在相对较低(例如每个细胞10,000、50,000、100,000或250,000个拷贝)到相对较高(例如每个细胞1,000,000个拷贝)的范围内。本文引用的拷贝数仅仅是引导,并且是基于悬浮培养的HEK细胞文库中所观察到的拷贝数,所述HEK细胞具有约10 $\mu\text{m}$ 的半径<sup>95</sup>。以指定密度呈现于大细胞表面上的多肽的拷贝数高于以相同密度在小细胞上呈现时的拷贝数,因此在较大或更小细胞产生的文库中观察到的绝对拷贝数可以相应地改变。参见实例3b。

[0185] 实施本发明的方法时,人们在各种克隆之间作比较时通常对相对呈现水平感兴趣,在此情况下,结合剂的绝对数(每个细胞的绝对拷贝数)不是关键,而是能够根据结合剂在细胞表面上展示的不同水平将克隆排序或区分开来。

[0186] 来自文库的代表性克隆样品的表面呈现水平能够加以测定,并且用于评估存在于文库克隆中的表面呈现水平的平均值(方式、均值或中值)和展度(范围)。

[0187] 能够利用任何适合方法测定多肽存在于细胞表面上的水平,以便比较其相对数量。在细胞之间测量结合剂呈现相对差异的优选方式是将细胞暴露于携带可检测(例如荧光)标记的试剂,允许所述试剂结合到结合剂,并且检测细胞上的标记的相对数量,其中来自可检测标记的信号较强(例如较多荧光单位)指示细胞上存在较高的结合剂呈现水平。当可检测标记是荧光时,可以使用荧光活化细胞分选仪(FACS)执行分选。或者或另外,在FACS的情况下,能够利用通过磁珠进行的选择。

[0188] 人们通常将选择结合到结合剂恒定区的试剂,以便所有结合剂都能够独立于其序列而同等地被标记。使用抗体和包括Fc区的其它结合剂时,宜用结合Fc区的试剂进行标记。举例来说,在结合剂是IgG抗体的情况下,可以使细胞与可检测试剂接触,所述可检测试剂结合到IgG Fc区,例如经标记的抗IgG抗体。适当时,能够对所述方法进行调适,例如其中文库包括其Fc区的序列不同的结合剂(一种结合Fc的非多样化部分的试剂),或能够利用结合剂分子的其它区域。可以将肽表达标签(例如血球凝集素(HA)、c-Myc)并入结合剂多肽(例如并入抗体支架中),从而能够通过检测结合到标签的试剂来检测所展示的结合剂。这已经在选择正确组装的抗体的上下文中预先描述,以便基于抗体表达水平将抗原结合信号归一化<sup>24,62</sup>。用于检测结合剂的表面呈现水平的试剂可以单独使用,即,无需根据其它特点(例如抗原结合或靶标特异性)进行协同检测或协同选择。因此,在一些实施例中,使用结合结合剂恒定区的试剂检测结合剂呈现水平的步骤不包含使用经标记的靶标检测结合剂。在其它实施例中,可以执行多参数选择。文库中的表达结合剂的克隆可以展现多种多样的表面呈现水平,反映了高水平的克隆间变化。对于来自文库的代表性细胞样品来说,表面呈现水平能够相对于观察到呈现水平的频率作图,例如在FACS检测之后。关于此的一些实施例结合图31描述于实例10并且绘示于图31中。检测样品上中值表面呈现水平或峰值表面呈现水平,并且能够观察且/或计算呈现水平的展度和其相对于中值或峰值的偏差。在表面呈现水平相对于细胞计数的图中,峰值表面呈现能够作为最高峰方便地可视化。

[0189] 在克隆群体中观察到的表面呈现水平的展度可以由相对于峰值(mode)的变异度表示。在克隆之间,克隆群体的表面呈现水平可以显示广泛的变异,特别是其中结合剂的结构(例如不同一级序列)展现较大变异度,使得一些克隆展示结合剂的密度比其它克隆高得

多。在一些克隆群体中,表面呈现水平的克隆间变异可以是较低的,但本发明方法可以用于设定阈值呈现水平以及将较高呈现克隆与较低呈现克隆区分开来。展度的一种统计学度量是四分位数间距,即上四分位数(75%)与下四分位数(25%)之间的差异。处于75%(表示为上四分位数的下边界)的拷贝数与处于25%(表示为下四分位数的上边界)的拷贝数可以相差至少1.5倍、至少2倍、至少2.5倍、至少3倍、至少3.5倍、至少4倍、至少4.5倍或至少5倍。

[0190] 在一些实施例中,当细胞群逐渐地富集结合剂存在高度表面呈现的克隆时,可以观察到表面呈现水平的范围和上述参考点之间的表面呈现差异倍数随着表面呈现水平的每一轮选择而减小。在一些实施例中,可以观察到峰值随着较高表面呈现的每一次富集而逐渐增加。选择之后,测定或观察结合剂群体的表面呈现水平增加的方法可以包括测定或观察平均(例如中值或峰值)拷贝数的增加,如通过例如FACS或本文所述的类似方法所测量。可以观察到所选群体(或其克隆)的平均表面呈现水平高于表达对比多肽的未选群体或克隆的平均表面呈现水平,例如高于表达群体所来源的亲本结合剂的克隆的平均表面呈现水平。举例来说,所选克隆或所选群体的峰值或中值拷贝数可以比对比克隆或亲本克隆高至少5%、至少10%、至少20%或至少25%,并且可以基于表面呈现水平的改善来选择所述克隆。

[0191] 为了选择表达结合剂的较高表面呈现的克隆,或为了选择富集编码展现较高表面呈现的结合剂的克隆的克隆群体,根据结合剂在细胞上的表面呈现水平能够将细胞分选到收集部分和丢弃部分。表面呈现高于预定阈值的细胞分选到收集部分并且表面呈现低于预定阈值的细胞分选到丢弃部分。表面呈现任选地成为在这个步骤期间执行选择/富集的唯一基础。这能通过使用标记所有结合剂的可检测试剂、任选地在不检测结合剂识别靶标的情况下促进。

[0192] 通过丢弃展现结合剂的较低表面呈现的细胞部分,从群体中耗竭表达具有不良可开发性特征的结合剂的克隆。选择剩余细胞的全部或一部分提供了所选细胞群,所选细胞群富集的克隆与其它克隆相比表达结合剂具有较高表面呈现的结合剂。从而能够鉴定所选细胞、克隆或群体与起始群体或文库相比且与未选定的细胞、克隆或群体相比具有更佳的可开发性特征。

[0193] 如上文所提及,结合剂在细胞表面上的绝对数将变化,并且熟练的技术人员将确定系统在使用时的适合阈值。举例来说,阈值可以预先确定以在表达最高表面呈现的结合剂的群体内选择特定百分比的克隆,例如基于文库中的初始测试样品。可以设定阈值以选择例如前50%的细胞、前30%、前25%、前20%、前15%、前10%或前5%的细胞。

[0194] 来自文库的代表性细胞样品的表面呈现水平已测定时,并且样品的拷贝数的峰值、中值、展度和/或四分位数间距已观察或计算时,能够设定适当的阈值拷贝数。使用FACS时,这将对应于每个细胞的阈值荧光强度,例如如果用于收集细胞的FACS阈值设定成收集中值荧光强度对应于样品峰值的文库细胞,则将收集具有所述荧光水平或更高的细胞。对于表面呈现水平的甚至更大富集来说,用于收集细胞的FACS阈值可以设定成收集荧光强度对应于样品中的75%细胞的荧光强度的文库细胞。

[0195] 阈值可以表示每个细胞所呈现的至少约100,000、至少约500,000或至少约1,000,000的结合剂数量。

[0196] 选择和富集

[0197] 选择克隆、选择细胞或选择群体可以包括在实体上将克隆、细胞或群体与其它克隆或细胞分离或从更宽的群体或文库中分离出来。所选克隆、细胞或群体可以分离形式提供。

[0198] 在选择包括实体分离多种细胞或克隆的情况下,这将典型地包括产生收集部分和丢弃部分。收集部分将富集展示具有已通过所述方法选择的特征的结合剂的克隆。富集意谓着这些克隆在群体中的相对丰度增加。富集是相对于选择步骤之前的细胞池或克隆池(例如文库或起始群体)来说的。通过在选择期间丢弃一部分细胞(例如展现结合剂存在较低表面呈现的细胞、具有较低亲和力的细胞等),排除表达具有较少期望特征(例如不良可开发性特征)的结合剂的克隆。因此,作为选择步骤的结果,群体中的具有较少期望特征的细胞/结合剂的相对丰度减小。所选细胞的全部或一部分(收集部分)在必要时可以前送到进一步选择方法,并且任选地选择一种或多种克隆用于独立的个别培养。任选地,一种或多种所选细胞或结合剂或所选群体可以用于建立衍生物文库,如在本文别处所述。

[0199] 选择克隆的收集部分时,所述方法的操作人员可以采用设定百分比或分率的“最好”或“最佳”克隆。这种原理的实施例已参照根据高度表面呈现进行的选择详细论述。应理解,当针对其它特征(例如对靶分子或非靶分子的结合)进行选择时,相同原理也适用。因此,当选择表达识别靶标的结合剂的克隆(例如亲和力选择)时,操作人员可以选择对靶标展现最高结合的设定百分比或分率的克隆(例如在使用经标记的靶标的方法中根据结合到克隆的可检测标记的数量所测量)。当选择表达对非靶分子显示减少的(或缺少的)非特异性结合的结合剂的克隆时,操作人员可以选择设定百分比或分率的对非靶分子展现最低结合的克隆(例如在使用经标记的非靶分子(例如本文提及的各种多反应性探针中的任一种)的方法中根据结合到克隆的可检测标记的数量所测量)。一般来说,当根据“良好”特征(例如表面呈现水平、对靶标的亲和力)进行正向选择时,人们将选择所述特征的水平高于其它克隆的克隆;而当针对“不良”特征(例如非特异性结合)进行选择时,人们将选择展现所述特征的水平低于其它克隆的克隆。

[0200] 如在本文别处所论述,选择阈值可以由操作人员根据现有情形确定,并且可以根据对群体中的样品(例如来自文库的克隆的代表性样品)的初始评估来引导。在所期望特征的情况下,阈值可以设定成富集具有最高信号水平的克隆,例如前50%、前30%、前25%、前20%、前15%、前10%或前5%的细胞(例如基于所结合的可检测标记的数量,其代表着表面呈现水平或靶标结合水平,或基于如通过AC-SINS所测定的颗粒间距离,或基于溶液中的单体比例)。所选前%的克隆变成收集部分,而其它丢弃。在针对非期望特征进行选择时(例如基于结合到多反应性探针的可检测标记的数量,或基质上的阻滞程度以检测非特异性结合),阈值可以设定成富集具有最低信号水平的克隆,例如选择一定百分比(例如50%、30%、25%、20%、15%、10%或5%)的展现最低数量的所结合可检测标记或基质上的最低阻滞程度的克隆。收集的实例百分比当然只是指导原则,并且只要操作人员遵守所述方法的原理,则数值精确度并不重要。

[0201] 量化和测量差异

[0202] 本发明的各种方法涉及比较结合剂和/或其编码克隆的特性,例如如下特性:呈现水平、结合、浓度(例如临界浓度)、溶解度极限等等。方法可以包括鉴定与群体中的其它结合剂相比(例如与一种或多种其它结合剂相比,或与群体平均值(均值)相比)更佳的结合

剂,和/或一种或多种特征相对于他们所来源的亲本分子改善的结合剂。也可以与基准结合剂进行比较,在抗体实例中,基准结合剂可以是NIST RM 8671抗体(Saro D等人,所提出的NIST单克隆抗体的可开发性评估(Developability Assessment of a proposed NIST monoclonal antibody)<sup>37)</sup>。用于选择和改善的期望特性论述于本文中别处且包括较大溶解度、较大临界浓度、较低非特异性结合和/或较高的靶标亲和力。相对术语,例如“较大”或“较小”、“较高水平”或“较低水平”、“较佳”或“较差”、“较多”或“较少”,通常是指相关实验背景下所观察到的差异,其允许基于其特性区分结合剂或克隆。差异可以具有统计显著性。差异可以任选地用%术语量化,例如至少10%、至少25%或至少50%的差异。或者,差异倍数可以认为是例如至少1.5倍、至少2倍、至少3倍、至少5倍、至少10倍、至少20倍或至少100倍。

[0203] 靶标结合

[0204] 可以选择结合到所关注的靶分子的结合剂,所关注的靶分子任选地可以是另一种多肽,例如受体、酶和/或疾病相关的多肽,例如肿瘤相关抗原。其它靶分子类别包括核酸、碳水化合物、脂质和小分子。示例性的结合剂和靶标详述于本文中别处。经典实例是抗体分子文库,其可以根据对所关注的靶抗原的结合来筛选。其它实例包括针对靶标MHC:肽复合物来筛选TCR文库或针对靶标TCR筛选MHC:肽复合物文库。

[0205] 选择编码识别靶标的结合剂(同源结合剂)的克隆可以包括使如本文所述的展示文库与靶标接触,借此使结合剂暴露于靶标,从而允许同源结合剂(如果存在)识别靶标,并且检测靶标是否被同源结合剂识别。然后可以选择展示同源结合剂的一种或多种克隆。

[0206] 靶标可以可溶形式提供。靶标可加以标记以有助于检测,例如其可以携带荧光标记或其可以生物素化。表达靶标特异性结合剂的细胞可以使用直接或间接标记的靶分子鉴定,其中结合剂捕捉经标记的分子。举例来说,经由结合剂:靶标相互作用结合到荧光标记的靶标的细胞能够通过流式细胞术检测和分选,以分离出所期望的细胞。涉及细胞测量术的选择要求靶分子被荧光直接标记或被能够用二级试剂检测到的分子标记,例如生物素化靶标能够添加到细胞中并且对细胞表面的结合能够被荧光标记的抗生素链菌素(例如抗生素链菌素-藻红蛋白)检测到。另一种可能性是使靶分子或结合到靶标的二级试剂固定于固体表面,例如磁珠或琼脂糖珠粒,以允许结合靶标的细胞富集。举例来说,经由结合剂:靶标相互作用结合到生物素化靶标的细胞能够在抗生素链菌素涂布的底物(例如抗生素链菌素涂布的珠粒)上分离。磁珠通过磁力回收珠粒而便于捕捉已经结合到生物素化抗原的细胞。最优靶标浓度可以预先确定或可以通过使用一系列浓度且与背景对照进行比较而凭经验确定。

[0207] 根据对靶标的结合进行选择的方法包括根据细胞上的靶标结合水平将细胞分选到收集部分和丢弃部分中,借此将呈现所结合靶标高于预定阈值的细胞分选到收集部分中并且将所结合靶标低于预定阈值的细胞分选到丢弃部分中。阈值可以相对于不展示同源结合剂的阴性对照细胞设定。在FACS中,当分选展示同源结合剂和非同源结合剂的细胞时,典型地将观察到阴性对照峰。可以设定阈值,以便将展示荧光的水平在统计上显著高于阴性对照的所有细胞分选到收集部分,或可以选择较高阈值以在选择同源结合剂和较大富集展示同源结合剂的克隆时实现较大置信度。如已经参照细胞表面呈现结合剂的测定所论述,可以使用确定适合阈值的细胞样品执行校准。能够实现从非结合剂中富集结合剂。富集较

高亲和力结合剂的方法在本文中别处进一步描述。

[0208] 针对靶标进行的选择可以作为另一步骤并入本发明的其它方法之前或之后,或纳入本发明的其它方法内。通过诱发起始域突变构建变异体文库以改善可开发性的各方面(例如溶解度或低特异性或最优FcRn结合),可能会有损于一些文库成员对靶标的结合(例如如果诱发接触CDR突变)。举例来说,根据结合剂呈现水平选择的克隆随后可以针对靶标选择。通过对以高度表面呈现水平展示同源结合剂的克隆使用协同选择,也能够同时测定表面呈现水平和靶标结合。此类方法可以包括如在本文中别处所述使用包含可检测标记的试剂测定结合剂呈现水平,并且可以进一步包括使结合剂暴露于靶标,其中所述靶标用第二试剂标记,所述第二试剂包含能够检测靶标结合的第二可检测标记。在使用荧光标记的情况下,可以利用FACS以同时根据结合剂呈现和靶识别来分选细胞。可以选择发出不同波长的荧光的标记,以便能够区分其不同信号。因此,本发明的方法可以包括使用针对每一者的不同可检测标记协同检测结合剂呈现水平和靶标结合水平。在其它实施例中,在不测定细胞上的结合剂呈现水平的情况下检测靶标结合,例如使用经标记的靶标检测结合剂的步骤不包含使用结合结合剂恒定区的试剂检测结合剂呈现水平。

[0209] 检测同源结合剂识别靶标之后,可以回收含有编码同源结合剂的DNA的所选克隆的细胞。然后可以鉴定编码结合剂的DNA,扩增且/或以分离的形式提供,借此获得编码识别靶标的结合剂的DNA。

[0210] 多重选择

[0211] 并行执行多重选择的构思能够扩展到基于本文所述的任何两个或更多个特征对细胞进行协同分选。在以上论述中,通过同时测定结合剂的表面呈现水平和结合剂结合靶标的水平以及协同选择展示展现较高表面呈现的同源结合剂的克隆来对细胞同时进行协同选择。还能够并行使用本发明的其它方法。方法可以包括同时测定以下中的任何两种或多种:

[0212] (i) 表面呈现水平

[0213] (ii) 非特异性结合到非靶分子的水平

[0214] (iii) 靶标结合水平,

[0215] (iv) FcRn结合水平,

[0216] 和相应地协同选择克隆。FACS使用发射不同波长的荧光的多重标记能够执行并行选择。

[0217] 可以通过连续或并行执行多种类型的选择来获得优势和协同作用,例如将针对溶解度进行的选择与针对非特异性结合进行的选择组合。对文库中的克隆群体应用的每次选择产生的进化压力有利于满足选择标准(例如表面呈现水平高)的变异体。根据单一参数进行的重复选择可以将进化驱向这个特征(例如高溶解度),但以其它质量(例如用于结合靶标的亲和力)为代价,使其它质量处于耗竭或损耗的风险中<sup>31</sup>。例如当使溶解度增加的突变也使亲和力减小时,这会发生。选择方法的审慎组合可以将群体的进化导向表达具有多重期望特征的多肽的克隆,从而允许在多肽药物所需的需求的完整范围内鉴定出性能最优(或至少可接受)的多肽。

[0218] 可以在根据增加的表面呈现水平进行选择之前或之后执行针对非特异性结合的选择。举例来说,人们可以执行根据结合剂在溶液中的溶解度和/或对自缔合的抗性将结合

剂区分或排序且/或富集在溶液中展现较大溶解度和/或较大自缔合抗性的结合剂的方法，从而产生所选克隆群体，然后从所选群体中筛选表达结合剂的克隆，所述结合剂对结合一种或多种非靶分子展现低倾向，借此鉴定出表达也具有低非特异性结合的结合剂的一种或多种克隆。

[0219] 在一些实施例中，针对非特异性结合的筛选能够与针对表面呈现水平的筛选同时整合。举例来说，能够执行细胞双重暴露于 (i) 用于结合所呈现的所有结合剂的试剂，其携带可检测标记(例如荧光标记的抗Fc)；和(ii) 一种或多种非靶分子(例如避免非特异性结合的肝素或其它分子)，其携带不同的可检测标记(例如不同荧光波长的可检测标记)。克隆的双重染色允许基于表面呈现水平和非特异性或脱靶结合来区分克隆。可以设定分选阈值以收集展现较大表面呈现水平和对非靶分子的结合较低的细胞。这将消除或至少减少具有不良可开发性的抗体在所选克隆中的发生率。

[0220] 类似地，根据对非靶分子(例如肝素或避免非特异性结合到的其它分子)的结合进行的选择能够与根据FcRn结合进行的选择在本文所述的方法中组合。

[0221] 在一些情形中，作为可开发性的某些方面的互相关联的结果，针对一种特征进行的选择(例如针对表面呈现水平高进行的选择)将富集具有多种有益质量的结合剂。举例来说，一些克隆可以表达与非靶分子(例如蛋白多糖)在细胞表面上相互作用的多反应性结合剂，例如他们可以非特异性结合到肝素。在高等真核细胞展示系统中，表达结合剂的细胞克隆本身可以表达相同或相似的非靶分子(例如细胞内源的蛋白多糖)，所产生的效应是表面展示的结合剂与细胞上所展示的一种或多种分子在细胞上发生非特异性相互作用。这会引起结合剂在细胞内内化，从而展现较低的表面呈现水平，从而取消选择表达其的克隆。在这种情况下，选择展示结合剂的较高表面呈现水平的克隆将并行富集或选择表达在多个方面具有更佳可开发性质量(例如较大可溶性和较小倾向于非特异性结合)的结合剂的克隆。

[0222] 其不同实例是，通过富集自聚集倾向较低的结合剂，方法可以通过从选择中排除倾向于形成免疫原性聚集体的那些结合剂来促进选择体内免疫原性程度较低的结合剂(如果用于医药配制物)。

[0223] 除此之外，如所提及，通过包括其它经标记的检测试剂(例如经标记的FcRn、经标记的靶标、经标记的其它多反应性探针)能够合并其它尺度的并行选择。一系列不同标记(例如不同波长的荧光团)的随时可用性，和FACS机器执行多重检测和分选的能力能够有助于此类并行选择的设计。

#### 亲和力选择

[0225] 选择针对靶标的结合剂时，能够基于亲和力选择结合剂通常是适用的，其允许富集表达高亲和力结合靶标的结合剂的克隆。

[0226] 亲和力通常用 $K_D$ (平衡解离常数)表示。 $K_D$ 是结合剂与其靶标(例如抗体与其抗原之间)之间相互作用的 $k$ (解离)/ $k$ (结合)比率。 $K_D$ 数值涉及结合剂的浓度并且因此， $K_D$ 数值越低(浓度越低)，则结合剂的亲和力越高。以抗体识别其抗原的方式特异性识别其靶标的结合剂可以称为同源结合剂。同源结合剂识别靶标宜是高亲和力相互作用。结合剂：靶标相互作用的 $K_D$ 可以小于 $1\mu M$ ，优选小于 $10nM$ 。

[0227] 本发明的方法可以包括从细胞群中富集编码(和呈现)针对靶标的较高亲和力结合剂的那些细胞。通过使用其上结合剂以相对较低拷贝数呈现的真核细胞能够增强选择的

严格度,以驱动亲和力选择。选择结合到所关注靶标的结合剂的方法可以利用各含有编码结合剂的DNA的高等真核细胞克隆的文库,其中结合剂呈现于细胞表面上,其中所编码的结合剂从弱活性启动子表达且/或在细胞表面上以相对较低的拷贝数表达。这可以来自弱活性启动子驱动的表达或是转录物不稳定性或非最优剪接、翻译、输送到表面或滞留于细胞表面上的结果。

[0228] 提供高等真核细胞的浓稠悬浮液(例如 $10^7/\text{ml}$ ,其具有高水平的表面呈现(例如 $6 \times 10^5$ 个结合位点/细胞呈现相对较高浓度的抗体(在此实例中为 $10\text{nM}$ ))。在所述情形中,即使当使用低浓度抗原试图驱动选择严格度(例如 $0.1\text{nM}$ )时,抗体的高浓度也将驱动结合,此限制了高亲和力克隆与低亲和力克隆之间的相对富集(参见实例8a和实例8b)。优选 $1\text{nM}$ 或更小的结合剂输入浓度。即使细胞密度较低,在高密度的已固定结合剂存在下仍存在靶标再结合的潜在问题。这个问题在基于表面的亲和力测量(例如表面等离子体共振(BIAcore手册))中已得到特别充分的认知和记录,并且具有使高亲和力克隆与低亲和力克隆之间的差异性结合减少的作用。因此,拷贝数可以在例如每个细胞 $100\text{-}100,000$ 范围内。在优选实施例中,拷贝数是每个细胞约 $60,000$ 或更少,任选地是每个细胞约 $50,000$ 或更少,或每个细胞约 $40,000$ 或更少。为了促进检测,拷贝数可以是每个细胞至少 $100$ 、至少 $1,000$ 或至少 $10,000$ 。因此,拷贝数可以在例如每个细胞 $1,000\text{-}60,000$ 范围内。拷贝数和在表面呈现结合剂的背景下测定拷贝数的方法在本文中别处论述。

[0229] 文库暴露于靶标(例如通过将靶标添加到文库中的表达结合剂的细胞的悬浮液中),使结合剂与靶标接触,从而允许同源结合剂(如果存在)识别靶标。展示同源结合剂的细胞结合到靶标。通过使用浓度有限的靶标,较高亲和力结合剂优先与靶标结合。与展示较高亲和力结合剂的细胞相比,展示不识别靶标或以较低亲和力识别靶标的结合剂的细胞将不结合靶标或每个细胞展示较少的靶分子。然后能够分离出结合到靶标的细胞,借此选择富集展示同源结合剂的细胞的细胞群。

[0230] 使用浓度降低的靶标能够任选地重复选择程序,以逐渐地增加选择严格度以及增加较高亲和力克隆的富集度。在进一步富集高亲和力结合剂之前,可以将突变引入所选群体的结合剂中以产生变异体。衍生物文库的产生在本文中别处描述。

[0231] 所用靶标浓度可以低于所述方法设法分离的结合剂的 $K_d$ 相互作用。

[0232] 然后能够选择对靶标具有所期望亲和力的一种或多种克隆,并且任选地,然后能够回收编码DNA且由个别培养的重组细胞表达结合剂,如在本文中别处所述。

[0233] 为了在细胞上实现低水平的表面呈现,结合剂基因可以在低水平表达,例如可操作地连接到弱活性启动子或是转录物不稳定性或非最优剪接、翻译、输送到表面或滞留于细胞表面上的结果。诱导型启动子和其它可控表达系统详细地描述本文中别处。参见例如实例8a或实例8b。

[0234] 文库

[0235] 各含有编码结合剂的重组DNA的细胞克隆集合同形成文库。文库多样性是克隆所编码的不同结合剂的数目的函数。在药物发现中,有利的是提供大型多样化文库,以使鉴定出满足所有期望标准的结合剂的潜在性最大化。文库中的每种克隆可以通过将编码结合剂的DNA整合到细胞DNA中形成重组细胞来产生,如在本文中别处所述。可以将DNA并行引入许多细胞中以产生重组细胞群,其各自编码来自多样化谱系的至少一种结合剂。将供体DNA

整合到细胞DNA中之后,培养所得重组细胞以允许其复制,从而由初始产生的各种重组细胞产生细胞克隆。因此,每个克隆来源于一种其中整合(例如在通过位点特异性核酸酶或通过如本文所述的其它方法建立的整合位点处)有供体DNA的原始细胞。根据本发明的文库可以含有至少 $100$ 、 $10^3$ 、 $10^4$ 、 $10^5$ 、 $10^6$ 、 $10^7$ 、 $10^8$ 、 $10^9$ 或 $10^{10}$ 个克隆。

[0236] 根据本发明的文库可以具有以下特点中的任何一种或多种:

[0237] 多样性.文库可以编码且/或表达至少 $100$ 、 $10^3$ 、 $10^4$ 、 $10^5$ 、 $10^6$ 、 $10^7$ 、 $10^8$ 或 $10^9$ 种不同结合剂。不同序列的结合剂构成谱系。

[0238] 均一整合.文库可以由克隆组成,所述克隆含有在固定基因座整合的供体DNA或在细胞DNA中的有限数目个固定基因座整合的供体DNA。因此,文库中的每种克隆优选地含有位于固定基因座处或固定基因座中的至少一个处的供体DNA。优选地,克隆含有在细胞DNA中的一或两个固定基因座整合的供体DNA。如在本文中别处所解释,整合位点可以存在于位点特异性核酸酶的识别序列处。整合供体DNA以产生重组DNA详细地描述于本文中别处并且能够根据整合位点的数目产生不同结果。在用于产生文库的细胞存在单一潜在整合位点的情况下,文库将是克隆的文库,所述克隆含有在单一固定基因座整合的供体DNA。文库中的所有克隆因此在细胞DNA中的相同位置含有结合剂基因。或者,在存在多个潜在整合位点的情况下,文库可以是克隆文库,其含有在多个和/或不同固定基因座整合的供体DNA。优选地,文库中的每种克隆含有在第一和/或第二固定基因座整合的供体DNA。举例来说,文库可以包括其中供体DNA整合于第一固定基因座的克隆、其中供体DNA整合于第二固定基因座的克隆,和其中供体DNA整合于第一与第二固定基因座的克隆。在优选实施例中,文库的克隆中仅存在一个或两个固定基因座,但是在特定应用需要时可以将供体DNA整合于多个基因座。因此,在一些文库中,每种克隆可以含有在若干个固定基因座中的任何一个或多个处(例如三个、四个、五个或六个固定基因座)整合的供体DNA。

[0239] 对于含有在各别位点整合的结合剂亚单元的文库来说,所述文库的克隆可以含有在第一固定基因座整合的编码第一粘合剂亚单元的DNA和在第二固定基因座整合的编码第二结合剂亚单元的DNA,其中所述克隆表达包括第一与第二亚单元的多聚体结合剂。

[0240] 均一转录.在文库的不同克隆之间,结合剂转录的相对水平保持在可控的限值内,原因是供体DNA整合于可控数目个基因座以及不同克隆中的相同基因座(固定基因座)。结合剂基因的相对均一转录使得结合剂在文库中的克隆上产生相似水平的表达或文库中的克隆表达结合剂的水平相似。文库中的细胞表面上所展示的结合剂可以与相同细胞上所展示的其它结合剂一致(具有相同氨基酸序列)。文库可以由以下组成:各自展示结合剂谱系中的单个成员的细胞克隆,或每个细胞展示结合剂谱系中的多个成员的克隆。或者,文库可以包括:展示结合剂谱系中的单个成员的一些克隆,和展示结合剂谱系中的多个成员(例如两个)的一些克隆。优选地,文库的克隆表达结合剂谱系中的一或两个成员。

[0241] 举例来说,根据本发明的真核细胞克隆的文库可以表达至少 $10^3$ 、 $10^4$ 、 $10^5$ 、 $10^6$ 、 $10^7$ 、 $10^8$ 或 $10^9$ 种不同结合剂的谱系,例如IgG、Fab、scFv或scFv-Fc抗体片段,每个细胞含有在细胞DNA中整合的供体DNA。供体DNA编码结合剂且可以进一步包括用于选择其中整合有供体DNA的细胞的基因元件。文库中的细胞可以含有编码外源位点特异性核酸酶的DNA。

[0242] 文库中的细胞表面上所展示的结合剂可以与相同细胞上所展示的其它结合剂一致(具有相同氨基酸序列)。文库可以由以下组成:各自展示结合剂谱系中的单个成员的细

胞克隆,或每个细胞展示结合剂谱系中的多个成员的克隆。或者,文库可以包括:展示结合剂谱系中的单个成员的一些克隆,和展示结合剂谱系中的多个成员(例如两个)的一些克隆。相应地,根据本发明的文库可以包括编码结合剂谱系中的超过一个成员的克隆,其中供体DNA整合于双重复固定基因座或多个独立固定基因座。

[0243] 如果相应克隆仅表达一种结合剂,则最容易鉴定结合剂的相应编码DNA。典型地,供体DNA分子将编码单一结合剂。结合剂可以是多聚体,因此供体DNA分子包括与多聚体结合剂的不同亚单元对应的多个基因或开放阅读框架。

[0244] 如所提及,根据本发明的文库可以编码至少 $100$ 、 $10^3$ 、 $10^4$ 、 $10^5$ 或 $10^6$ 、 $10^7$ 、 $10^8$ 、 $10^9$ 或 $10^{10}$ 种不同结合剂。在结合剂是多聚体的情况下,多样性可以由结合剂的一个或多个亚单元提供。多聚体结合剂可以是一个或多个可变亚单元与一个或多个恒定亚单元的组合,其中恒定亚单元在文库的所有克隆中是相同的(或多样性较有限)。在产生多聚体结合剂文库时,组合多样性是可能的,其中可以将第一谱系的结合剂亚单元与第二谱系的结合剂亚单元中的任一种配对。

[0245] 本文中别处进一步描述了根据本发明的文库的这些和其它特点,并且适用于其构建和用途的文库和方法的实例也阐述于W02015/166272(Iontas Limited)中,所述文献的内容并入本文中供参考。可以鉴定出适用于将编码DNA靶向整合于细胞染色体中的基因座,并且若干个实例在所属领域中已知。AAVS基因座可以如本文中所说明来使用。其它适合的整合位点包括ROSA26、HPRT和FUT8基因座(例如在CHO细胞内)。虽然靶向整合可以具有优势,但是随机整合是一种适合的替代方案并且可以用于许多情形中。除核酸酶介导的靶向整合之外,用于产生表面表达的多肽文库的方法还包括用编码多肽的载体转染真核细胞,例如使用慢病毒、腺病毒、腺相关病毒或转座子,或使用基因组嵌入式重组酶位点,例如F1p、Bxb2重组酶,或内源性隐蔽重组酶位点,例如φ重组酶位点。通过这些或任何其它技术建立的文库可以在本文所述的方法中使用。本发明扩展到纯形式的文库,其为缺乏其它真核细胞的文库克隆群体;或与其它真核细胞混合的文库。其它细胞可以是相同类型(例如相同细胞系)的真核细胞,或不同细胞。其它优势可以通过将根据本发明的两个或更多个文库组合或将根据本发明的文库与细胞的第二文库或第二群体组合来获得,以促进或拓宽筛选,或用于其它用途,如本文中所述或熟练的技术人员显而易知。

[0246] 根据本发明的文库、从所述文库获得的一种或多种克隆,或编码来自文库的结合剂的DNA已引入其中的宿主细胞可以在细胞培养基中提供。可以培养细胞并且然后浓缩以形成用于方便运输或储存的细胞集结粒。

[0247] 文库通常在体外提供。文库可以存在于容器中,例如含有悬浮于培养基中的文库细胞的细胞培养瓶,或包含真核细胞(包括文库)的集结粒或浓缩悬浮液的容器。文库可以构成容器中的真核细胞的至少75%、80%、85%或90%。选择步骤可以针对混合培养物中的文库执行,从而促进高通量。

[0248] 作为协同培养文库的克隆的混合物的一个替代方案,个别地培养克隆有时可能是方便的,所述克隆各自独立地存在于各别烧瓶或其它容器中。可以使用个别培养物,其中对数目相对较少的克隆进行比较,例如其中亲本结合剂已经发生突变而产生一种或多种变异体(例如多达10种变异体),并且利用本发明比较变异体的质量与亲本结合剂的质量和/或比较彼此间的质量。

[0249] 本文所述的选择方法可以应用于初始文库,即,尚未经历基于亲和力的选择的文库。因此,在执行任何基于亲和力的选择之前,可以利用方法使文库富集具有较高溶解度和/或较低非特异性结合的克隆(反之,使文库中的展现低溶解度和/或较高非特异性结合的克隆耗竭)。已经根据较大可开发性克隆“预选”的此类文库然后高度适于使用所关注的靶标执行基于亲和力的选择。与从尚未根据可开发性预筛选的文库选出的克隆相比,亲和力选择步骤获得的克隆更可能展现良好溶解度、高临界浓度、低非特异性结合和/或其它可开发性质量。

[0250] 真核细胞

[0251] 根据本发明的真核细胞优选高等真核细胞,其在此定义为基因组大于酿酒酵母基因组的细胞,酿酒酵母具有 $12 \times 10^6$ 个碱基对(bp)的基因组尺寸。高等真核细胞可以具有例如大于 $2 \times 10^7$ 个碱基对的基因组尺寸。这包括例如哺乳动物、禽鸟、昆虫或植物细胞。本发明的真核细胞优选地缺乏细胞壁。优选地,他们不是酵母细胞或其它真菌细胞。所述细胞优选哺乳动物细胞,例如小鼠或人类。所述细胞可以是原代细胞或可以是细胞系。中国仓鼠卵巢(CHO)细胞通常用于抗体和蛋白质表达,但本发明可以使用任何替代的稳定细胞系。本文中的若干实例使用HEK293细胞。

[0252] 由于制造用于研究、诊断和治疗应用的抗体典型地在这些细胞中进行,因此使用高等真核细胞(例如哺乳动物细胞)呈现结合剂具有优势。对展现相同表达环境和相同翻译后修饰的哺乳动物细胞实施药物发现将更好地给下游制造指出潜在问题或好处,从而允许早期鉴定出具有最优表达特性的克隆。

[0253] 相比之下,细菌和酵母细胞不能充分地再现高等真核细胞的糖基化、表达和分泌机制。因此,在哺乳动物细胞上展示能够有助于鉴定出具有更佳呈现水平或稳定性特性、影响未来研究使用或下游制造的克隆。能够在哺乳动物细胞表面上展示大型抗体文库将允许根据结合和呈现特性直接筛选数百万个克隆,所述克隆具有在抗体开发的发现阶段期间使用制造细胞系的潜能。

[0254] CHO细胞系最初在1957年分离<sup>63</sup>且这种细胞系的衍生物已经成为大部分治疗抗体的生产细胞系<sup>64</sup>。举例来说,通过CHO细胞表达生产的赫赛汀(Herceptin)(已批准用于治疗乳癌的抗HER-2抗体)每年超过一公吨。与人类细胞相比,CHO细胞因其不传播大部分人类病原性病毒而具有用于生产供施用于人类的产品的优势。另外,他们允许外来DNA整合到其基因组中且快速稳定地生长。抗体和其它多肽结合剂的特性,包括生物物理学特性、稳定性药物动力学和免疫原性,能够受到其翻译后修饰的影响,例如糖基化<sup>65</sup>,所述翻译后修饰是在细胞的分泌路径内获取,例如内质网(ER)和高尔基体。在此,单糖单元,例如甘露糖(Man)、半乳糖(Gal)、岩藻糖(Fuc)、N-乙酰基葡萄糖胺(GlcNAc)、N-乙酰基半乳糖胺(GalNAc)和唾液酸,共价连接到特定氨基酸。“O连接型”或“N连接型”糖基化是指连接到丝氨酸或苏氨酸残基中的氧原子的聚糖或连接到天冬酰胺残基中的酰胺氮原子的聚糖。糖基化的复杂度能够通过线性或分支化的糖基化以及单糖单元内的不同位置处的糖苷键联的原子位置和构象 $\alpha$ 或 $\beta$ 引入。用于表达的宿主细胞系能够影响到抗体糖基化特征的确切本质<sup>65</sup>。因此,在抗体和治疗蛋白筛选期间,在尽可能接近最终生产细胞系的宿主细胞系中生产重组蛋白是一种优势。这确保了待筛选的多肽的翻译后修饰和因此其特性与在其大规模制造期间获得的那些相同或尽可能相似。由于大部分人类治疗抗体是在CHO细胞中产生,因此在CHO宿主细胞

系中进行高等真核展示是一种优势。实例12展现了对CHO细胞文库中的候选多肽药物进行基于可开发性的选择。

[0255] 在本发明的方法和用途中,一般来说,多种细胞克隆可以是至少1000种克隆的文库,其任选地在单个容器内、在相同培养基中一起培养。其可以是初始文库,即,编码结合剂谱系的文库,所述结合剂谱系尚未预先根据对靶标的结合而经历选择。在早期发现中,情况通常如此,但使用编码结合剂谱系的文库可为所期望的,所述结合剂谱系是预先根据对靶标的结合进行一轮或多轮选择的结果。举例来说,可以将噬菌体展示文库的选择结果引入真核细胞文库中。

[0256] 本发明在本说明书中具体参照哺乳动物细胞描述,并且在实例中使用哺乳动物细胞说明本发明。然而,除非上下文另外要求,否则应理解,可以改用其它高等真核细胞。举例来说,可以使用昆虫或鸡细胞。

[0257] **结合剂**

[0258] 根据本发明的“结合剂”是一种结合分子,其代表着另一种分子的特异性结合搭配物。特异性结合搭配物的典型实例是抗体-抗原和受体-配体。本发明的许多原理扩展到在传统上可以不视为“结合剂”的多肽,例如酶、辅因子、凝血因子和其抑制剂、补体因子和其抑制剂等。在许多情况下,此类多肽代表着希望以大规模生产和/或以高浓度提供的候选临床药物。其可开发性质量因此是重要的考虑因素。举例来说,因子VIII用于血友病,但具有相当大的可开发性问题。本发明的与可开发性有关的方面能够应用于所有此类多肽。因此,除非上下文另外要求,否则本发明不应解释为受限于经典的特异性结合分子,如抗体,并且应理解为通常扩展到设计成与一种或多种其它分子相互作用的多肽。

[0259] 本发明涉及作为多肽的结合剂,即,氨基酸的聚合物,其由编码DNA在细胞中表达并且任选地经历翻译后修饰,例如裂解、糖基化等。结合剂可以包括例如大约10到30个氨基酸的短肽。结合剂还可以是较长多肽并且可以任选地包括多个亚单元。

[0260] 结合剂优选地包括哺乳动物(例如人类)起源的多肽。结合剂可以包括人类抗体(任选地为包括人类抗体的嵌合抗原受体(CAR))、人类TCR或其它受体,或其它人类多肽。结合剂可以是可溶性肽或多肽(例如细胞因子、趋化因子、补体蛋白质或补体调控剂、酶(包括用于工业用途的酶)、凝血因子(例如因子VIII))。结合剂可以是哺乳动物(例如人类)膜蛋白,或可以是已经工程改造以包括一个或多个TM域或其它膜锚的哺乳动物(例如人类)可溶性蛋白质/肽。结合剂可以是原生的、天然存在的多肽,或(通常)是合成变异体。举例来说,人类因子VIII多肽的文库可以包括与人类因子VIII具有至少70%氨基酸序列一致性(例如至少80%或至少90%)的结合剂。

[0261] 由文库编码的结合剂谱系通常享有共同结构且具有一个或多个多样性区域。所述文库因此允许选择具有所期望的分子结构类别的成员,例如肽或scFv抗体分子。在根据本发明的结合剂的文库或群体中,结合剂多肽可以因此享有共同结构(例如相关的二级和/或三级结构,任选地包括高度相似或一致氨基酸序列的区域-“恒定区”)并且可以具有一个或多个氨基酸序列多样性区域-“可变区”。

[0262] 考虑抗体谱系能够对此进行说明。这些可以是具有共同结构类别的抗体,例如IgG、Fab、scFv-Fc或scFv,其序列的一个或多个区域不同。抗体的互补决定区(CDR)典型地具有序列可变性,互补决定区是主要涉及抗原识别的区域。本发明中的结合剂谱系可以是

一个或多个CDR不同的抗体分子的谱系,例如所有六个CDR中或一个或多个特定CDR(例如重链CDR3和/或轻链CDR3)中可存在序列多样性。

[0263] 本文中别处更详细地描述了抗体和其它结合剂。然而,本发明的潜力扩展超出抗体展示而包括肽或工程化蛋白质(包括受体、配体、个别蛋白质结构域和替代蛋白质支架)的文库展示<sup>66-68</sup>。实例是具有单体结合域的多肽,例如锚蛋白重复蛋白(DARPins)和脂笼蛋白(lipocalins)、亲和抗体(affibodies)和肽呈现支架蛋白(adhIrons)。本发明还能够结合复杂的多聚体结合剂使用。举例来说,T细胞受体(TCR)在T细胞上表达并且已经进化而识别在抗原呈现细胞上与MHC分子复合呈现的肽。可以产生编码且表达TCR谱系的文库,并且可以筛选以鉴定对MHC肽复合物的结合。

[0264] 对于多聚体结合剂来说,编码结合剂的供体DNA可以作为一个或多个DNA分子提供。举例来说,在个别抗体VH和VL域分开表达的情况下,这些域可以在供体DNA的单独分子上编码。供体DNA在多个整合位点整合到细胞DNA中,例如VH的结合基因在一个基因座整合,而VL的结合基因在第二基因座整合。引入编码单独结合剂亚单元的供体DNA的方法更详细地描述于本文中别处以及W02015/166272(Iontas Limited),所述文献并入本文中供参考。或者,多聚体结合剂的亚单元或一部分均可以在供体DNA的同一分子上编码,所述供体DNA在固定基因座整合。

[0265] 结合剂可以是包含抗原结合位点的抗体或非抗体蛋白质。抗原结合位点可以通过将肽环布置于非抗体蛋白质支架(例如纤维结合蛋白或细胞色素B等)上或通过使蛋白质支架内的环氨基酸残基随机化或突变以赋予针对所期望靶标的结合来提供。(Haan及Maggos(2004)BioCentury,12(5):A1-A6<sup>69,70</sup>,。用于抗体模拟物的蛋白质支架公开于W0/0034784中,其中发明人描述了包括具有至少一个随机化环的纤维结合蛋白III型域的蛋白质(抗体模拟物)。一个或多个肽环(例如一组抗体VH CDR环)所植入的适合支架可以由免疫球蛋白基因超家族的任何结构域成员提供。支架可以是人类或非人类蛋白质。

[0266] 抗原结合位点在非抗体蛋白质支架中的使用先前已有评述(Wess,L.于:《生物世纪:伯恩斯坦关于生物商业的报告(BioCentury,The Bernstein Report on BioBusiness)》,12(42),A1-A7,2004)。典型的是具有稳定主链和一个或多个可变环的蛋白质,其中所述一个或多个环的氨基酸序列发生特定或随机突变以建立用于结合靶抗原的抗原结合位点。此类蛋白质包括来自金黄色葡萄球菌的蛋白质A的IgG结合域、运铁蛋白、四连接素(tetranectin)、纤维结合蛋白(例如第10个纤维结合蛋白III型域)和脂笼蛋白。其它方案包括限制性小肽(例如基于“打结素(knottin)”)和环肽支架<sup>71</sup>。鉴于其小尺寸和复杂性(尤其关于二硫键的正确形成),因此使用真核细胞选择基于这些支架的新颖结合剂可能存在优势。鉴于这些肽在自然界中的共同功能,因此基于这些支架的结合剂的文库可能有利于产生特别适用于阻断离子通道和蛋白酶的高亲和力小结合剂。W02017/118761(Iontas Limited)描述了各自包含融合蛋白的结合成员的文库,所述融合蛋白含有插入受体多样性支架域(例如抗体恒定或可变域)内的供体多样性支架域,例如富含半胱氨酸的蛋白质。如W02017/118761中所述的此类结合剂和文库可以用于本发明并且所述文献并入本文供参考。因此,在一些实施例中,根据本发明的结合剂是“打结抗体”,其包括插入抗体可变域内的富半胱氨酸蛋白质。参见本文中的实例15。

[0267] 除抗体序列和/或抗原结合位点之外,结合剂还可以包括其它氨基酸,例如形成肽

或多肽,例如折叠域,或赋予分子除结合抗原能力之外的另一种功能特征。结合剂可以携带可检测标记,或可以与毒素或靶向部分或酶偶联(例如经由肽键或接头)。举例来说,结合剂可以包括催化位点(例如酶域中的催化位点)以及抗原结合位点,其中所述抗原结合位点结合到抗原并且从而使催化位点靶向抗原。催化位点可以抑制抗原的生物功能,例如通过裂解。

[0268] 任选地,文库中的结合剂是多肽群体,其热稳定性(例如根据解链温度所测定)不能预测结合剂在溶液中的溶解度或自缔合抗性。当考虑文库中的具有至少10mg/ml的溶解度或临界浓度(测定方法在本文中别处描述)的所有结合剂时,可以应用热稳定性与溶液中的溶解度/自缔合抗性之间的这种无关。当考虑表面呈现为每个细胞至少 $10^3$ 个、至少 $10^4$ 个、每个细胞至少 $10^5$ 个或每个细胞至少 $10^6$ 个的所有结合剂(测定方法描述于本文中别处,例如FACS设门)时,其可为适用的。任选地,当考虑文库中的表面表达结合剂的总群体时,其可为适用的。

[0269] 抗体

[0270] 抗体是优选的结合剂。它们可能是完整的抗体或免疫球蛋白(Ig),具有四个多肽链-两条相同的重链和两条相同的轻链。重链和轻链形成对,每个对具有包含抗原结合位点的VH-VL域对。重链和轻链还包含恒定区:轻链CL以及重链CH1、CH2、CH3和有时CH4。两条重链在柔性铰链区通过二硫键连接。抗体分子可以包含VH和/或VL域。

[0271] 抗体分子的最常见原生形式是IgG,其为由两条相同重链和两条相同轻链组成的异四聚体。重链和轻链是由具有保守二级结构的模块化结构域构成,所述模块化结构域由以下组成:四链逆平行β片层和三链逆平行β片层,其通过单个二硫键稳定。抗体重链各自具有一个N端可变域(VH)和3个相对保守的“恒定”免疫球蛋白域(CH1、CH2、CH3),而轻链具有一个N端可变域(VL)和一个恒定域(CL)。二硫键使个别结构域稳定且形成共价键以将四条链接合成稳定复合物。轻链的VL和CL与重链的VH和CH1缔合,并且这些元件能够单独表达以形成Fab片段。CH2和CH3域(也称为“Fc域”)与另一对CH2:CH3对缔合,得到四聚体Y形分子,其中来自重链和轻链的可变域位于“Y”的尖端。CH2和CH3域负责与免疫系统中的效应细胞以及补体组分发生相互作用。重组抗体先前已以IgG形式或Fab(由VH:CH1的二聚体和轻链组成)表达。此外,可以使用称为单链Fv(scFv)的人工构建体,其由编码VH和VL片段的DNA与编码柔性接头的DNA以基因方式融合而组成。

[0272] IgG是用于治疗用途的优选抗体类别之一。使用高等真核细胞(尤其是哺乳动物细胞)的一个优势是,它可以与IgG形式的抗体一起发挥作用。这使得药物发现和筛选能够直接在用于制造IgG的生产细胞类型中进行。

[0273] 结合剂可以是人抗体分子。因此,在存在恒定域的情况下,这些优选人恒定域。

[0274] 结合剂可以是抗体片段或较小的抗体分子形式,例如单链抗体分子。例如,抗体分子可以是scFv分子,其由通过接头肽连接的VH域和VL域组成。在scFv分子中,VH和VL域形成VH-VL对,其中VH和VL的互补决定区共同形成抗原结合位点。

[0275] 包含抗体抗原结合位点分其它抗体片段包括(但不限于)(i)Fab片段,其由VL、VH、CL和CH1域组成;(ii)Fd片段,其由VH和CH1域组成;(iii)Fv片段,其由单一抗体的VL和VH域组成;(iv)dAb片段<sup>72-74</sup>,其由VH或VL域组成;(v)分离的CDR区;(vi)F(ab')2片段,一种包含两个相连的Fab片段的二价片段;(vii)scFv,其中VH域与VL域通过肽接头连接,所述肽接头

允许两个结构域缔合而形成抗原结合位点<sup>75,76</sup>; (viii) 双特异性单链Fv二聚体(PCT/US92/09965)和(ix)“双功能抗体”，通过基因融合构建的多价或多特异性片段(WO94/13804<sup>77</sup>)。Fv、scFv或双功能抗体分子可以通过并入连接VH和VL域的二硫桥来稳定化<sup>78</sup>。

[0276] 包括一个或多个抗体抗原结合位点的各种其它抗体分子已经工程改造，包括例如Fab<sub>2</sub>、Fab<sub>3</sub>、双链抗体、三链抗体、四链抗体和微型抗体(小免疫蛋白质)。抗体分子以及其构建和使用方法已有描述<sup>79</sup>。

[0277] 结合片段的其它实例是Fab'，其与Fab片段不同之处是重链CH1域的羧基端添加了几个残基，包括来自抗体铰链区的一个或多个半胱氨酸；并且不同于Fab' -SH，其是恒定域的半胱氨酸残基具有游离巯基的Fab'片段。

[0278] dAb(域抗体)是抗体的小单体抗原结合片段，即，抗体重链或轻链的可变区。VH dAb天然地存在于骆驼科(例如骆驼、大羊驼)中并且可以通过用靶抗原使骆驼科免疫、分离抗原特异性B细胞且直接克隆来自个别B细胞的dAb基因来产生。dAb也能在细胞培养中产生。其小尺寸、良好溶解度和温度稳定性使得其在生理学上特别适用且适于选择和亲和力成熟。用于治疗用途的骆驼科VH dAb正以名称“纳米抗体™”开发。

[0279] 通过从借助于在适合的表达载体内合成和组装寡核苷酸而产生的基因表达可以建立合成抗体分子，例如Knappik等人或Krebs等人所述<sup>80,81</sup>。

[0280] 双特异性或双功能抗体形成第二代单克隆抗体，其中两个不同可变区组合成同一分子<sup>63</sup>。其用途在诊断领域和治疗领域中已凭其募集新效应功能或使若干分子靶向肿瘤细胞表面的能力得到证明。在使用双特异性抗体的情况下，这些抗体可以是传统双特异性抗体，其能够以多种方式制造，例如以化学方式或由杂交融合瘤制备，或可以是上文所提及的双特异性抗体片段中的任一个<sup>82</sup>。这些抗体能够通过化学方法<sup>83,84</sup>或体细胞方法<sup>85,86</sup>获得，但类似地且优选地通过基因工程技术获得，所述基因工程技术迫使异二聚发生并且因此促进了所寻求抗体的提纯过程<sup>87</sup>。双特异性抗体实例包括BiTE™技术的那些抗体，其中能够使用具有不同特异性的两种抗体的结合域且其通过短柔性肽直接连接。这是两种抗体在单个短多肽链上的组合。双链抗体和scFv能够在没有Fc区的情况下仅使用可变域构建而成，潜在地减少了抗个体基因型反应的影响。

[0281] 在本发明的一些实施例中，结合剂是双特异性抗体并且其编码克隆包含两条不同抗体重链和任选的两条不同抗体轻链或优选一条共同轻链。重链之间成功的异二聚对引起细胞表面呈现双特异性抗体，各抗体包含重链的异二聚对，并且任选地，每条重链与一条轻链成对，任选地与一条共同轻链成对(即，轻链与具有相同氨基酸序列的每条重链成对)。在包括Fc区的情况下，本发明可以用于评估可能会改善异二聚的Fc序列变体。在此前的研究中，在Fc区已经工程改造以改善异二聚的情况下，可开发性特征已因变化而受损<sup>67</sup>。本发明提供了根据可开发性(例如溶解度和多特异性)以及其异二聚潜力筛选此类变体的机会。

[0282] 双特异性抗体能够作为完整IgG、作为双特异性Fab' 2、作为Fab' PEG、作为双链抗体或作为双特异性scFv构建而成。另外，使用所属领域中已知的常规方法能将两种双特异性抗体连接以形成四价抗体。

[0283] 与双特异性全抗体相反，双特异性双链抗体也可以是特别适用的。能够容易地选择具有适当结合特异性的双链抗体(和许多其它多肽，例如抗体片段)。如果双功能抗体的

一个臂欲保持恒定(例如针对所关注抗原的特异性),则能够制备其中另一臂改变的文库且选择具有适当特异性的抗体。双特异性全抗体可以通过如以下文献中所述的替代工程化方法制备:Ridgeway等人(《蛋白质工程(Protein Eng.)》,9,616-621,(1996))。

[0284] 双特异性抗体的替代形式包括包含免疫球蛋白的mAb<sup>2</sup>(“mAb平方”)分子,所述免疫球蛋白的CH3环区已经工程改造以提供抗原结合位点(参见例如WO2006072620、WO2008003103、WO2008003116)。经修饰的CH3区域称为Fcab。Fv区的抗体抗原结合位点提供了一种结合特异性,而Fcab中的结合位点提供了不同结合特异性(或其它价数)。

[0285] 根据本发明的文库可以用于选择结合所关注的一种或多种抗原的抗体。从文库中选择详细地描述于本文中别处。在选择之后,抗体可以经工程改造成不同形式和/或含有其它特点。举例来说,可以将所选抗体转化成不同形式,例如上述抗体形式之一。所选抗体和包含所选抗体分子的VH和/或VL CDR的抗体是本发明的一个方面。抗体和其编码核酸可以分离的形式提供。

[0286] 以抗体分子为起始物,利用例如酶(例如胃蛋白酶或木瓜蛋白酶)消化等方法和/或通过化学还原使二硫桥裂解能够获得抗体片段。通过所属领域的技术人员众所周知的基因重组技术或通过肽合成、借助于例如自动肽合成仪或通过核酸合成和表达能够以另一方式获得抗体片段。

[0287] 可以采用单克隆抗体和其它抗体以及使用重组DNA技术中的技艺生产出结合靶抗原的其它抗体或嵌合分子。此类技术可能会涉及将编码抗体的免疫球蛋白可变区或CDR的核酸(例如DNA)引入不同免疫球蛋白的恒定区,或恒定区和构架区。参见例如EP-A-184187、GB 2188638A或EP-A-239400,和大量后续文献。

[0288] 抗体分子可以选自文库并且然后加以修饰,例如通过化学修饰(例如聚乙二醇化)或通过并入脂质体能够延长抗体分子的体内半衰期。

[0289] 结合剂可以包括展现序列多样性、任选地在一个或多个互补决定区中展现序列多样性的抗体可变域。结合剂还可以或替代地包含任选地展现序列多样性的抗体恒定区或Fc区。Fc区的特点进一步论述于下文。

#### [0290] Fc区

[0291] 结合剂可以是包含Fc区的多肽。结合剂可以是包含Fc区的抗体、打结抗体或其它多肽,任选地为融合蛋白。Fc区可以是或可以包括结合剂的恒定区,即,具有高度相似或一致的氨基酸序列,如在文库的结合剂之间所比较。在一些情况下,抗体的恒定域(例如Fc区,或CL或CH1域)可以是或可以包括可变氨基酸序列,从而在文库中的结合剂谱系中展现序列多样性。序列多样性可以任选地存在于Fc的CH3域中。举例来说,结合剂可以包括Fcabs或mAb<sup>2</sup>,其中Fcabs的结合环在文库内具有多样性。mAb<sup>2</sup>可以包括多种多样的Fcab区域和Fv区域中的非变异或多种多样的抗原结合位点。结合剂的氨基酸序列多样性可以局限于Fc区,任选地局限于CH3域。

[0292] 使用本发明的展示文库,可以根据改变的功能且根据可开发性标准(任选地同时根据改变的功能和可开发性标准)筛选Fc域文库。

[0293] 已经采取多种工程化方案对Fc域与其相互作用搭配物的相互作用进行工程改造。举例来说,“臼包杵”方案修改了两个成对Fc序列,使得其在细胞中的共表达(或其在共培养的细胞中的表达)主要引起两个变异型Fc域之间形成异二聚体,从而有利于产生双特异性

抗体。不幸的是,此类突变会影响可开发性<sup>88</sup>。本发明可以用于评估Fc变异体(包括含有候选“臼包杵”突变的Fc域)的可开发性潜力。

[0294] Fc工程化还已经用于改变Fc  $\gamma$  受体的亲和力或特异性,例如建立Fc  $\gamma$  受体相互作用减少的“剔除式变异体”。

[0295] 抗体恒定域的突变和具有所期望结合质量的变异体的选择能够降低稳定性和可制造性。另外,本发明可以用于评估变异型Fc域的文库以富集具有改善的特性的那些变异型Fc域。

[0296] 出于正向或负向调节半衰期的目的,也已经进行修饰以增加或减小Fc与FcRn的相互作用<sup>89</sup>。举例来说,M252Y/S254T/T256E的突变三联体(所谓的“YTE突变”)已经增加IgG半衰期,从而延长半衰期且减少给药频率和成本。

[0297] 应认识到,突变引入Fc域也能对变异体的生物物理学特性产生有害的影响,引起聚集和不良可开发性。举例来说,Borrok等人(2017)<sup>90</sup>评述了影响与FcRn相互作用的突变和影响与其它Fc受体相互作用的突变。他们描述了一种抗体,所述抗体将增加半衰期的突变(M252Y、S254T、T256E,所谓的“YTE突变”)和减少CH2域与Fc  $\gamma$  受体相互作用的其它突变(L234F、L235E、P331S,所谓的TM突变体)组合<sup>90</sup>。与野生型相比,此TM-YTE突变体具有降低的热稳定性、较大的构象柔性、增加的自缔合、不良的溶解性和不良的聚集特征。通过选择候选突变,他们能够建立新的FQQ-YTE变异体(L234F/L235Q/K322Q/M252Y/S254T/T256E),所述变异体具有显著改善的构象和胶体稳定性,同时保持延长的半衰期且不具有抗体依赖性细胞介导的细胞毒性和补体依赖性细胞毒性活性。

[0298] 本发明提供了根据Fc域的改变的结合特性以及可开发性标准(例如自聚集或低特异性)同时筛选大量变异体的机会。因此,在本发明的各种实施例中,结合剂可以包括展现序列多样性(例如在CH3域的一个或多个氨基酸残基中展现序列多样性)的Fc区。Fc域外部的一个或多个可变区(例如抗体重链可变域和/或轻链可变域)中展现序列多样性且任选地具有未变异序列的恒定Fc域的结合剂也可以加以筛选以鉴定可变区序列对FcRn相互作用的影响(参见实例9)。在可变区不直接结合到受体的情况下,此类影响可以是间接的,例如影响分子的构象或以其它方式较远地影响结合。

[0299] 可以将根据多肽的FcRn结合特征进行的筛选整合于药物发现中。方法可以包括在序列变异体的文库内根据最优的pH依赖性FcRn相互作用进行选择。本文描述了鉴定或选择因与FcRn的相互作用的性质而使得体内半衰期延长或缩短的结合剂。

[0300] 根据本发明的方法可以包括:

[0301] 提供多种真核(例如哺乳动物)细胞克隆,其各自含有编码包含Fc域的结合剂的DNA,

[0302] 在体外,在使所述结合剂呈现于细胞表面上的条件下培养所述克隆,

[0303] 使所述克隆在低pH(例如约pH 6.0)或中性pH(例如约pH 7.4)暴露于FcRn受体,以允许Fc域识别FcRn,

[0304] 在较高pH(例如约pH 7.4)洗脱并且选择表达结合剂的一种或多种克隆,所述结合剂在约pH 6.0展现的亲和力结合高于在较高pH(例如约pH 7.4)展现的结合。

[0305] 所选克隆能够从洗脱部分获得。来自所洗脱的克隆的结合剂可以鉴定为在体内具有延长的半衰期。

[0306] 或者,如果期望在较高pH的结合得以保持,则可以收集在切换到较高pH之后滞留的结合剂,例如为了缩短半衰期或用于“扫描抗体”方案<sup>32</sup>。在此类情况下,所述方法可以包括:

[0307] 提供多种真核(例如哺乳动物)细胞克隆,其各自含有编码包含Fc域的结合剂的DNA,

[0308] 在体外,在使所述结合剂呈现于细胞表面上的条件下培养所述克隆,

[0309] 使所述克隆在低pH(例如约pH 6.0)暴露于FcRn受体,以允许Fc域识别FcRn,

[0310] 在较高pH(例如约pH 7.4)下洗涤并且选择表达结合剂的一种或多种克隆,所述结合剂在约pH 6.0展现的亲和力结合低于或类似于在较高pH(例如约pH 7.4)展现的结合。

[0311] 所选克隆能够从在较高pH洗涤的情况下未洗脱的滞留部分中获得。此类克隆能够在更极端的pH(高于pH 7.4或低于pH 6)洗脱。可以选择在较低pH(例如约pH 6.0)下进行较高亲和力结合,以便在选择期间,通过减小在所述pH使用的FcRn的浓度来增加选择的严格度。

[0312] 可以优先选择在两个pH之间展现较大亲和力差异的结合剂,原因是这些结合剂可以显示最大的半衰期延长。反之,如果寻求较短半衰期,则选择表达结合剂的克隆,所述结合剂在两个pH之间对FcRn的亲和力存在较小差异(或无显著差异)。所述方法因此可以适于根据较短或较长半衰期进行选择。从而在克隆群体中富集表达具有所期望的半衰期的结合剂的克隆。

[0313] 对生物素化FcRn的结合可以在pH 6.0进行以确保结合发生,然后用pH增加的缓冲液洗涤之后,收集洗脱的样品。或者,可以使用流式分选、根据荧光标记的滞留或损失进行选择。基于FcRn的亲和色谱方法联合pH梯度洗脱已被描述<sup>91</sup>用于表征抗体。此类方法可以用于细胞上所展示的抗体变异体的文库,其中可以收集具有所期望的pH依赖性结合特性的克隆并且回收抗体基因。

[0314] 本发明的一些实施例使用抗Fc检测试剂测定结合剂的表面呈现水平。此类试剂仍然可以联合具有Fc序列多样性的结合剂使用,条件是所述多样性不影响可检测试剂的结合,例如能够确保检测试剂结合Fc区时,其中结合剂享有共同序列。在替代实施例中,结合剂包含不展现序列多样性的Fc区。

[0315] 细胞培养和结合剂表达

[0316] 为了提供针对所关注的靶标和/或根据可开发性特征筛选选用的结合剂谱系,可以培养从编码DNA表达结合剂的文库。如所论述,结合剂可以含有跨膜域、膜锚,或可以与膜结合的搭配物分子结合用于细胞外展示。培养用于表达结合剂的细胞通常涉及在适合的培养基中、任选地在悬浮培养基中且在有助于细胞生长的温度(例如37摄氏度,对于哺乳动物细胞来说)下培育细胞。从编码DNA表达多肽结合剂是在启动子(和任选地,其它元件,例如增强子)的控制下起始,并且在一段时间之后开始观察到结合剂的表面呈现,并且在例如12小时内应该是可检测的,但结合剂呈现可能需要更长的持续时间(例如24小时或48小时)以在细胞表面上达到最终或平衡浓度。

[0317] 启动子

[0318] 在根据本发明的细胞中,编码结合剂的DNA可操作地连接到启动子以便表达。可以使用异源启动子,这意味着其不是自然界中与编码DNA有关的启动子,例如可以使编码抗体

的DNA可操作地连接到除来自免疫球蛋白基因座的启动子之外的启动子。通常方便的是将启动子整合到细胞DNA中,任选地以顺式整合(与编码结合剂的序列位于同一供体DNA上),但替代方式是从宿主细胞内源的启动子表达所整合的结合剂DNA。

[0319] 在编码结合剂的DNA的表达处于强启动子(例如已经最大限度地诱导表达的组成型启动子或诱导型启动子)控制下的情况下,可以实现高水平的结合剂呈现。反之,在编码结合剂的DNA的表达处于弱活性启动子(例如已经最低限度地诱导表达或仅展现基本活性水平的弱启动子或诱导型启动子)控制下的情况下,可以实现低水平的结合剂呈现。启动子的强度能够使用报道基因量化并且测定报道体的表达水平(例如GFP的表达)能够作为可以量化的荧光检测。与CMV启动子相比,弱活性或弱启动子可以显示例如全部活性或组成型启动子的活性的约1-10%。

[0320] 控制细胞表面呈现结合剂的方便方法是提供在细胞中可操作地连接到编码结合剂的DNA的诱导型启动子。四环素诱导型启动子是适合的,并且多种这些启动子是可获得的。Gossen等人描述了第一代rtTA蛋白质(EP0804565和Gossen等人,1995<sup>92</sup>)。将VP16活化域与来自大肠杆菌的突变型Tet抑制因子融合以产生转录反式活化剂“rtTA”,其特异性DNA结合需要某些四环素(Tc)衍生物。多西环素(Doxycycline)是这种系统的诱导剂,并且其添加到其中基因表达处于rtTA控制下的所培养细胞中能够使诱导型启动子表达增强1000倍。第二代“TetO/CMV”启动子,命名为pTight或Ptet-14,其使用优化的7TetO间距和截短的最小CMV启动子设计,展现了降低的基础表达水平(克隆技术:pTRE-Tight载体(Clontech:pTRE-Tight Vectors),《克隆技术(Clontechiques)》2003,18(3):13-14)。启动子序列是:

[0321] TTTCGTCTCACACGAGTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTACTCCCTATCAGTG  
ATAGAGAACGATGTCGAGTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTACTCCCTATCAGTGATAGAG  
AACGTATGTCGAGTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTATCCCTATCAGTGATAGAGAACGTA  
TGTCGAGTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGGTAGGCCTGTACGGTGGGAGGCCTATATAAGCAG  
AGCTCGTTAGTGAACCGTCAGATGCC

[0322] 这种诱导型启动子或多种其它可调节启动子可以在本发明中用于控制结合剂呈现水平。可以使用任何诱导型系统,其中DNA结合域与能够结合诱导剂分子的蛋白域融合,引起蛋白质构象变化或DNA识别序列的亲和力变化。这将引起转录的去抑制或活化,促成蛋白质表达。举例来说,在T-Rex或二甲氨基二硫代甲酸铜(cumate)开关系统的情况下,诱导剂结合到抑制因子蛋白质引起DNA结合损失和转录的去抑制。或者,诱导剂结合到与转录活化域融合的DNA结合域(融合蛋白)将引起DNA结合和转录因子的募集,对于Tet-on<sup>92,114</sup>或GAL4 GeneSwitch系统<sup>121</sup>来说,情况如此。

[0323] 上文提到的四环素诱导型启动子系统可以转化成更和谐、均一且可滴定的诱导,这是通过使TetO重复序列的数目减少到一与六之间<sup>118</sup>或通过修饰rTA蛋白质来达成。实例8b描述了第三代诱导型tet启动子系统,其实现了改善的表达范围。细节绘示于图37中。还可以使用替代的诱导型表达系统,包括二甲氨基二硫代甲酸铜开关<sup>119</sup>、T-Rex<sup>120</sup>或GAL4系统<sup>121</sup>。

#### 结合剂和编码核酸的回收

[0325] 从文库中选择所关注的结合剂或克隆之后,共同的下一个步骤将是分离、鉴定和/或扩增编码结合剂的核酸(例如DNA或RNA)。任选地,可能希望修饰编码结合剂的核酸,例如

重构结合剂和/或将编码序列插入不同载体内。编码所展示结合剂的核酸(DNA或RNA)能够从所选细胞中回收,克隆到表达载体中并且所编码的多肽以分泌形式表达或用于展示。这能够对个别克隆进行。

[0326] 当供体DNA分子群体用于建立含有相同序列的多个拷贝的文库时,可以获得含有编码相同结合剂的DNA的两种或更多种克隆。情况也可能是克隆可以含有编码超过一种不同结合剂的供体DNA,例如位点特异性核酸酶存在超过一种识别序列时,如在本文中别处所详述。因此,文库的多样性(按照所编码或表达的不同结合剂的数目)可以不同于所得克隆的数目。

[0327] 文库中的克隆优选地含有编码结合剂谱系中的一种或两种成员的供体DNA且/或优选地表达结合剂谱系中的仅一种或两种成员。当要鉴定克隆和/或编码特定结合剂(针对指定靶标筛选文库时所鉴定)的DNA时,每个细胞有限数目个不同结合剂是一种优势。当克隆编码结合剂谱系中的单个成员时,这是最简单的。然而,如果选自文库的克隆编码少数不同结合剂(例如克隆可以编码结合剂谱系中的两个成员),则也可以直接鉴定所期望结合剂的相关编码DNA。如本文中别处所论述,编码一种或两种结合剂的克隆特别方便地通过选择位点特异性核酸酶的识别序列来产生,由于二倍体细胞含有双重复固定基因座(每个染色体拷贝存在一个),因此所述识别序列在二倍体基因组中的每个染色体拷贝中存在一次,并且供体DNA可以整合于一个或两个固定基因座。因此,文库的克隆可以各自表达结合剂谱系中的仅一个或两个成员。

[0328] 在结合剂是抗体分子的情况下,方法可以包括从克隆的细胞中分离出编码抗体分子的DNA、扩增编码至少一个抗体可变区(优选VH和VL域)的DNA,以及将DNA插入载体中以提供编码抗体分子的载体。具有恒定域的多聚体抗体分子可以转化成单链抗体分子以便以分泌的可溶形式表达。抗体可以不同形式呈现,但无论哪一种形式均使抗体得到选择,一旦分离出抗体基因,则可以将其再配置成许多不同形式。一旦分离VH或VL域,则能够将其再克隆到包含所需搭配物结构域的表达载体中。

[0329] 可以将编码所选结合剂的DNA整合于宿主细胞染色体中用于表达。重组蛋白的表达典型地涉及将编码所期望蛋白质的基因引入细胞中置于所述细胞表达的启动子控制下。举例来说,所述基因可以处于细胞巨大病毒增强子/启动子的控制下并且可以引入哺乳动物细胞,例如常用的人类胚肾293(HEK293)细胞或中国仓鼠卵巢(CHO)细胞。将表达构筑体引入宿主细胞中的标准方法是众所周知的。为了产生分泌型可溶性蛋白质,将前导序列置于所编码的基因之前,从而将所编码的蛋白质导向内质网。在缺乏跨膜域的情况下,所编码的基因分泌到培养基中,从培养基中提纯且浓缩。

[0330] 所期望的结合剂可以遵循本发明的任何方法,以分离的形式在溶液中提供,例如在从稳定转染以便表达结合剂的宿主细胞中分泌之后。然后可以测试可溶性结合剂的特性以证实其性能,例如评估可开发性性状,例如溶解度、自缔合、非特异性结合、FcRn相互作用等,或评估亲和力。适用于测定这些特点中的每一个的分析提供于本文的相关章节中,并且可以实施以证实结合剂展现所期望的特性且/或显示相关特征改善,例如其与亲本分子相比改善。

[0331] 结合剂的医药配制物

[0332] 使用本发明方法获得的结合剂可以提纯且/或分离的形式提供,并且可以配制成

包含一种或多种额外组分的组合物。组合物可以含有适合的载体、赋形剂和并入配制物中以改善传递、递送、耐受性等的其它试剂。实例配制物描述于《雷氏药学大全 (Remington's Pharmaceutical Sciences)》。旨在体内使用的结合剂可以根据所期望的施于患者的途径配制,例如注射液(任选地为水溶液)。各种递送系统是已知的,并且能够用于施用包含结合剂的医药组合物。施用方法包括皮内、肌内、腹膜内、静脉内、皮下、鼻内、硬膜外和口服途径。组合物可以通过任何方便的途径施用,通过输注或大剂量注射,通过上皮或粘膜皮肤内层(例如口腔粘膜、直肠和肠粘膜等)吸收来施用,并且可以与其它生物活性剂一起施用。施药可以是全身的或局部的。

[0333] 结合剂或包含结合剂的组合物可以装在医疗容器中,例如小瓶、注射器、IV容器或注射装置。在一个实例中,提供了一种试剂盒,其包括结合剂、包装和治疗方法的使用说明书。所述方法可以包括皮下施用于患者。

[0334] 可以将结合剂配制成组合物,任选地配制成水性缓冲溶液,其包含浓度为至少(或高于)以下的结合剂:50mg/ml、60mg/ml、70mg/ml、80mg/ml、90mg/ml或100mg/ml。结合剂的浓度可以是50-200mg/ml之间,例如50-150mg/ml之间,例如50与100mg/ml之间。

[0335] 用于浓缩所表达的重组蛋白的许多方法已为所属领域的技术人员所知,而且可以包括柱色谱(例如亲和色谱、离子交换色谱)和超滤<sup>27</sup>。

[0336] 条项

[0337] 以下编号的各项代表着本发明的声明并且是说明书的一部分。

[0338] 1. 一种结合剂在体外培养的高等真核细胞上的表面呈现水平作为所述结合剂在溶液中的溶解度和/或其自缔合抗性的预测指标的用途。

[0339] 2. 根据第1项的用途,其中所培养的细胞是表达结合剂的多样化谱系的展示文库的克隆。

[0340] 3. 一种高等真核细胞的培养文库用于体外选择结合剂以实现在溶液中的较高溶解度和/或较低自缔合倾向的用途,

[0341] 其中所述文库是各含有编码结合剂的DNA的高等真核细胞克隆的文库,其中已编码的所述结合剂呈现于细胞表面上。

[0342] 4. 一种结合剂发现方法,其中结合剂在展示文库的已培养高等真核细胞克隆表面上的表面呈现水平是用作所述结合剂在溶液中的溶解度和/或其自缔合抗性的预测指标。

[0343] 5. 一种根据结合剂在溶液中的溶解度和/或自缔合抗性将它们区分或排序和/或富集在溶液中展现较大溶解度和/或较大自缔合抗性的结合剂的方法,其包括(i)提供各含有编码结合剂的DNA的高等真核细胞克隆的文库,

[0344] (ii) 在用于表达所述结合剂的条件下体外培养所述克隆,其中所述结合剂呈现于细胞表面上,

[0345] (iii) 测定所述结合剂在所述克隆上的表面呈现水平,任选地通过用包含可检测(例如荧光)标记的试剂标记所述结合剂来测定,

[0346] (iv) 选出相较于其它克隆展现结合剂存在更高的表面呈现的一种或多种克隆,以及

[0347] (v) 鉴定出由所述一种或多种所选克隆编码的在溶液中具有良好溶解度和/或自缔合抗性的结合剂,且任选地提供所选克隆用于一个或多个进一步筛选步骤。

[0348] 6. 根据第1项到第3项中任一项的用途,或根据第4项或第5项的方法,其中所述结合剂是含有跨膜域的多肽。

[0349] 7. 根据第4项或第5项的方法,包括通过用包含可检测(例如荧光)标记的试剂标记所述结合剂来测定所述结合剂在所述克隆上的表面呈现水平,其中所述试剂结合到所述结合剂的恒定区,任选地其中所述结合剂包括Fc区且所述试剂结合到所述Fc区。

[0350] 8. 根据第4项到第7项中任一项的方法,包括根据结合剂在细胞上的表面呈现水平将细胞分选到收集部分和丢弃部分,其中将展示表面呈现大于预定阈值的细胞分选到所述收集部分并且将展示表面呈现低于所述预定阈值的细胞分选到所述丢弃部分。

[0351] 9. 根据第8项的方法,其中丢弃部分包括表达临界浓度为至少10mg/ml的对比多肽的细胞,且其中所述收集部分包括表达临界浓度为所述丢弃部分中的所述对比多肽的至少1.5倍高的结合剂的细胞。

[0352] 10. 根据第8项或第9项的方法,其中分选是利用荧光活化细胞分选仪(FACS)进行。

[0353] 11. 根据第4项到第10项中任一项的方法,其中步骤(ii)包括将所述文库中的所述克隆作为混合物在一个容器中培养。

[0354] 12. 根据第4项到第10项中任一项的方法,其中步骤(ii)包括在独立容器中培养所述文库中的每种克隆。

[0355] 13. 根据第4项到第12项中任一项的方法,其中所述结合剂是亲本结合剂的序列变体。

[0356] 14. 根据第13项的方法,其中所述亲本结合剂已鉴定为需要改善其在溶液中的溶解度或自缔合抗性。

[0357] 15. 根据第13项或第14项的方法,其中所述方法包括产生所述亲本结合剂的序列变体以及将编码所述序列变体的DNA整合到高等真核细胞的细胞DNA中以提供含有编码所述结合剂的DNA的细胞克隆的文库,

[0358] 任选地,其中所述方法包括分析所述亲本的多肽序列、鉴定出经预测会促进自缔合和/或减小溶解度的一个或多个氨基酸残基,以及使所述一个或更多个氨基酸残基产生突变。

[0359] 16. 根据第13项到第15项中任一项的方法,其中所述亲本结合剂在磷酸盐缓冲盐水溶液(PBS)中的临界浓度小于50mg/ml且/或在磷酸盐缓冲盐水溶液(PBS)中的溶解度极限小于50mg/ml,

[0360] 且/或其中所述方法包括鉴定出临界浓度和/或溶解度极限为所述亲本结合剂的至少1.5倍高的由所述一种或多种所选克隆编码的结合剂。

[0361] 17. 根据第4项到第16项中任一项的方法,包括基于结合剂在所述细胞克隆上的表面呈现水平来预测所述结合剂的亲水性和/或鉴定出具有较强亲水性的一种或多种所选克隆的结合剂。

[0362] 18. 一种体外筛选展示结合剂的高等真核细胞文库以从所述文库中富集表达结合剂的细胞的方法,所述结合剂在哺乳动物体内结合一种或多种非靶分子的倾向低,所述方法包括

[0363] (i) 提供各含有编码结合剂的DNA的高等真核细胞克隆的文库,

[0364] (ii) 在用于表达所述结合剂的条件下体外培养所述克隆,其中所述结合剂呈现于

细胞表面上，

[0365] (iii) 使所述结合剂暴露于所述一种或多种非靶分子，允许所述结合剂与所述一种或多种非靶分子之间发生结合，

[0366] (iv) 丢弃对所述一种或多种非靶分子展现较大结合的细胞，

[0367] (v) 选择对所述一种或多种非靶分子展现较低结合的细胞，以提供所选细胞群，所选细胞群富集表达对所述非靶分子结合倾向低的结合剂的克隆；以及任选地

[0368] 提供所选群体用于一个或多个进一步筛选步骤。

[0369] 19. 根据第18项的方法，其包括

[0370] (iii) 使所述结合剂暴露于包括所述一种或多种非靶分子的基质，允许结合到所述基质，

[0371] (iv) 丢弃对所述基质展现较大结合的细胞，

[0372] (v) 选择对所述基质展现较低结合的细胞，以提供所选细胞群，所选细胞群富集表达对所述非靶分子结合倾向低的结合剂的克隆；以及任选地

[0373] 提供所选群体用于一个或多个进一步筛选步骤。

[0374] 20. 根据第18项或第19项的方法，其中所述非靶分子包括DNA、肝素、乙酰肝素硫酸盐、软骨素硫酸盐、伴随蛋白、玻糖醛酸，或糖萼的一种或多种组分。

[0375] 21. 根据第18项到第20项中任一项的方法，其中对非靶分子的所述结合是低亲和力非特异性结合。

[0376] 22. 根据第18项到第21项中任一项的方法，其包括将所述文库中的所述克隆作为混合物在一个容器中培养以及使所述混合物暴露于所述基质。

[0377] 23. 根据第18项到第21项中任一项的方法，包括将所述文库中的每种克隆在独立容器中培养。

[0378] 24. 根据第18项到第23项中任一项的方法，其中所述结合剂是亲本结合剂的序列变异体。

[0379] 25. 根据第24项的方法，其中所述亲本结合剂已鉴定为需要减少对一种或多种非靶分子的结合。

[0380] 26. 根据第24项或第25项的方法，其中所述方法包括产生所述亲本结合剂的序列变异体以及将编码所述序列变异体的DNA整合到高等真核细胞的细胞DNA中以提供含有编码所述结合剂的DNA的细胞克隆的文库，

[0381] 任选地其中所述方法包括分析所述亲本的多肽序列、鉴定经预测会促进非特异性结合的一个或多个氨基酸残基，以及使所述一个或多个氨基酸残基产生突变。

[0382] 27. 根据第24项到第26项中任一项的方法，其中所述亲本结合剂对一种或多种非靶分子展现显著的结合。

[0383] 28. 根据第18项到第27项中任一项的方法，其包括

[0384] (iii) 使所述结合剂暴露于呈现所述一种或多种非靶分子的细胞或珠粒。

[0385] 29. 根据第28项的方法，其包括检测表达结合剂的细胞与呈现所述一种或多种非靶分子的所述细胞或珠粒的相互作用。

[0386] 30. 根据第29项的方法，其包括使用AC-SINS检测颗粒间距离以及选择由展现较高颗粒间距离的细胞所呈现的结合剂。

- [0387] 31. 根据第18项的方法,其中所述一种或多种非靶分子被可检测地标记。
- [0388] 32. 根据第31项的方法,其中所述一种或多种非靶分子被荧光标记。
- [0389] 33. 根据第32项的方法,其包括根据对所述一种或多种非靶分子的结合水平、利用FACS将细胞流式分选到收集部分和丢弃部分,其中将显示所述被标记非靶分子的荧光高于预定阈值的细胞分选到所述收集部分并且将显示所述被标记非靶分子的荧光低于所述预定阈值的细胞分选到所述丢弃部分。
- [0390] 34. 根据第4项到第33项中任一项的方法,其中编码所述结合剂的所述DNA是在强启动子的控制下表达。
- [0391] 35. 根据第34项的方法,其中所述启动子是组成型启动子。
- [0392] 36. 根据第35项的方法,其中所述启动子是CMV启动子。
- [0393] 37. 根据第34项的方法,其中所述启动子是已最大限度地诱导表达的诱导型启动子。
- [0394] 38. 根据第18项到第37项中任一项的方法,其进一步包括随后执行根据第5项到第17项中任一项的方法。
- [0395] 39. 根据第18项到第37项中任一项的方法,其进一步包括初始执行根据第5项到第17项中任一项的方法。
- [0396] 40. 一种选择针对靶标的一种或多种结合剂的方法,其包括
- [0397] 执行如第5项到第39项中任一项所定义的方法,其进一步包括
- [0398] 使所述结合剂暴露于所述靶标,允许同源结合剂识别所述靶标,其中展示同源结合剂的细胞结合到所述靶标,以及
- [0399] 选择一种或多种展示同源结合剂的克隆。
- [0400] 41. 根据任一前述项的方法,其包括
- [0401] (i) 同时测定所述结合剂的表面呈现水平和所述结合剂结合靶标的水平,以及协同选择展示同源结合剂的克隆,所述同源结合剂展现的表面呈现高于其它克隆;或
- [0402] (ii) 同时测定所述结合剂的表面呈现水平和非特异性结合到非靶分子的水平,以及协同选择展示结合剂的克隆,与其它克隆相比,所述结合剂展现更高的表面呈现和更低的非特异性结合;或
- [0403] (iii) 同时测定所述结合剂的靶标结合水平和非特异性结合到非靶分子的水平,以及协同选择展示同源结合剂的克隆,所述同源结合剂展现的非特异性结合低于其它克隆。
- [0404] 42. 一种鉴定以所期望亲和力识别靶标的结合剂的方法,所述方法包括
- [0405] (a) 提供各含有编码结合剂的DNA的高等真核细胞克隆的体外文库,其中所述结合剂呈现于所述细胞表面上,且其中被编码的所述结合剂从弱活性启动子表达且/或在细胞表面上以每个细胞100-60,000范围内的拷贝数表达,
- [0406] (b) 使所述文库暴露于所述靶标且允许同源结合剂识别所述靶标,其中展示同源结合剂的细胞结合到所述靶标,
- [0407] (c) 分离出结合到所述靶标的细胞以提供所选细胞群,所选细胞群富集展示同源结合剂的细胞,以及任选地
- [0408] (d) 使所选细胞群暴露于针对所述靶标的额外一轮或多轮选择,任选地其中逐渐

地减小靶标浓度以提高选择严格度,以及任选地

[0409] (e) 选出展示同源结合剂的一种或多种克隆,所述同源结合剂对所述靶标具有所期望的亲和力。

[0410] 43.根据第42项的方法,包括提供富集展示同源结合剂的细胞的所选细胞群,以及随后:

[0411] 提供处于强活性启动子控制下的编码结合剂的DNA,所述DNA来自高等真核细胞克隆的体外文库内的所选细胞群,以及

[0412] 执行根据第5项到第17项中任一项的方法。

[0413] 44.根据第42项的方法,其包括

[0414] 执行根据第5项到第17项中任一项所定义的方法以提供展示结合剂存在较高表面呈现的所选克隆,以及随后:

[0415] 在各含有编码结合剂的DNA的高等真核细胞克隆的体外文库中,从弱活性启动子表达所述结合剂,以及随后

[0416] 执行第42项的方法。

[0417] 45.一种鉴定识别靶标的结合剂的方法,其包括:

[0418] (i) 提供各含有编码结合剂的DNA的高等真核细胞克隆的文库,其中所述结合剂是在诱导型启动子的控制下表达以便呈现于细胞表面上,

[0419] (ii) 在所述诱导型启动子活性弱的条件下培养所述文库的细胞,

[0420] (iii) 使所述文库暴露于所述靶标,允许同源结合剂识别所述靶标,其中展示同源结合剂的细胞结合到所述靶标,

[0421] (iv) 选择展示同源结合剂的细胞,借此提供所选细胞群,

[0422] (v) 在增强从所述诱导型启动子表达结合剂的条件下培养所选细胞群,

[0423] (vi) 测定所述结合剂在所述多种克隆上的表面呈现水平,任选地通过用包含可检测(例如荧光)标记的试剂标记所述结合剂来测定,

[0424] (vii) 选出相较于其它克隆展现结合剂存在更高的表面呈现的一种或多种克隆。

[0425] 46.一种鉴定识别靶标的结合剂的方法,其包括:

[0426] (i) 提供各含有编码结合剂的DNA的高等真核细胞克隆的文库,其中所述结合剂是在诱导型启动子的控制下表达以便呈现于细胞表面上,

[0427] (ii) 在从所述诱导型启动子强烈表达结合剂的条件下培养所述文库,

[0428] (iii) 测定所述结合剂在所述多种克隆上的表面呈现水平,任选地通过用包含可检测(例如荧光)标记的试剂标记所述结合剂来测定,

[0429] (iv) 选出相较于其它克隆展现结合剂存在更高表面呈现的克隆群体,

[0430] (v) 在从所述诱导型启动子弱表达结合剂的条件下培养所选群体,

[0431] (vi) 使所述文库暴露于所述靶标,允许同源结合剂识别所述靶标,其中展示同源结合剂的细胞结合到所述靶标,

[0432] (vii) 选择一种或多种展示同源结合剂的克隆。

[0433] 47.根据第42项到第46项中任一项的方法,其中所述启动子是四环素诱导型启动子。

[0434] 48.根据第42项到第47项中任一项的方法,其中所述靶标是用可检测试剂标记,例

如荧光标记。

[0435] 49. 根据第48项的方法,其中所述方法包括根据所述细胞上的靶标结合水平将细胞分选到收集部分和丢弃部分中,其中将展示所结合靶标高于预定阈值的细胞分选到所述收集部分中并且将展示所结合靶标低于所述预定阈值的细胞分选到所述丢弃部分中。

[0436] 50. 根据第49项的方法,其中分选是利用荧光活化细胞分选仪(FACS)进行。

[0437] 51. 根据第5项到第50项中任一项的方法,其包括

[0438] 测定编码所述结合剂的所述DNA的序列,所述结合剂来自所述一种或多种所选克隆,以及

[0439] 提供编码所述结合剂的经分离的核酸。

[0440] 52. 根据第5项到第51项中任一项的方法,其进一步包括

[0441] 测定编码所述结合剂的所述DNA的序列,所述结合剂来自所述一种或多种所选克隆,以及

[0442] 在体外,在以可溶形式分泌所述结合剂的条件下,在宿主细胞中表达编码所述结合剂的DNA。

[0443] 53. 根据第52项的方法,其中分泌的所述结合剂是以至少1mg/ml的产量获得。

[0444] 54. 根据第52项或第53项的方法,其进一步包括提纯和/或浓缩所述结合剂以获得浓度为至少10mg/ml的所述结合剂的水溶液。

[0445] 55. 根据第54项的方法,其中所述浓度是至少50mg/ml。

[0446] 56. 根据第55项的方法,其中所述浓度是至少100mg/ml。

[0447] 57. 根据第52项到第56项中任一项的方法,包括将所述结合剂配制成包括医药学上可接受的赋形剂的组合物。

[0448] 58. 根据第57项的方法,包括将所述组合物提供于注射用的预装药注射器中,所述预装药注射器任选地存在于试剂盒内,所述试剂盒包括一种或多种额外的组件,例如针和/或产品信息说明书,包括通过注射施用所述组合物的说明。

[0449] 59. 一种鉴定与FcRn相互作用的结合剂的方法,所述方法包括

[0450] 提供多种高等真核细胞克隆,其各自含有编码具有Fc域的不同结合剂的DNA,

[0451] 在体外,在使所述结合剂呈现于细胞表面上的条件下培养所述克隆,

[0452] 使所述克隆在约pH 6.0和约pH 7.4暴露于FcRn受体,允许所述Fc域识别FcRn,

[0453] 选出表达结合剂的一种或多种克隆,所述结合剂在约pH 6.0展现的亲和力结合高于在约pH 7.4展现的亲和力、在约pH 6.0展现的亲和力结合低于在约pH 7.4展现的亲和力结合,或在约pH 6.0展现的亲和力结合与在约pH 7.4展现的亲和力结合大约相同,以及

[0454] 任选地提供所选克隆用于一个或多个进一步筛选步骤。

[0455] 60. 根据第59项的方法,其包括:选出表达结合剂的一种或多种克隆,所述结合剂在约pH 6.0展现的亲和力结合高于在约pH 7.4展现的亲和力结合;以及鉴定体内半衰期延长的由一种或多种所选克隆编码的所述结合剂。

[0456] 61. 根据第59项或第60项的方法,其中所述结合剂包括展现序列多样性、任选地在一个或多个互补决定区中展现序列多样性的可变域。

[0457] 62. 根据第59项到第61项中任一项的方法,其中所述结合剂包括展现序列多样性、任选地在其CH3域中展现序列多样性的Fc区域。

[0458] 63. 根据第59项到第61项中任一项的方法, 其中所述结合剂的所述Fc区域不展现序列多样性。

[0459] 64. 一种克隆、由克隆表达的结合剂, 或编码所述结合剂的核酸, 其大体上如本文所述且/或利用根据任一前述项的方法鉴定或选择。

[0460] 65. 一种各含有编码结合剂的DNA的高等真核细胞克隆的体外文库, 其中所述编码DNA任选地处于细胞DNA中的固定基因座, 且其中被编码的所述结合剂在细胞表面上以每个细胞100-1000范围内的拷贝数表达。

[0461] 66. 一种如第65项中所定义的文库用于对针对靶标的结合剂进行基于亲和力的选择的用途。

[0462] 67. 一种含有编码结合剂谱系的DNA的高等真核细胞克隆的体外展示文库, 其中结合剂是在四环素诱导型启动子的控制下表达以便呈现于细胞表面上。

[0463] 68. 一种生产含有编码结合剂谱系的DNA的高等真核细胞克隆文库的方法, 其包括

[0464] 提供编码所述结合剂的供体DNA分子, 和高等真核细胞,

[0465] 将所述供体DNA引入所述细胞中, 借此建立含有整合于所述细胞DNA中的供体DNA的重组细胞,

[0466] 其中所述结合剂是在四环素诱导型启动子的控制下表达以便呈现于细胞表面上, 以及

[0467] 培养所述重组细胞以产生克隆,

[0468] 借此提供含有编码所述结合剂谱系的供体DNA的高等真核细胞克隆文库。

[0469] 69. 根据第68项的方法, 其中所述重组细胞是通过将所述供体DNA引入所述细胞中且在所述细胞内提供位点特异性核酸酶来建立, 其中所述核酸酶使细胞DNA中的识别序列裂解以建立整合位点, 在所述位点处, 所述供体DNA整合于所述细胞DNA中, 整合是通过所述细胞内源的DNA修复机制发生。

[0470] 70. 根据第68项或第69项的方法, 其进一步包括诱导供体DNA从所述四环素诱导型启动子表达以及在表达所述结合剂的条件下培养所述细胞, 从而获得结合剂在细胞表面上的呈现。

[0471] 71. 根据第69项或第70项的方法, 其进一步包括在如任一前述项所定义的方法或用途中使用所述文库。

[0472] 72. 根据任一前述项的用途、方法或文库, 其中所述高等真核细胞是哺乳动物细胞。

[0473] 73. 根据任一前述项的用途、方法或文库, 其中所述哺乳动物细胞是人类细胞系或CHO细胞系。

[0474] 74. 根据任一前述项的用途、方法或文库, 其中所述高等真核细胞处于悬浮培养中。

[0475] 75. 根据任一前述项的用途、方法或文库, 其中编码所述结合剂的DNA整合于所述细胞DNA中的固定基因座处。

[0476] 76. 根据任一前述项的用途、方法或文库, 其中所述结合剂是抗体。

[0477] 77. 根据第76项的用途、方法或文库, 其中所述抗体是全长免疫球蛋白。

[0478] 78. 根据第77项的用途、方法或文库, 其中所述抗体是IgG。

[0479] 79. 根据第76项到第78项中任一项的用途、方法或文库,其中所述抗体包含与跨膜域融合的重链,和轻链。

[0480] 80. 根据第1项到第79项中任一项的用途、方法或文库,其中所述结合剂是融合蛋白,所述融合蛋白包括插入受体多样性支架域内的供体多样性支架域,任选地包括与所述融合蛋白缔合的搭配域,

[0481] 其中所述供体多样性支架域包括供体支架和供体相互作用序列并且所述受体多样性支架域包括受体支架和受体相互作用序列。

[0482] 81. 根据第80项的用途、方法或文库,其中所述融合蛋白是打结抗体,其包括插入抗体可变域内的富半胱氨酸肽。

[0483] 82. 根据第81项的用途、方法或文库,其中所述打结抗体包括与跨膜域融合的抗体重链,和抗体轻链。

[0484] 83. 根据任一前述项的用途、方法或文库,其中所述结合剂包括展现序列多样性、任选地在一个或多个互补决定区中展现序列多样性的抗体可变域。

[0485] 84. 根据任一前述项的用途、方法或文库,其中所述结合剂包括展现序列多样性、任选地在其CH3域中展现序列多样性的Fc区域。

[0486] 85. 根据任一前述项的用途、方法或文库,其中所述结合剂具多特异性,包括用于第一靶标的第一结合位点和用于第二靶标的第二结合位点。

[0487] 86. 根据任一前述项的用途、方法或文库,其中所述文库包括至少 $10^3$ 种克隆。

[0488] 87. 根据任一前述项的用途、方法或文库,其中所述文库是初始文库。

[0489] 88. 根据第1项到第86项中任一项的用途、方法或文库,其中所述文库中的所述克隆已经根据对所选靶标的结合预选出来。

[0490] 89. 根据第88项的用途、方法或文库,其中所述靶标是人类多肽。

[0491] 90. 根据第88项或第89项的用途、方法或文库,其中所述文库中的所述克隆已经根据对两种不同靶标的双特异性结合预选出来。

[0492] 所属领域的技术人员将认识到,或者使用不超出常规的实验就能够确定本文所述的发明的具体实施例的许多等效物。此类等效物希望被所附权利要求书涵盖。

[0493] 本发明的其它方面和实施例是由上述方面和实施例通过用术语“由……组成”替换术语“包含”而得到以及由上述方面和实施例通过用术语“基本上由……组成”替换术语“包含”而得到。

[0494] 应了解,除非上下文另有需求,否则本申请公开了上述方面和上述实施例中的任一个彼此间的所有组合。类似地,除非上下文另有需求,否则本申请公开了优选特征和/或任选特征单独或与其它方面中的任一个的所有组合。

[0495] 除非另有定义,否则本文使用的所有技术和科学术语都具有与本发明所属领域的普通技术人员通常所了解相同的含义。虽然类似于或等效于本文所述的任何组合物和方法能够用于实施或测试本公开的方法,但本文中描述了示例性的组合物和方法。本文所述的本公开的任一方面和实施例还可以加以组合。举例来说,本文公开的的任何从属或独立权利要求的标的物可以多次组合(例如,来自每个从属权利要求的一个或多个叙述可以基于它们所从属的独立权利要求而组合成单个权利要求)。

[0496] 如本文和权利要求书中所用,单数形式“一(a)”、“和(and)”和“所述(the)”包括复

数个提及物,除非上下文另有明确规定。因此,举例来说,提及“肽链”是提及一个或多个肽链并且包括所属领域的技术人员已知的其等效物。

[0497] 本说明书中提及的所有文献和序列数据库以全文引用的方式并入本文中用于所有目的。

[0498] “和/或”在本文中使用时被认为是具体公开了两种指定特征或组分中的每一种与或不与另一种指定特征或组分。举例来说,“A和/或B”将被认为是具体公开了(i)A、(ii)B和(iii)A和B中的每一个,就像各自在本文中单独列出一样。

[0499] 附图的简要说明

[0500] 本发明的实施例现将参考附图更详细地描述如下:

[0501] 图1.pINT17-BSD,用于表面表达的双重启动子抗体IgG表达盒

[0502] pINT17-BSD,所示示意图描绘关键特征(1-7500bp)

[0503] pINT17-BSD-D1.3,用于表面表达抗溶菌酶抗体D1.3的双重启动子抗体表达盒。所示为充分标注的核酸序列。特征:

[0504] AAVS左同源臂9-812

[0505] 杀稻瘟菌素抗性基因853-1254

[0506] pEF启动子1522-2705

[0507] BM40前导序列2745-2799

[0508] 人源化D1.3 VL 2799-3130

[0509] 人类Cκ3138-3443

[0510] BGH聚腺苷酸3468-3682

[0511] CMV启动子3701-4273

[0512] 具有内含子4299-4426的小鼠VH前导序列

[0513] 人源化D1.3 VH 4432-4779

[0514] 优化的人类IgG1 CH1-CH3 4780-5775

[0515] Myc标签5776-5805

[0516] PDGFR锚5806-5961

[0517] BGH聚腺苷酸6011-6225

[0518] AAVS右同源臂6288-7124

[0519] f1复制起点7282-7695

[0520] pUC复制起点7916-8590

[0521] 卡那霉素(Kanamycin)抗性基因9310-10104

[0522] 图2.IgG在细胞表面上的表达。

[0523] 通过在FL3通道中使用前向散射和染色,使分析聚焦于活细胞。FL3通道中染色呈阳性的细胞(代表吸收7-AAD的非活细胞)加以排除。在AAVS TALEN存在下,用pINT17抗体转染HEK293细胞。利用杀稻瘟菌素选择稳定群体。细胞在转染后第14天,用抗Fc PE (FL2) 染色。描绘荧光强度(抗Fc-PE,x轴)的图相对于细胞计数(y轴)绘制并且包括CNT0607 (a)、MEDI1912 (b) 和Ang2mAb (c) 对。对于所有图来说,图包括染色的HEK293细胞(虚线)、亲本抗体(虚线)、改善的突变体(黑色实线)。

[0524] 图3.MEDI-1912和MEDI-1912-STT的制备型尺寸排阻色谱。通过瞬时转染Expi293

细胞来表达抗体,随后进行蛋白质A亲和提纯和透析。使用PBS (pH 7.4) 操作缓冲液将提纯的MEDI-1912 (0.5ml, 1.1mg/ml) 或MEDI-1912-STT (0.5ml, 1.7mg/ml) 装载于与AKTA Pure系统连接的Superdex 200 10/300柱上。洗脱体积 (ml) 在x轴上相对于280nm吸光度 (mAU) 在y轴上作图。MEDI-1912和MEDI-1912-STT的洗脱体积 (Ve) 分别是10.3ml和11.7ml。MEDI-1912显示的洗脱体积较早,表明自身相互作用。

- [0525] 图4. MED-1912可变重 (VH) 链的DNA和蛋白质序列。
- [0526] 建立文库所用的引物在核酸序列上方标识。
- [0527] 图5. 基于抗体表达对展示抗体群体的混合细胞进行的FACS分离。
- [0528] 通过核酸酶定向整合到HEK293细胞的AAVS基因座中而靶向MEDI-1912与MEDI-1912\_STT IgG基因的均等混合物。转染后第15天,通过使用BD Influx分选仪的FACS、基于抗体表达来分离这个混合细胞群。细胞用藻红蛋白 (PE) 标记的抗Fc和别藻蓝蛋白 (APC) 标记的NGF-生物素/抗生蛋白链菌素染色。通过使用前向散射和染色使分析聚焦于活细胞。 $\lambda_{em}=450/40$ 、 $\lambda_{exc}=355$ 通道中染色呈阳性的细胞 (代表着吸收7-AAD的非活细胞) 加以排除。
- [0529] a. 抗Fc-PE的荧光强度相对于混合型输入群体 (虚线) 以及展示MEDI-1912 (灰线) 和MEDI-1912\_STT (黑线) 的单克隆HEK293细胞系的细胞计数的直方图。
- [0530] b. 点阵图显示抗Fc-PE的荧光强度 (x轴),其代表抗体表达水平,相对于y轴上的抗原结合荧光强度 (NGF-生物素/NGF-生物素/抗生蛋白链菌素-APC) 所绘制。为了分离高表达和低表达抗体群体而选择的门标识为P5和P6并且在点阵图上以黑框显示。总事件计数为 $3.9 \times 10^6$ 个细胞并且在P5和P6门中分选的细胞数目分别是 $2.5 \times 10^5$ 和 $2.8 \times 10^5$ 个细胞。
- [0531] 图6. 富集根据哺乳动物细胞展示水平所选的抗体。
- [0532] 将门5和6中分选的细胞 (图5) 扩增,制备基因组DNA且通过PCR分离抗体VH基因。通过Nextgen测序来分析两个抗体群体,以测定MEDI-1912和MEDI-1912\_STT在两个门内群体中的比例。直方图显示MEDI-1912 (阴影条柱) 和MEDI-1912\_STT (黑色实心条柱) 在低表达抗体群体 (门6) 和高表达抗体群体 (门5) 中的频率百分比。
- [0533] 图7. 基于抗体表达,对MEDI-1912文库抗体群体进行的FACS分离。
- [0534] 通过核酸酶定向整合到HEK293细胞的AAVS基因座中而靶向MEDI-1912 IgG基因文库,其中利用NNS寡核苷酸定向诱变使编码W30、F31和L56的密码子随机突变。转染后15天,使用BD Influx分选仪、通过FACS、基于抗体表达和抗原结合来分离这个混合细胞群。细胞用藻红蛋白 (PE) 标记的抗Fc和别藻蓝蛋白 (APC) 标记的NGF-生物素/抗生蛋白链菌素染色。通过使用前向散射和染色使分析聚焦于活细胞。 $\lambda_{em}=450/40$ 、 $\lambda_{exc}=355$ 通道中染色呈阳性的细胞 (代表着吸收7-AAD的非活细胞) 加以排除。点阵图显示抗Fc-PE的荧光强度 (x轴),其代表抗体表达水平,相对于展示亲本MEDI-1912的单克隆细胞系 (a)、展示MEDI-1912\_STT的单克隆细胞系 (b) 和MEDI-1912氨基酸位置30、31和56随机文库 (c) 在y轴上的抗原结合荧光强度 (NGF-生物素/NGF-生物素/抗生蛋白链菌素-APC) 绘制。选用于分析的门标识为P5和P6并且在点阵图上以方框显示。
- [0535] 图8. 所选MEDI-1912变异体的序列分布。
- [0536] 哺乳动物展示且FACS基于Fc表达和NGF结合设门 (P5门,图7) 之后,MEDI-1912VH文库在MEDI-1912VH位置30、31和56的氨基酸一致性频率的直方图。氨基酸 (单字母码) 在x轴

上相对于邻近氨基酸30(阴影条柱)和31(黑色条柱(a))和氨基酸56(b)在y轴上的频率(出现率百分比)作图。位置56的亮氨酸从分析中排除。

[0537] 图9.博可珠单抗(Bococizumab)小鼠亲本抗体5A10 VH(A)和VL(B)与人源化中间抗体5A10-i和博可珠单抗的比对。

[0538] CDR用序列上方的条柱指示并且待突变的残基用粗体加下划线突出显示。互补位残基(对直接结合到PCSK9有贡献的氨基酸)用斜体加下划线突出显示。点表示与亲本小鼠mAb 5A10一致。

[0539] 图10.抗体哺乳动物展示博可珠单抗和亲本人源化中间体抗体5A10-i的表达

[0540] 编码博可珠单抗或5A10-i IgG的靶向载体pINT17通过TALE核酸酶整合于Hek293细胞的AAVS基因座中。转染后(dpt)第1、8或21天,将细胞( $10^6$ )用抗Fc-PE染色,并且使用iQue Intellicyte流式细胞仪、通过流式细胞术分析 $10^6$ 个细胞。死细胞从分析中排除。直方图显示转染后(dpt)第1、8和21天,博可珠单抗和5A10-i细胞展示群体的荧光强度(抗Fc-PE)相对于细胞计数,其中野生型HEK293细胞的染色作为阴性对照纳入。

[0541] 图11.博可珠单抗VH与人类生殖系序列(IMGT)的比对。

[0542] 针对人类VDJ数据库,对博可珠单抗VH(查询\_1)进行Ig基本局部比对检索(IgBLAST)。结果展示为与查询序列的比对,按照一致性百分比的次序(第二栏),所述查询序列涵盖框架区1(FR1)、互补决定区1(CDR1)、FR2、CDR2和FR3。人类生殖系显示于第1栏中。与博可珠单抗VH序列一致的残基(.)和差异(单氨基酸码)显示于多序列比对中。

[0543] 图12.编码博可珠单抗VH变异体和VL的DNA序列,和终止密码子模板。与原始“野生型”博可珠单抗(列f)相比,变异是以粗体加下划线突出显示。侧接的5' 和3' VH限制位点(NcoI和XhoI)或VL限制位点(NheI和NotI)加下划线。VL终止密码子以粗体加下划线突出显示。

[0544] 图13.博可珠单抗文库经MACS提纯后的分析型流式细胞术分析

[0545] 经博可珠单抗文库转染的Hek293细胞在抗Fc或PCSK9i转染后第7天经MACS提纯。(a)流式细胞术点阵图显示了抗Fc表达(FL2,x轴),其相对于MACS提纯后文库、HEK293细胞和未分选文库、博可珠单抗和5A10i转染物(9dpt)的PCSK9结合(FL4,y轴)绘制。(b)荧光强度(抗Fc,FL2,x轴)相对于细胞计数绘制的直方图。图(从上到下)是HEK293对照、博可珠单抗、5A10i、博可珠单抗文库、抗PCSK9MACS提纯的博可珠单抗文库,和抗Fc MACS提纯的博可珠单抗文库。

[0546] 图14.预先经MACS提纯的博可珠单抗文库的BD Influx分选仪点阵图,其基于抗原结合或抗Fc。

[0547] 通过核酸酶定向整合到HEK293细胞AAVS基因座中而靶向博可珠单抗IgG基因文库。这个混合细胞群首先基于PCSK9结合(a)或抗Fc(b),经MACS提纯。转染后的第16天,基于抗体表达和抗原结合,使用BD Influx分选仪、通过FACS分离富集MACS的文库。细胞用藻红蛋白(PE)标记的抗Fc和别藻蓝蛋白(APC)标记的PCSK9-生物素/抗生蛋白链菌素染色。通过使用前向散射和染色使分析聚焦于活细胞。 $\lambda_{em}=450/40$ 、 $\lambda_{exc}=355$ 通道中染色呈阳性的细胞(代表着吸收7-AAD的非活细胞)加以排除。点阵图显示抗Fc-PE的荧光强度(x轴),其代表抗体表达水平,相对于y轴上的抗原结合荧光强度(PCSK9-生物素/抗生蛋白链菌素-APC)所绘制。选用于分析的门标识为P5和P6并且在点阵图上以方框显示。

[0548] 图15. 哺乳动物展示选择之后的博可珠单抗VH分布。

[0549] 对未选定的随机输入克隆(84)、所分选的抗原(75)和所选择的Fc(85)进行测序并且测定VH一致性。直方图显示相对于输入的出现率百分比(白色实心条柱)而在x轴上绘制的VH生殖系一致性,抗原MACS、随后用所选的抗原和抗Fc FACS(黑色实心条柱)和抗Fc MACS、随后用所选的抗原和抗Fc FACS(灰色实心条柱)哺乳动物细胞展示所选群体。

[0550] 图16. 哺乳动物展示选择之后的博可珠单抗VL序列分析。

[0551] 对未选定的随机输入克隆(84)、所分选的抗原(75)和所选择的Fc(85)进行测序并且测定VL一致性。计算3种突变密码子的平均pI和脂肪族指数。这显示哺乳动物展示所选的抗体的pI与脂肪族指数均减小。

[0552] 图17. 表格列举了哺乳动物展示的博可珠单抗克隆,其根据抗原结合、通过MACS富集,随后根据抗体展示水平(抗Fc)与抗原结合进行FACS富集。

[0553] 对克隆进行测序并且VH CDR1和CDR2和VL CDR2和CDR3单字母氨基酸序列相对于博可珠单抗原始序列显示有变异,所述变异以粗体和红色突出显示。保持博可珠单抗序列的靶标氨基酸加有下划线。在捕捉ELISA中实施抗体(包括原始亲本抗体博可珠单抗和5A10-i)对抗原的结合并且结合信号(荧光单位)显示于第2栏中。如Liu等人,2014<sup>30</sup>此前所述实施AC-SINS分析并且第3栏显示相较于非抗体PBS对照(nm)存在最大吸光度波长漂移。所选人类VH生殖系也在第6栏中用字母指出,如实例5中详述:

[0554] a: VH Y33A-IGHV1-3\*01

[0555] b: VH Y33D-IGHV1-8\*01

[0556] c: VH S52N、F54S、R57S-IGHV1-46\*01

[0557] d: VH Y33A、S52N、F54S、R57S(a. 和c.:组合的突变体)

[0558] e: VH Y33D、S52N、F54S、R57S(b. 和c.:组合的突变体)

[0559] f: 博可珠单抗“野生型”序列

[0560] 图18.. 抗PCSK9 IgG1抗体的HPLC-SEC。

[0561] 通过瞬时转染Expi-293细胞来表达抗体,通过蛋白质A色谱进行亲和力提纯且透析。使用Agilent 1100HPLC仪器,以0.35ml/min的流量将样品(2μl,1mg/ml)装载于Agilent AdvancedBio SEC 300A,2.7um,4.6x300 mm柱(安捷伦科技(Agilent Technologies),目录号PL1580-5301)。显示所选抗体的滞留时间相对于吸光度的图。从黑色到逐渐变浅的灰色是:5A10-i、884\_01\_G01(根据哺乳动物细胞展示所鉴定)、博可珠单抗和阿利库单抗(Alirocumab)。

[0562] 图19. 凝胶过滤分析:尼沃单抗(Nivolumab)(a)和韦森单抗(Vesencumab)(b)。

[0563] 通过瞬时转染Expi293细胞来表达抗体,随后进行蛋白质A亲和提纯且透析。使用PBS(pH 7.4)操作缓冲液将提纯的尼沃单抗(0.5ml,1.3mg/ml)或韦森单抗(0.5ml,1.2mg/ml)装载于与AKTA Pure系统连接的Superdex 200 10/300柱上。洗脱体积(ml)在x轴上相对于280nm吸光度(mAU)在y轴上作图。尼沃单抗和韦森单抗的洗脱体积(Ve)分别是12.0ml和13.7ml。

[0564] 图20. 尼沃单抗(a)和韦森单抗(b)在4℃储存2周之后的稳定性测定。

[0565] 通过尺寸排阻色谱来提纯韦森单抗和尼沃单抗(参见图19)且用PBS将其浓度调节到0.5mg/ml(pH 7.4)。然后在4℃储存抗体2周。在20℃下,根据制造商说明书,使用

Zetasizer APS(马尔文仪器公司(Malvern Instruments),英国马尔文(Malvern,UK)进行动态光散射测量。利用爱因斯坦-斯托克斯方程式(Einstein-Stokes equation)评估流体动力学半径且相对于分散强度作图。相较于韦森单抗(b)的多个聚集体峰,观察到尼沃单抗(a)存在单个单分散峰。

[0566] 图21.人类血清对细胞表面上的IgG的结合。

[0567] 通过在FL3通道中使用前向散射和染色,使分析聚焦于活细胞。FL3通道中染色呈阳性的细胞(代表着吸收7-AAD的非活细胞)加以排除。细胞在AAVS TALEN存在下用pINT17-尼沃单抗或pINT17-韦森单抗转染。利用杀稻瘟菌素选择稳定群体。转染后的第20天,细胞用抗Fc PE (FL2) 和经Dylight 633 (325 0000, Innova) 标记的人类血清(H4522, Sigma) 染色。图为未转染的HEK293细胞(a)、经pINT17-尼沃单抗(b) 或pINT17-韦森单抗(c) 转染的HEK293细胞。

[0568] 图22.亲和力、抗原浓度和抗体浓度之间的关系

[0569] 在 $K_D$  10nM(虚线)或0.1nM(实线)的不同浓度的抗体存在下,使用0.1nM抗原的复合物的浓度。

[0570] 较高亲和力抗体( $K_D$  0.1nM)相对于较低亲和力( $K_D$  10nM)在不同抗体浓度下结合到0.1nM抗原的双向相对选择率

[0571] 即使抗原浓度低(“严格”),但在抗体浓度高的情况下,选择率相对极小,但这随着抗原浓度降低而增加。

[0572] 图23.控制抗体展示水平的剪接受体/供体变异体。

[0573] 图示为pINT17-J9、J10、J29和J30变异体的从HindIII位点到5'内含子的核酸序列,包括剪接供体区域。人类原始“野生型”序列是J9并且J9在剪接点与J10、J29和J30不同的核苷酸加有下划线。

[0574] 图24.pINT17-J30,双重启动子抗体IgG表达盒,用于减少展示表面的表达图示了XhoI(4804)与SbfI(8387)限制位点之间的经标注的核酸序列。特征:

[0575] IgG1 CH1-3 4805-5808

[0576] 剪接接合点5801-5802

[0577] 内含子5802-7104

[0578] M1外显子7105-7239

[0579] BGH pA 7264-7478

[0580] AAVS右同源臂7544-8381

[0581] 3' β-球蛋白隔离子8421-8492

[0582] 图25.使用替代跨膜域和剪接变异体使IgG在细胞表面上的表面表达减少。

[0583] 通过在FL3通道中使用前向散射和染色,使分析聚焦于活细胞。FL3通道中染色呈阳性的细胞(代表着吸收7-AAD的非活细胞)加以排除。细胞在AAVS TALEN存在下用靶向pINT17的载体转染。利用杀稻瘟菌素选择稳定群体。细胞在转染后第27天,用抗Fc PE (FL2) 染色。流式细胞术点阵图中的各图包括pINT17-J9-尼沃单抗(a)、pINT17-J10-尼沃单抗(b)、pINT17-J29-尼沃单抗(c)、pINT17-J30-尼沃单抗(d) 和pINT17-BSD-尼沃单抗(e)。

[0584] 图26.根据由pINT17-BSD或pINT17-J30靶向载体表达的抗体,量化细胞表面上的IgG展示水平

[0585] 如经小鼠 IgG-PE 标记染色的 Quantum Simply 细胞抗小鼠 IgG 珠粒 (目录号 815, Bangs 实验室有限公司) 的制造商说明书中所述进行校准珠粒 FL2 染色。(a) 标记的直方图显示校准珠粒组的染色, 其中标识为 1、2、3 和 4 的峰分别代表 12257、72745、283360、886417 的珠粒拷贝数。空白峰代表不具有捕捉抗体的珠粒。(b) 校准图显示相对于拷贝数 (y 轴) 绘制的中值荧光强度 (x 轴)。来自 pINT17-BSD 或 pINT17-J30 表达盒的展示 337\_1\_C08 (c) 和尼沃单抗 (d) 的细胞系用抗 Fc-PE ( $5\mu\text{l}, 0.1\text{mg/ml}; 10^5$  个细胞) 染色。通过在 FL3 通道中使用前向散射和染色, 使分析聚焦于活细胞。FL3 通道中染色呈阳性的细胞 (代表着吸收 7-AAD 的非活细胞) 加以排除。直方图显示 pINT17-BSD 和 PDGFR TM (标有黑色实线)、pINT17-J30 (标有虚线) 和野生型 HEK293 细胞系 (灰色实线) 作为分别展示 337\_1\_C08 (c) 和尼沃单抗 (d) 的细胞的荧光强度相对于细胞计数。

[0586] 图 27. 通过减少细胞展示拷贝数能够分离哺乳动物展示的对其靶标具有不同亲和力的抗体。

[0587] 展示尼沃单抗和 337\_1\_C08 抗体的 Hek293 细胞分别用 50nM 细胞追踪绿和 50nM 细胞追踪红标记。(a) 展现使用 J30 剪接变异体的展示和 (b) 展现使用由 pINT17-BSD 载体编码的 PDGFR 跨膜域的展示。将经标记的细胞均等混合且基于抗原结合进行 MACS 分选。使用 intellicyt 流式细胞仪分析已分选的细胞。点阵图在 x 轴上展现尼沃单抗 (FL1) 且在 y 轴上展现 337\_1\_C08 (FL4)。图 i、ii、iii 和 iv 分别展现 MACS 提纯所用的 10nM、1nM、0.1nM 抗原和无抗原。

[0588] 图 28. pINT18-Tet1, 用于减少展示表面表达的诱导型启动子抗体 IgG 表达载体。图示了 AsIS1(5) 和 SbfI(7672) 限制位点之间的经标注的核酸序列。包含复制起点和卡那霉素抗性基因的 AsIS1 和 SbfI 位点 (7673-10922 和 1-4) 外部的载体主链与 pINT17-BSD 一致 (图 1)。

[0589] 主要特点:

- [0590] AAVS 左同源臂 9-812
- [0591] 杀稻瘟菌素抗性基因 853-1254
- [0592] CMV 启动子 1540-2112
- [0593] 反向 Tet 活化因子 (tTA) CDS 2164-3168
- [0594] SV40 pA 3178-3395
- [0595] tetO 七聚体 3679-3932
- [0596] 最小 CMV 启动子 (PminCMV) 3946-4005
- [0597] BM40 前导序列 4016-4066
- [0598] 抗 PD1 MK3475 VL 4068-4413
- [0599] 人类 Ck4421-4738
- [0600] 弗林蛋白酶 (Furin) 裂解位点 4745-4756
- [0601] P2A 肽 4757-4816
- [0602] 具有内含子 4829-4960 的小鼠 VH 前导序列
- [0603] 抗 PD1 MK3475 VH 4962-5321
- [0604] 优化的人类 IgG1 CH1-CH3 5322-6317
- [0605] Myc 标签 6318-6347

- [0606] PDGFR锚6348-6503
- [0607] BGH聚腺苷酸6553-6767
- [0608] AAVS右同源臂6829-7666
- [0609] 3' β-球蛋白隔离子7706-7777
- [0610] f1复制起点7824-8237
- [0611] pUC复制起点8458-9132
- [0612] 卡那霉素抗性基因9852-10646
- [0613] 图29.诱导型哺乳动物展示表达
- [0614] 直方图展现经pINT18-Tet1-377\_1\_C08和TALE核酸酶共转染的HEK293细胞系和在杀稻瘟菌素存在下历时20天所选择的稳定细胞群的染色结果。将样品分成20mL每毫升含有 $5 \times 10^5$ 个细胞,且用20ng/mL、2ng/mL和0ng/mL多西环素诱导。诱导后的第24小时,使用来自多西环素诱导的每个样品的 $1 \times 10^6$ 个细胞进行流式染色。细胞使用抗Fc-PE和TOPRO-3活力染色剂染色。直方图显示FL2通道(抗Fc-PE)上的荧光强度,其相对于细胞计数作图。HEK293 WT对照(灰色实线)、经0ng/mL多西环素(黑色虚线)、2ng/mL多西环素(黑色虚线)和20ng/mL多西环素(黑色实线)诱导的HEK293-pINT18-Tet1-377\_1\_C08稳定细胞系。
- [0615] 图30.细胞展示的抗体对FcRn的结合。
- [0616] 表达布瑞金单抗(Briakinumab)和乌斯坦金单抗(Ustenkinumab)的HEK293细胞使用不同缓冲液用与抗生蛋白链菌素PE(11nM)预偶联的生物素化FcRn(50nM)染色:
- [0617] a.Hek293 WT
- [0618] b.抗生蛋白链菌素PE-对照
- [0619] c.用缓冲液pH6.0染色的细胞
- [0620] d.用缓冲液pH7.4染色的细胞
- [0621] 图31.FACS根据展示水平分离展示抗间皮素IgG的HEK293细胞。
- [0622] 通过核酸酶介导的基因转移,将一群抗间皮素抗体基因整合到HEK293细胞的人类AAVS基因座中。通过抗人类Fc-PE染色,根据抗体展示水平,利用FACS分离出展示抗体的HEK293细胞的多克隆群体。转染后的第16天,使用BD Influx分选仪,基于抗体表达,利用FACS分离出MACS富集的文库。细胞用藻红蛋白(PE)标记的抗Fc染色。通过使用前向散射和染色使分析聚焦于活细胞。 $\lambda_{em}=450/40$ 、 $\lambda_{exc}=355$ 通道中染色呈阳性的细胞(代表着吸收DAPI的非活细胞)加以排除。直方图显示抗Fc-PE的荧光强度(x轴),其代表抗体表达水平,相对于y轴上的细胞计数作图。选用于分析的门标识为P4、P6和P5,分别代表低、中和高水平展示群体。
- [0623] 图32.分别来源于高水平展示组(实线)和低水平展示组(虚线)的两种抗间皮素IgG1抗体克隆的HPLC-SEC。
- [0624] 图33.DNA结合剂的DNA结合和耗竭(使用MACS)。(A) HEK293细胞(实心,灰色)或展示优特金单抗(虚线)、布瑞金单抗(长虚线)和阿玛西单抗(点线)的经生物素化DNA染色、经抗生蛋白链菌素PE检测的HEK293细胞的叠加图;(B)代表三种抗体细胞群的混合物的点阵图,所述细胞群展示优特金单抗(未标记,Q4)、阿玛西单抗(经CellTrace远红标记,X轴)和布瑞金单抗(经CellTrace CFSE标记,Y轴),在MACS分选之前经DNA染色;(C)显示DNA结合剂耗竭的流过物点阵图。

[0625] 图34.用肝素-FITC(x轴)和抗人类Fc APC(y轴)双重染色。(a)显示优特金单抗(灰色)和布瑞金单抗(黑色)重叠图的点阵图。(b)显示优特金单抗(灰色)和加尼图单抗(黑色)重叠图的点阵图。重叠图内的门显示细胞是肝素的高表达者和非结合者。

[0626] 图35.伴侣蛋白与DyLight 633的偶联物(x轴)和抗人类Fc PE(y轴)的双重染色。(a)显示优特金单抗(灰色)和布瑞金单抗(黑色)的重叠图的点阵图,其经Hsp70-DyLight 633和抗人类Fc PE双重染色。(b)显示优特金单抗(灰色)和布瑞金单抗(黑色)的重叠图的点阵图,其经Hsp90-DyLight 633和抗人类Fc PE双重染色。重叠图内的门显示细胞是伴侣蛋白的高表达者和非结合者(Hsp70和Hsp90)。(c和d)重叠直方图显示冷自鲁单抗(lenzilumab)和贝伦妥单抗(brentuximab)分别结合Hsp70和Hsp90。

[0627] 图36.经抗人类Fc PE和各种多反应性探针染色的抗体的直方图。通过核酸酶介导的基因整合来建立展示抗体选择的稳定单克隆HEK293细胞系。图示为在HEK293细胞表面上利用以下探针所展示的不同抗体的细胞计数(y轴)相对于荧光强度(x轴)的直方图:(a)抗人类Fc-PE、(b)使用抗生蛋白链菌素PE检测的生物素化DNA、(c)肝素-FITC、(d)抗生蛋白链菌素PE、(e)Hsp70-DyLight 633、(f)Hsp90-DyLight 633和(g)FcRn与抗生蛋白链菌素PE的预偶联物。

[0628] 图37.pINT17-Tet-D1.3,一种诱导型抗体IgG哺乳动物展示表达载体。图示为AAVS同源臂与无启动子杀稻瘟菌素基因之间从Bg1II到BstZ171限制位点的充分标注核酸序列。编号始于Bg1II限制位点。下文列出了关键特征。

[0629] BGH聚腺苷酸223-9(反向链)

[0630] 人类Ck544-236(反向链)

[0631] D1.3 VL 877-549(反向链)

[0632] 具有内含子1168-883的人类VL前导序列(反向链)

[0633] TRE3G启动子1230-1618

[0634] CMV启动子1237-1809

[0635] 具有内含子1644-1782的VH前导序列

[0636] D1.3 VH 1783-2127

[0637] IgG1 CH1-3 2125-3120

[0638] Myc标签3121-3150

[0639] PDGFR锚3151-3306

[0640] BGH聚腺苷酸3356-3570

[0641] pEF启动子3621-4955

[0642] rtTA-3G 5063-5809

[0643] SV40聚腺苷酸5832-6274

[0644] 图38.诱导型IgG哺乳动物展示细胞系。1549\_02\_D06(1)、1535\_01\_E03(2)和337\_1\_C08(3)、博可珠单抗(4)、884\_01\_G01(5)、5A10i(6)和阿利库单抗(7)。转染后第27天,通过添加0ng/ml(a)、2ng/ml(b)、4ng/ml(c)或100ng/ml(d)多西环素来诱导细胞系。诱导后第24小时,细胞用抗Fc-PE染色。荧光强度(抗Fc,FL2,x轴)相对于细胞计数绘制的直方图。

[0645] 图39.诱导型IgG哺乳动物展示细胞系:细胞表面IgG代谢

[0646] 通过AAVS TALE核酸酶介导的基因整合和杀稻瘟菌素选择,使用具有抗PD1抗体:

1549\_02\_D06(1)、1535\_01\_E03(2)和337\_1\_C08(3)以及抗PCSK9抗体博可珠单抗(4)、884\_01\_G01(5)、5A10i(6)和阿利库单抗(7)的VH和VL的pINT17-Tet建立稳定HEK293细胞系。转染后的第27天,通过添加100ng/ml多西环素来诱导细胞系。诱导后第48小时,细胞用抗Fc-PE染色。荧光强度的直方图(抗Fc,FL2,x轴)相对于细胞计数绘示。

[0647] 图40.展示抗PD1抗体1549\_02\_D06(对于PD-1, $K_D=2.9\text{nM}$ )和337\_1\_C08(对于PD-1, $K_D=74\text{nM}$ )的细胞系分别用(a)0ng/ml、(b)2ng/ml、(c)4ng/ml和(d)100ng/ml多西环素诱导。显示了FL1通道中的荧光(y轴)相对于前向散射(FSC,x轴)的点阵图。展示1549\_02\_D06的经标记的细胞显示于每个点阵图的上象限中并且展示337\_1\_C08的未标记细胞显示于下象限中。图i、ii、iii和iv展现了MACS提纯分别使用的0.1nM、1nM、10nM浓度的PD-1-生物素。图iv展现输入的MACS前群体。每个细胞群的百分比显示于每个象限内。

[0648] 图41.优特金单抗(灰色)和布瑞金单抗(黑色)的双重染色群体的重叠点阵图。50nM FcRn-Avi标签与抗生蛋白链菌素PE(x轴)和抗人类Fc APC(y轴)的预偶联物的双重染色。图内的门代表作为FcRn非结合剂的优特金单抗,其能够通过FACS从FcRn结合剂分选出来。

[0649] 图42.抗间皮素重链可变(VH)域抗体群体的生殖系分析。图绘制了每种VH生殖系的输入和哺乳动物低、中和高展示门控群体的出现频率。

[0650] 图43.抗间皮素轻链可变 $\kappa$ (VL $\kappa$ )域抗体群体的生殖系分析。图绘制了每种VL $\kappa$ 生殖系的输入和哺乳动物低、中和高展示门控群体的出现频率。

[0651] 图44.抗间皮素轻链可变 $\lambda$ (VL $\lambda$ )域抗体种群的生殖系分析。图绘制了每种VL $\lambda$ 生殖系的输入和哺乳动物低、中和高展示门控群体的出现频率。

[0652] 图45.抗间皮素IgG1抗体的HPLC-SEC。通过瞬时转染Expi-293细胞来表达抗体,通过蛋白质A色谱进行亲和力提纯且透析。使用Agilent 1100HPLC仪器,以0.35ml/min的流量将样品( $2\mu\text{l}, 1\text{mg/ml}$ )装载于Agilent AdvancedBio SEC 300A, 2.7um, 4.6x300 mm柱(安捷伦科技,目录号PL1580-5301)。图示为以下所选抗间皮素抗体在215nm的吸光度相对于滞留时间的图:来源于高水平展示组的932\_01\_A03(黑线);和来源于低水平展示组的930\_01\_A12(交替的点和虚线)、930\_01\_B02(长虚线)、930\_01\_C12(短虚线)。

[0653] 图46.人类和CHO AAVS内含子1TALE-核酸酶(TALEN)靶标结合位点的比对。CHO AAVS内含子1DNA序列获自ENSEMBL标注的CHO-K1谷氨酸盐合成酶(GS)基因剔除细胞系寄存号:CHOK1GS\_HDv1:scaffold\_52:2374828:2406177:1。编号提及人类PPP1R12C内含子1起点。粗体指示人类TALEN靶点的左臂和右臂。星号指示人类与CHO序列之间的同源性并且虚线(-)指示缺失。下划线和斜体指示pINT17靶向载体内的AAVS左同源臂和右同源臂的末端。这种比对用于设计pINT17-CHO靶向载体内的CHO AAVS同源臂并且为了比较,可用于设计CRISPR/Cas9向导RNA。编号1到3的有义或反义CRISPR向导RNA识别位点分别显示于序列上方或下方。

[0654] 图47.载体pINT17-BSD-CHO(双重启动子抗体IgG表达盒,用于CHO细胞上的表面表达)内的CHO AAVS同源臂。图示为载体内的左和右CHO AAVS同源臂的标注DNA序列。所有其余特征,包括此图中未示出的那些特征,包含双重启动子抗体表达盒,都如关于载体pINT17-BSD所述(图1)并且列举如下

[0655] 特征:

- [0656] CHO AAVS左同源臂9-899
- [0657] 杀稻瘟菌素抗性基因942-1343
- [0658] pEF启动子1611-2794
- [0659] BM40前导序列2834-2885
- [0660] 人源化D1.3 VL 2888-3219
- [0661] 人类Cκ3227-3532
- [0662] BGH聚腺苷酸3468-3682
- [0663] CMV启动子3790-4362
- [0664] 具有内含子4388-4515的小鼠VH前导序列
- [0665] 人源化D1.3 VH 4521-4868
- [0666] 优化的人类IgG1 CH1-CH3 4869-5864
- [0667] Myc标签5865-5894
- [0668] PDGFR锚5895-6050
- [0669] BGH聚腺苷酸6100-6314
- [0670] CHO AAVS左同源臂6376-7266
- [0671] f1复制起点7424-7837
- [0672] pUC复制起点8058-8732
- [0673] 卡那霉素抗性基因9452-10246
- [0674] 图48.通过TALEN或CRISPR/Cas9核酸酶介导的基因整合使抗体展示于CHO细胞表面上。荧光强度(抗Fc, FL2, x轴)相对于细胞计数绘制的直方图。图(从上到下)是CHO对照、pINT17-BSD-CHO V2-尼沃单抗减核酸酶、pINT17-BSD-CHO V1-尼沃单抗减核酸酶、pINT17-BSD-CHO V1-尼沃单抗加CHO TALEN、pINT17-BSD-CHO V1-尼沃单抗加CRISPR3、pINT17-BSD-CHO V1-尼沃单抗加CRISPR2、pINT17-BSD-CHO V1-尼沃单抗加CRISPR1。
- [0675] 图49.抗体在CHO表面上的展示水平。荧光强度(抗Fc, FL2, x轴)的直方图,其相对于CHO细胞的细胞计数绘制(实心图):(a)博可珠单抗(实线)和884\_01\_G01(虚线)。(b)MEDI-1912(实线)和MEDI-1912-STT(虚线)。
- [0676] 图50.利用哺乳动物展示来建立“可开发性增强的”群体用于随后的结合选择
- [0677] a.抗PD1 337\_1\_C08 VH(i)和VL(ii)链的序列。核苷酸序列图示在密码子上方有翻译单字母氨基酸码。对CDR进行标注(加有下划线)并且以粗体突出显示的CDR3氨基酸经历定点诱变。
- [0678] b.基于抗体在细胞上的高、中和低展示水平,通过FACS分离抗PD1抗体VH和VL CDR3哺乳动物展示文库且通过用抗Fc-PE染色、通过分析型流式细胞术加以分析。图示(从上到下)为细胞计数(y轴)相对于Fc表达(x轴)的直方图:高(i)、中(ii)和低(iii)抗PD1群体。作为参考,图示为起始抗Fc MACS群体(iv)、“野生型”337\_1\_C08亲本克隆(v)和不展示抗体的HEK293细胞(vi)。
- [0679] 图51.pINT17-Bi-CMV-艾密次单抗(Emicizumab),一种含有双向CMV和延伸因子(pEF)启动子的质体,用于细胞表面表达双特异性“臼包杵”共同轻链IgG艾密次单抗。这是一种三顺反子靶向载体,其具有驱动三种基因表达的三种启动子:抗FIXa重链、抗FX重链和共同轻链。图示为AAVS同源臂之间从BgIII到BstZ171限制位点的充分标注的核酸序列。编

号始于Bg1II限制位点。下文列出了关键特征。

- [0680] BGH聚腺苷酸222-8(反向链)
- [0681] 人类C $\kappa$ 546-232(反向链)
- [0682] 艾密次单抗VL 876-547(反向链)
- [0683] 具有内含子1143-884的人类VL前导序列(反向链)
- [0684] 最小CMV启动子1230-1167(反向链)
- [0685] CMV启动子1237-1809
- [0686] 具有内含子1835-1973的小鼠VH前导序列
- [0687] 艾密次单抗抗FIXa VH 1974-2339
- [0688] 艾密次单抗抗FIXa CH1-3 2340-3317
- [0689] Myc标签3318-3347
- [0690] PDGFR锚3348-3503
- [0691] BGH聚腺苷酸3553-3767
- [0692] pEF启动子3818-5152
- [0693] 具有内含子5260-5401的人类VH前导序列
- [0694] 艾密次单抗抗FX VH 5402-5758
- [0695] 艾米珠单抗抗FX CH1-3 5759-6733
- [0696] Myc标签6734-6763
- [0697] PDGFR锚6764-6916
- [0698] SV40聚腺苷酸6942-7384
- [0699] 图52.FIXa和FX对HEK293细胞表面上所展示的双特异性抗体艾密次单抗的结合
- [0700] pINT17-Bi-CMV-艾密次单抗或pINT17-BSD-抗FIXa用于在编码AAVS TALEN的质体存在下转染HEK293细胞。转染后第24小时,分析细胞的抗体展示和结合抗原FIXa或FX的能力。直方图描绘了细胞计数相对于荧光强度,从左到右,其使用以下各物染色:抗Fc-APC、FX-生物素或FIXa-生物素,其与抗生蛋白链菌素-PE预偶联,或单独的抗生蛋白链菌素-PE,用于HEK293细胞展示(a)双特异性艾密次单抗(黑色虚线)、(b)抗FIXa IgG(实心黑线)、(c)HEK293细胞。
- [0701] 图53.艾密次单抗VL与亲本艾密次单抗VL的比对。
- [0702] CDR由序列上方的条柱指示。点指示与最终艾密次单抗VL的一致性。促成正电荷补丁的残基用粗体突出显示。
- [0703] 图54.HEK293细胞表面上的打结抗体展示与其生物物理学特性之间的关系
- [0704] (a)HEK293细胞在AAVS TALEN存在下用pINT17-打结抗体转染。利用杀稻瘟菌素选择稳定群体。转染后的第7天,细胞用抗Fc PE (FL2) 染色且通过流式细胞术分析。直方图描绘了细胞计数(y轴)相对于荧光强度(抗Fc-PE,x轴)的图且包括KB\_A12\_EETI-II(黑色实线)、KB\_A12\_Hstx1(虚线)和KB\_A12\_ProTxIII(虚线)。KB\_A12\_Hstx1(虚线)和KB\_A12\_ProTxIII(虚线)的迹线重叠并且似乎合并了。
- [0705] 打结抗体是通过瞬时转染Expi293细胞来表达并且通过蛋白质A亲和色谱来提纯。打结抗体如上文所述通过HPLC-SEC分析并且显示了(b)KB\_A12\_EETI-II、(c)曲妥珠单抗(Trastuzumab)和(d)KB\_A12\_ProTx-III的吸光度相对于洗脱体积的图。

[0706] 图55.打结抗体的突变体文库含有展示水平与亲本打结抗体相比改善的克隆。展示打结抗体的HEK293细胞用抗Fc-PE染色并且通过流式细胞术分析。图示为三个文库(抗Fc MACS提纯之后)的细胞计数相对于荧光强度与其相关亲本打结抗体对照展示细胞系相比的直方图。(从左到右):(a) KB\_A12 ProTxIII文库组A(虚线)与KB\_A12 ProTxIII对照(实心线)、(b) KB\_A12 ProTxIII文库组B(虚线)与KB\_A12ProTxIII对照(实心线)、(c) KB\_A12 HsTx1文库(虚线)与KB\_A12 HsTx1对照(实心线)。

[0707] 实例

[0708] 实例1.构建靶向载体用于IgG格式化抗体的可溶性表达和细胞表面展示

[0709] 为了能够使结合剂分子(包括抗体)展示于高等真核细胞表面上和其后续的基因选择,可以利用载体将结合剂基因靶向宿主基因组中的特定位置。载体可以编码可选标记物以实现稳定细胞系的选择并且这种可选标记物可以编码赋予针对杀稻瘟菌素、G418/遗传霉素(Genticin)、潮霉素(hygromycin)、嘌呤霉素(puromycin)或博莱霉素(zeocin)的抗性的基因。靶向载体可以含有驱动编码可选标记物的基因表达的外源启动子。或者,可以将转基因整合于细胞DNA中位于内源启动子下游的位置,以使正确整合的转基因能够被优先选择。靶向载体也将编码同源臂以允许相关染色体基因座和启动子发生同源重组,从而驱动结合剂分子和聚腺苷酸化(pA)位点的表达。使结合剂分子基因与编码前导序列的DNA融合,以允许经由内质网(ER)分泌到细胞表面和膜锚,例如跨膜域或糖基化磷脂酰肌醇(GPI)锚。

[0710] 在此使用的靶向载体的示意图显示于图1a中并且充分标注的DNA序列显示于图1b中。质体包括侧接表达盒的AAVS同源臂,以允许转基因在人类AAVS位点同源重组。AAVS基因座最初鉴定为腺相关病毒的共同整合位点并且是异源基因在人类细胞中插入和表达的“安全港”基因座<sup>93</sup>。核酸酶介导在AAVS位点内裂解之后,靶向载体中的AAVS同源臂通过同源重组促进表达盒整合。杀稻瘟菌素基因在载体内缺乏启动子,但其前面存在剪接受体,所述剪接受体与AAVS基因座的上游外显子建立同框融合。抗体重链和轻链表达盒的细节描述于下文和图1中。

[0711] 利用此前所述的载体(WO2015166272A2)、通过添加实现其后续组装的限制位点、通过聚合酶链反应(PCR)扩增所选片段来构建靶向载体pINT17-BSD(图1a和1b)。pINT17-BSD的各种元件的起点现已描述。编码AAVS-左同源臂、剪接受体、杀稻瘟菌素抗性基因、聚腺苷酸化位点和延伸因子1 $\alpha$ 启动子(pEF1 $\alpha$ )的DNA来源于pD2质体(WO2015166272A2),其通过添加5' AsiSI和3' Bg1II限制酶、通过PCR扩增(1511bp)而产生。利用pD2质体、通过添加5' IgG1 CH3同源序列和3' -HindIII位点对编码Myc标签和PDGFR跨膜域的DNA进行PCR扩增。利用此前描述的pINT3质体(WO2015166272A2)、通过添加5' Bg1II位点和3' 的编码Myc标签的DNA对编码以下的DNA进行PCR扩增:轻链BM40前导序列、抗溶菌酶抗体D1.3的可变轻链(VL)<sup>94</sup>、人类恒定轻链(CL)、牛生长激素(BGH)pA、即刻早期细胞巨大病毒启动子(CMV启动子)、被内含子分离的小鼠重链前导序列、抗溶菌酶抗体D1.3的可变重链(VH)和IgG1抗体重链恒定域1到3(IgG1 CH1-3)。通过PCR组装将编码pINT17-BSD的4446bp区域(从Bg1II到HindIII)的两个片段组合,以将PDGR跨膜域直接添加到CH3末端。利用pD2质体(WO2015166272A2)、通过添加HindIII和SbfI限制位点、对编码AAVS右同源臂的区域(从HindIII到SbfI)进行PCR扩增(1168bp)。载体主链从来源于pSF-EF1 $\alpha$ (Oxford Genetics OG43)的SbfI到AsiSI位点包

含f1和pUC复制起点以及卡那霉素抗性基因。图1b中所示的实例编码人类抗溶菌酶抗体的VL和VH,但使用标准分子生物学技术(例如使用侧接限制酶置换VL和VH基因)能够方便地用其取代其它特异性。

[0712] 实例2.比较3对抗体在亲本克隆和可开发性增强的克隆上的表面呈现水平

[0713] 我们检查了三对抗体,其中原始亲本抗体具有不良的可开发性特征和其再经工程改造的子系分子,所述子系分子经改变可改善其自身相互作用和交叉相互作用特性。所述组包括CNT0607(针对介白素IL-13的单克隆抗体)和其经修饰的对应物CNT0607 W100A<sup>14</sup>。CNT0607在中性pH是难溶的,在高浓度下沉淀于PBS缓冲液中并且显示自身相互作用,如在亲和力捕捉自身相互作用纳米颗粒光谱法(AC-SINS)分析中所测量<sup>39</sup>。CNT0697的结构测定揭露了重链CDR3中的疏水性补丁。VH CDR3突变W100A使其抗体溶解度与交叉相互作用色谱(CIC)特征均得到改善<sup>47</sup>。CIC测量了对固定于柱基质上的人类血清多克隆抗体的结合。第二个实例是单克隆抗体,命名为Ang2mAb,其靶向血管生成素2(Tie2受体的可溶性配体和病理性血管生成的调节因子)。然而,Ang2mAb据报道展现不良的表达和聚集。对19种变异体进行结构构建模和实验筛选的组合能够对更佳的表达Ang2mAb C49T<sup>6</sup>进行工程改造,其使不成对的半胱氨酸残基发生突变。最终,我们纳入MEDI-1912,一种抗神经生长因子(NGF)抗体,其通过TrkA和p75受体抑制信号传导<sup>7</sup>。MEDI-1912可以潜在地用于治疗慢性疼痛,但在溶液中显示沉淀和聚集以及不良的药物动力学特征。MEDI-1912以皮摩尔浓度亲和力结合到NGF并且由命名MEDI-578的“祖父级”抗体进行亲和力成熟,所述“祖父级”抗体在自聚集方面表现良好。通过氢/氘交换-质谱学(HDX-MS)和分子建模,鉴定出由VH CDR1和CDR2内的残基产生的VH域上存在疏水性补丁。由此允许预测负责自缔合和随之发生的聚集的氨基酸。这继而能够设计出具有突变W30S、F31T和L56T的三重突变体(MEDI-1912\_STT),所述突变使自身相互作用界面中断,同时保持针对NGF的效能和亲和力<sup>7</sup>。

[0714] 将编码CNT0607、CNT0607-W100A、Ang2mAb、Ang2mAb-C49T、MEDI-1912和MEDI-1912\_STT重链和轻链可变域的合成DNA(序列参见表1)克隆到哺乳动物展示载体pINT17-BSD中(载体图谱和序列参见实例1),证实DNA序列并且制备转染质量的质体DNA。转染前一天,经悬浮液调适的HEK293细胞以 $5 \times 10^5$ 个细胞/毫升接种于HEK FreeStyle 293表达培养基中。当细胞在10ml中达到 $1 \times 10^6$ 个细胞/毫升的密度时,进行PEI转染。将具有pINT17的抗体基因(1ug)、左和右TALEN质体(各5ug)混合且在未补充的HEK FreeStyle 293表达培养基(1ml)中稀释。添加25000Da MW线性聚乙烯亚胺(PEI)(10ul, 1mg/ml, Polysciences),在室温下培育10分钟。然后将质体DNA/PEI混合物添加到HEK293悬浮细胞( $1 \times 10^6$ 个细胞/毫升,存在于10ml HEK FreeStyle 293表达培养基中)中。以7 $\mu$ g/ml的浓度转染之后的第48小时开始杀稻瘟菌素选择。在实验期间,所述群体保持在选择下。转染后(dpt)的第15天之后,用抗人类Fc PE(Biolegend)将细胞染色。然后通过以下方案将展示抗体的单克隆细胞系染色。使展示抗体的HEK293细胞系或野生型HEK293细胞(一百万个细胞)集结(200g, 3分钟, 在微量离心管(Eppendorf tube)(1.5ml)中)。使集结粒再悬浮于PBS(1ml)中且离心(600g, 2.5分钟)。使集结粒再悬浮于含有抗Fc PE(5 $\mu$ l, Biolegend)的1%BSA/PBS(100 $\mu$ l)中。混合物在4°C下避光培育30分钟。添加0.1%BSA/PBS(900 $\mu$ l)并且使细胞集结(600g, 2.5分钟)。将细胞再悬浮于0.1%BSA/PBS(1ml)中,且重复此洗涤步骤一次。将细胞再悬浮于含有7-AAD(5 $\mu$ l/百万个细胞)的0.1%BSA/PBS(200 $\mu$ l)中。使用Intellicyte iQue筛选仪分析经标

记的细胞(50 $\mu$ l)。流式细胞术分析(图2)显示所有三对抗体的经改善的子系分子的改善展示水平与存在问题的原始亲本分子相比存在增加的抗体展示水平。

链	蛋白质序列
[0715]	Ang2 VH QVQLVESGGGVVQPGRSLRLSCAASGFTFTNYGMHWGRQAPGKGLEWVAVI SHDGNN KYYVDSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAREGIDFWSGLNWFDPWG QGTLTVSS
	Ang2 VL EIVLTQSPGTLSLSPGERATLSCRASQSITGSYLAWYQQKPGQAPRLLICGASSWAT GIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYSSSPITFGQGTRLEIK
	Ang2 VL C49T EIVLTQSPGTLSLSPGERATLSCRASQSITGSYLAWYQQKPGQAPRLLI <b>T</b> GASSWAT GIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYSSSPITFGQGTRLEIK
	CNT607 VH QVQLVESGGGLVQPGGSLRLSCAASGFTFNSYWINWVRQAPGKGLEWVSGIAYDSSN TLYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARGLGAFHWDMQPDYWGQ GTLTVSSAS
	CNT607 VH W100A QVQLVESGGGLVQPGGSLRLSCAASGFTFNSYWINWVRQAPGKGLEWVSGIAYDSSN TLYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCARGLGAF <b>H</b> ADMQPDYWGQ GTLTVSS
	CNT607 VL SYELTQPPSVSVPQTARISCSGDNIGGTFVSWYQQKPGQAPVLVIYDDNDRPSGI PERFSGSNSGNTATLTISGTQAEDADEYYCGTWDMVTNNVFGGGTKLTVL
	MEDI-1912 VH QVQLVQSGAEVKKPGSSVKVSCKASGGTFWFGAFTWVRQAPGQGLEWMGGI IPIFGL TNLAQNPFQGRVTITADESTSTVYMELSSLRSEDTAVYYCARSSRIYDLNPSLTAYYD MDVWGQGTMVTVSS
	MEDI1912 VH STT QVQLVQSGAEVKKPGSSVKVSCKASGGTF <b>ST</b> GAFTWVRQAPGQGLEWMGGI IPIF <b>G</b> T TNLAQNPFQGRVTITADESTSTVYMELSSLRSEDTAVYYCARSSRIYDLNPSLTAYYD MDVWGQGTMVTVSS
	MEDI-1912 VL QSVLTQPPSVAAPGQKVТИSCSGSSSDIGNNYVSWYQQLPGTAPKLIYDNNKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDSL SAWVFGGGTKLTVL
	韦森单抗 VH EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSQISPAGGY TNYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGELPYRMSKVMDVWG QGTLTVSS
	韦森单抗 VL DIQMTQSPSSLSASVGDRVТИCRASQYFSSYLAWYQQKPGKAPKLLIYGASSRASG VPSRFSFGSGSGTDFTLTISSLQPEDFATYYCQQYLGSPPTFGQGTKVEIK

[0716] 表1.测试抗体的VH和VL基因的蛋白质序列.显示了可变抗体重链和轻链的氨基酸序列(单字母码)。可变域加有下划线。抗体对之间的变异残基以粗体突出显示。

[0717] 实例3a.高浓度下的细胞表面呈现水平与自身相互作用之间的关系

[0718] 为了检查实例2中描述的抗体的特性,进行了抗体表达和提纯。将编码CNT0607、CNT0607-W100A、Ang2mAb、Ang2mAb-C49T、MEDI-1912和MEDI-1912\_STT重链和轻链可变域的合成DNA(序列参见表1)克隆到基于pINT3的双重启动子IgG可溶性表达载体(WO2015166272A2)中并且通过DNA测序来证实克隆正确。

[0719] 制备质体DNA并且使用所述质体DNA、使用转染试剂ExpiFectamine、根据制造商说明书(A14525,赛默飞世尔科学(ThermoFisher Scientific))转染Expi293细胞(30ml最终培养体积规模)。转染之前的第24小时,以 $2 \times 10^6$ 个细胞/毫升的密度接种存在于25.5ml Expi293表达培养基中的细胞。使质体DNA(30 $\mu$ g)在Opti-MEM培养基(1.5ml)中稀释并且使

ExpiFectamine 293试剂(80 $\mu$ l)在Opti-MEM培养基(1.5ml)中稀释并且在室温下培育5分钟。然后将稀释的质体DNA(30 $\mu$ g,存在于1.5ml Opti-MEM培养基中)添加到稀释的ExpiFectamine 293试剂(80 $\mu$ l ExpiFectamine存在于1.5ml Opti-MEM培养基中)中并且在室温下培育20分钟。将细胞在37°C、5%CO<sub>2</sub>、5%湿度下培育且以130rpm搅拌(25mm回转摆幅,ISF1-X,Climo-Shaker,Kuhner)。表达5天后,通过离心(2000g,20分钟)收获培养物上清液。

[0720] 50ml离心管中的培养物上清液通过添加1/10体积的PBS(pH 7.4)进行pH调节并且添加蛋白质A琼脂糖FF树脂(300 $\mu$ l,Generon,PC-A100)并且在室温下搅拌培育1小时。50ml管以2000g离心5分钟以收集珠粒并且丢弃上清液,得到约1ml珠粒浆液。将此浆液再悬浮,装载于玻璃料柱上(Proteus“1步分批式”midi离心柱,Generon,GEN-1SB08),离心(50g,1分钟,4°C)且丢弃流过物。柱用2×PBS(2×10ml)洗涤,随后在每个洗涤步骤之后离心(50g,1分钟,4°C)。使用洗脱缓冲液(900 $\mu$ l,0.2M甘氨酸,pH 2.6)洗脱抗体,添加到柱基质中且立即使用中和缓冲液(300 $\mu$ l,1M Tris-HCl,pH 8)中和洗脱液。然后通过离心(50g,1分钟,4°C)将抗体从蛋白质A琼脂糖柱直接洗脱到中和缓冲液中。抗体通过转移到GeBAflex最大管(8kDa分子量截止值,Generon,D045)中进行缓冲交换并且在4升PBS中透析并且在4°C下培育至少3-18小时。利用第二个4L PBS透析步骤重复这个透析步骤。通过测量280nm吸光度和使用1.4的估测消光系数估算浓度、利用比尔-兰伯特定律(Beer-Lambert Law)计算来测定抗体产量和浓度。

[0721] 通过瞬时表达产生的多肽产量可以被视为可开发性潜力指标。虽然在实例2的3对抗体之间比较瞬时转染的表达产量显示亲本抗体的表达较低,但是通过比较瞬时转染的产量不能明显简单地看出可开发性潜力的显著差异(表2)举例来说,亲本CNT0607抗体的表达产量是34mg/L,而溶解度改善的CNT0607-W100A抗体表达产量是55mg/L。类似地,亲本抗体MEDI-1912的表达产量是33mg/L,相比之下,改善版本的表达产量是53mg/L。Ang2mAb产量是13mg/L,相比之下,工程化后代Ang C49T的产量是34mg/L。

[0722] 多肽的解链温度有时被当作预测“可开发性”的替代指标并且在一些情况下,已经选择解链温度改善的抗体以期望产生较大可开发性抗体<sup>9-11</sup>。使用PrometheusNT.4B(Nanotemper),根据制造商说明书测定解链温度(T<sub>m</sub>)和聚集起始温度(Tagg)。使用小毛细管吸收约8-10 $\mu$ l的0.5mg/ml抗体溶液。然后将毛细管夹持就位,以便通过Prometheus仪器进行荧光扫描用于热解链分析。利用Prometheus拟合软件测定解链起始温度和散射起始温度。抗体的解链温度(T<sub>m</sub>)与聚集温度(Tagg)相似(参见表2),均在抗体对组之间并且与临幊上批准的阳性对照抗PD1抗体尼沃单抗相比较。这个数据表明这个抗体组的解链温度和聚集温度不能预测自身相互作用和非特异性交叉相互作用的生物物理学特征。

[0723] 在制备型尺寸排阻色谱期间,MEDI-1912显示的洗脱特征比MEDI-1912\_STT更早(图3),表明其作为较高分子量物质存在并且倾向于自身相互作用。其余抗体的洗脱特征与尼沃单抗相似。为了通过动态光散射(DLS)能够测量抗体自身相互作用,利用超过滤来浓缩尺寸提纯的抗体。每个抗体对所实现的抗体浓度显示于表2中。这揭露了在抗体沉淀发生而堵塞超滤膜之前,将亲本抗体MEDI-1912和CNT0607分别浓缩而超过1.4mg/ml和1.8mg/ml是不可能的。相比之下,将溶解度增强的子系分子MEDI-1912\_STT和CNT0607\_W100A分别浓缩到29mg/ml和30mg/ml而无沉淀证据是有可能的。对于浓缩的Ang2mAb对来说,未观察到沉

淀。动态光散射 (DLS) 检测亲本抗体MEDI-1912和CNT0607的高阶聚集物质(表2),而子系分子MEDI-1912\_STT和CNT0607\_W100A则检测不到,如根据计算的多分散指数(PDI)和累积量(或z-平均值)大小所判断。举例来说,亲本CNT0607和MEDI 1912的PDI分别是0.22和0.15,而子系分子的PDI分别是0.1和0.12,表明单分散状态较均匀(表2)。类似地,亲本MEDI-1912的平均粒度分别是22nm,而子系分子MEDI-1912-STT的平均粒度是13nm,表明低阶聚集状态(表2)。因此,出现显著的自身相互作用,从而在较低浓度下发生可检测的自身相互作用并且在较高浓度下发生沉淀。

克隆ID	抗体	T <sub>m</sub> ( °C )	T <sub>agg</sub> ( °C )	表达产量(mg/L)	C-Max(mg/ml)	Zav(nm)	PDI
[0724]	1 CNT0607 -亲本	62. 6	75. 6	33. 6	1. 8	16. 2	0. 22
	2 CNT0607 (W100A)	61. 4	75. 4	55. 2	>30	14. 8	0. 1
	3 MEDI-1912-亲本	71. 5	73. 2	33. 2	1. 4	21. 5	0. 15
	4 MEDI-1912 (STT)	70. 8	74. 8	52. 8	>29	12. 5	0. 05
	5 Ang2mAb-亲本	63. 8	65. 0	13. 4	>21	13. 1	0. 12
	6 Ang2mAb C49T	66. 5	65. 2	34. 4	>18	12. 7	0. 09
	14 尼沃单抗	67. 6	67. 3	102. 8	>50	nd	nd

[0725] 表2.IgG生物物理学特性.使用Prometheus NT.4B(Nanotemper),根据制造商说明书测定解链温度(Tm)和聚集起始温度(Tagg)。通过瞬时转染30ml规模的Expi293细胞(赛默飞世尔)来测定表达产量(按照每升培养体积所表达的抗体数量(mg)),随后进行亲和纯化(蛋白质A)并且利用280 nm吸光度和1.4的估测抗体消光系数测定经提纯的抗体的产量。在Superdex 200 10/300上使用AKTA Pure系统、使用PBS(pH 7.4)操作缓冲液、通过尺寸排阻色谱进一步提纯抗体。使用Nano SDLS(马尔文仪器公司,英国马尔文)对样品进行动态光散射测量并且使用zetasizer软件(马尔文仪器公司,英国马尔文)计算多分散指数(PDI)和累积量(或z-平均值)大小(Zav)。

[0726] 这个实例证明就三个不同抗体对来说,哺乳动物细胞展示抗体的水平和其生物物理学特性之间存在非常明显的关系。记录有与自身相互作用、交叉相互作用和不良药物动力学(MEDI-1912)有关的问题的特异性针对IL-13的亲本抗体(CNT0607)、Angiopoietin2(Ang2mAb)和神经生长因子(MEDI-1912)与其溶解度增强的子系分子相比,都产生了较低的细胞展示水平(图2)。

#### [0727] 实例3b.支持细胞表面抗体浓度的理论

[0728] 这个实例展示的潜在推理可以有助于了解当多肽滞留于细胞表面上时真核细胞中的强烈多肽表达能够潜在地实现高局部浓度的本发明方案。

[0729] 在此所述的工作中,经悬浮液调适的HEK293细胞用于抗体展示。HEK293细胞系具有哺乳动物(人类)起源。细胞大致呈球形,半径为10微米<sup>95</sup>。如果我们将悬浮液HEK293细胞当作球体,则我们能够计算出抗体在其表面上所占据的体积。(我们为了这个实例假定细胞表面的所有区域同等可及以便抗体展示。如果不是这样,则实现抗体的较高局部浓度)。球体的半径(r)能够根据公式 $4/3\pi r^3 = 4.18r^3$ 计算。抗体高度取150埃(Å)(=15nm)时,其将以 $4.18(r + 150 \text{ } \text{\AA})^3$ 的较大球体存在。因此,抗体体积是这种球体体积与细胞体积之间的差

值。

- [0730] 10微米半径的细胞体积是：
- [0731]  $4.18 \times 10^{-15} \text{ m}^3$  ( $4180 \times 10^{-15}$  升)。
- [0732] 包括抗体的较大球体的体积是：
- [0733]  $4.198 \times 10^{-15} \text{ m}^3$  ( $4198 \times 10^{-15}$  升)。
- [0734] 因此,所展示的抗体占据的体积是
- [0735]  $0.018 \times 10^{-15} \text{ m}^3$  ( $18 \times 10^{-15}$  升)。
- [0736] 以类似方式,我们能够计算不同尺寸细胞的体积差。
- [0737] 知道抗体/细胞的数目、分子量和所占体积,我们于是能够计算在细胞表面上所达成的浓度。使用阿沃伽罗常数 (Avogadro's constant),我们知道 $6 \times 10^{23}$ 个抗体分子/升将具有 $150,000 \text{ mg/ml}$ 的浓度。因此, $6 \times 10^{18}$ 个抗体分子/毫升将具有 $1.5 \text{ mg/ml}$  ( $10 \mu\text{M}$ ) 的浓度。利用这种方案计算表3中所示的浓度。

细胞半径	5 微米	10 微米	15 微米	20 微米
细胞体积 ( $\times 10^{-15}$ 升)	522.5	4180	14107.5	33440
细胞加抗体所占 的体积 ( $\times 10^{-15}$ 升)	527.2	4198	14150	33515
抗体所占的体积 ( $\times 10^{-15}$ 升)	4.7	18.8	42.3	75
[0738] 绝对数目 / 升 ( $\times 10^{18}$ ), 如果每 个展示 $10^6$ 个	212	53.2	23.64	13.3
浓度( $\text{mg/ml}$ ), 假 定展示 $10^6$ 个抗体	<b>53</b>	<b>13.2</b>	<b>5.9</b>	<b>3.3</b>
浓度( $\text{microM}$ ), 假 定展示 $10^6$ 个抗体	354	88	39	22
提供 $1 \text{ mg/ml}$ 的拷 贝数	<b>19,000</b>	<b>75,000</b>	<b>170,000</b>	<b>300,000</b>

- [0739] 表3.

[0740] 在这些计算中,细胞表面上的抗体数目取为 $10^6$ 个拷贝/细胞。实验测定拷贝数的方法详述本文中别处并且说明于实例7中。

[0741] 我们从表3中发现,10微米半径的细胞(例如悬浮培养的HEK293细胞)上估测展示 $10^6$ 个抗体/细胞,得到超过 $10 \text{ mg/ml}$ 的浓度。在此类浓度下,可能潜在地出现蛋白质自身相互作用的问题。因此,具有聚集倾向的抗体因通过内质网减少<sup>96</sup>和降解增加而在细胞表面上的呈现减少。结果,与非自身相互作用抗体相比,观察到倾向于自聚集的抗体存在较低展示水平。

[0742] 实例4. 基于表面呈现水平富集“可开发性增强”的抗NGF抗体

[0743] 实例3中描述了抗体的生物物理学特性(尤其自聚集)与其哺乳动物细胞展示水平之间的关系。这通过采用三个抗体对来说明,其中原始亲本抗体就自身相互作用和交叉相互作用而言具有不良的生物物理学特性并且这些特性是通过改变所选氨基酸以建立生物

物理学特性改善的子系分子来改善。在所有三种情况下，生物物理学特性改善的子系分子展示于HEK293细胞表面上的水平与存在问题的亲本抗体相比增加。在此实例中，我们证明可以从哺乳动物展示的混合克隆群体中富集生物物理学特性优良的抗体。此研究中选择在小鼠模型<sup>7</sup>中具有自身相互作用特性和不良药物动力学的亲本抗NGF MEDI-1912以及改善的子系MEDI-1912\_STT。进行模型实验，其中我们通过用编码亲本抗体和经修饰的抗体的等量哺乳动物展示质体(实例1)以及编码TALE核酸酶对(其将供体质体导向AAVS基因座)的质体转染HEK293细胞来建立展示MEDI-1912或MEDI-1912\_STT的HEK293细胞的混合群体，如此前所述(WO2015166272A2)。药物选择之后，进行荧光活化细胞分选(FACS)并且基于高抗体呈现水平来选择细胞，通过所选抗体基因的分离和序列分析，我们证明从混合群体中选择性富集生物物理学特性改善的抗体。

[0744] pINT17-MEDI-1912与pINT17-MEDI-1912\_STT(其描述参见实例2)以1:1比率混合并且使用核酸酶介导的基因靶向HEK293细胞而将此混合物整合于HEK293细胞基因组中。通过以200g离心10分钟来收获中对数期HEK293悬浮细胞(生长到 $1 \times 10^6$ 个细胞/毫升的细胞密度)并且以 $1 \times 10^8$ 个细胞/毫升的密度再悬浮于MaxCyte电穿孔缓冲液中。将由pINT17-MEDI-1912(1μg)、pINT17-MEDI-1912\_STT(1μg)、AAVS定向TALEN载体对(各10μg)组成的质体DNA混合物添加到HEK293细胞(100μl，总计 $1 \times 10^7$ 个细胞，存在于MaxCyte电穿孔缓冲液中)并且转移到OC100电穿孔光析管中并且使用MaxCyte STX电穿孔系统进行电穿孔。电穿孔之后，在37°C下回收细胞20分钟，在HEK FreeStyle 293表达培养基中稀释并且在120rpm、37°C、5%CO<sub>2</sub>下维持。以7μg/ml的浓度转染之后的第48小时开始杀稻瘟菌素选择。在实验期间，所述群体保持在选择下。转染后的第15天，细胞如实例2中所述染色，其例外之处是用DAPI染色剂置换7-AAD染色剂，使用NGF-生物素/抗生蛋白链菌素-APC染色剂检测抗原结合并且数量按比例扩大到将1千万个细胞染色。利用流式细胞术分析MEDI-1912/MEDI-1912\_STT混合型HEK293哺乳动物展示群体的抗体呈现水平(图5a)，并且这揭露了展示抗体水平不同的两个主要细胞群。这两个群体与单克隆MEDI-1912和MEDI-1912\_STT抗体展示水平相关，如重叠图所示(图5a)。混合群体通过FACS分选成两个群体：抗体低水平呈现组和高水平呈现组(分别为门5和6，图5b)。

[0745] 由FACS分选的细胞群(门5和6，图5b)制备基因组DNA。使用KOD热启动DNA聚合酶(默克密理博公司(Merck Millipore))，通过巢式PCR扩增编码IgG插入片段的DNA。使用以下基因组特异性引物进行外部PCR：正向：CCGGAACTCTGCCCTCTAAC；和反向：TCCTGGGATACCCCGAAGAG。使用KOD聚合酶(71086，默克)，根据制造商条件，使用外部PCR的PCR产物作为模板，以利用以下引物扩增所整合的IgG插入片段；正向：GAGGGCCTGGATCTTCTTCTC；和反向：GAAGTAGTCCTGACCAGGCA G。依循制造商说明书(20015964, Illumina)对PCR产物进行条形码编码并且使用Illumina MiSeq测序平台进行测序。分析约一百万个读段并且这揭露了根据抗体的高呈现水平(门5，图5b)分选的群体富集MEDI-1912\_STT抗体(96%，图6)。根据抗体低呈现分选的群体富集亲本MEDI-1912抗体(85%，图6)。这个实例证明通过基于细胞表面抗体展示水平选择克隆而从混合群体中选择性富集生物物理学特性改善的由哺乳动物展示的抗体。对此说明的例子是从稳定细胞系的混合群体中富集生物物理学特性比其亲本抗体MEDI-1912优良的MEDI-1912\_STT。

[0746] 我们证明混合群体中能够富集生物物理学特性优良的抗体并且我们表明仅基于

哺乳动物展示的呈现水平便能够从变异体的大型文库中鉴定出改善的抗体。如所论述, VH MEDI1912上的残基W30、F31和L56具有在这种抗体的表面上形成疏水性补丁的潜力<sup>7</sup>。这些残基选用于随机突变并且通过PCR组装诱变由合成DNA模板构建VH文库(引物的氨基酸和核酸序列和位置参见图4)。使用KOD聚合酶(71086-3,默克),根据制造商说明书,由VH模板扩增三种PCR产物:

[0747] a.VH1 (95bp),其使用引物MEDI-1912-F3 (CCATGGCCCAGGTTCAGCT G) 和MEDI1912\_W30NNS\_F31NNS (CTGTCGGACCCATGTAAAGGCGCCSNNNSNAAAGGTGCCGCCGCTTGCTTGCA) 扩增。

[0748] b.VH2 (102bp),其使用引物MEDI-1912-F (GGCGCCTTA CATGGGTCCGACAG) 和MEDI-1912\_L56NNS (CTGGAAGTTCTGGGCCAGA TTGGTSNNGCCGAAGATAGGGATGATGCCGCC) 扩增。

[0749] c.VH3 (213bp),其使用引物MEDI-1912F2 (ACCAATCTGG CCCAGAACTTCCAG) 和MEDI-1912-R (ACTCGAGACGGTGACCATTGTG) 扩增。

[0750] 将上文所列的三种PCR产物(VH1、VH2和VH3)组合(各10ng)并且使用KOD聚合酶(71086-3,默克),根据制造商说明书,使用外部引物MEDI-1912-F3和MEDI-1912-R,通过PCR反应组装。PCR产物用Nco I和Xho I消化并且与经Nco1/Not1消化的pINT17-MEDI-1912(pINT17哺乳动物展示载体(图1),其编码MEDI1912的VL)(100ng)接合。然后使用mini-Elute PCR提纯试剂盒(Qiagen)提纯这种接合混合物并且使提纯的接合混合物在50μl E.cloni 10G elite电感受态细胞(60061-1,Lucigen)中转化。使用0.1cm光析管对细胞进行脉冲,使用2ml回收培养基回收并且在37℃、250rpm下生长1小时。为了计算文库大小,将细胞1:1000稀释并且以10μl和100μl接种于10cm直径的2TY-卡那霉素培养板中。将剩余细胞离心并且接种于2×10cm直径的2TY-卡那霉素培养板中并且在37℃下培育整夜。对10ul培养板中的群落进行计数并且计算 $1.1 \times 10^6$ 的文库大小。由于使用NNS密码子通过使三个残基随机突变而构建的文库编码32,768种变异体,因此实验文库大小超过理论文库大小34倍。刮擦转化体培养板,通过读取600nm吸光度(OD600)来测量细胞密度,使用2个OD单位(2x OD600)的培养等效物接种50ml Circlegrow培养基,在37℃下在250ml有挡板的烧瓶中培养生长3到4小时,收获约400x OD600单位并且制备midiprep质体DNA(pINT17-MEDI-1912-文库)。

[0751] 核酸酶介导基因靶向HEK293细胞使用pINT17-MEDI-1912-文库。通过以200g离心10分钟来收获中对数期HEK293悬浮细胞(生长到 $1 \times 10^6$ 个细胞/毫升的细胞密度)并且以 $1 \times 10^8$ 个细胞/毫升的密度再悬浮于MaxCyte电穿孔缓冲液中。将由pINT17-MEDI-1912-文库(8μg)、AAVS定向TALEN载体对(各40μg)组成的质体DNA混合物添加到HEK293细胞(400μl,总计 $4 \times 10^7$ 个细胞,存在于MaxCyte电穿孔缓冲液中)并且转移到OC100电穿孔光析管中并且使用MaxCyte STX电穿孔系统进行电穿孔。电穿孔之后,在37℃下回收细胞20分钟,在HEK FreeStyle 293表达培养基中稀释并且在120rpm、37℃、5% CO<sub>2</sub>下维持。以7μg/ml的浓度转染之后的第48小时开始杀稻瘟菌素选择。在实验期间,所述群体保持在选择下。转染后的第15天,分析细胞,如实例3中所述染色。对HEK293展示的MEDI-1912-文库进行的流式细胞术分析(图7c)表明所述文库的混合群体内具有展示出抗体展示水平与MEDI-1912\_STT单克隆细胞系等效(图7b)且展示水平高于亲本MEDI-1912单克隆细胞系(图7a)的细胞。这表明就展示水平而言,克隆存在于MEDI-1912群体中等效于MEDI-1912\_STT。

[0752] 根据抗体呈现水平,通过FACS分选文库群体(图7c),从P5和P6门内群体中回收抗

体基因且通过“下一代测序”(NextGen)对VH基因进行测序,如上文所述。图8显示哺乳动物展示所选群体中的残基30、31和56的氨基酸一致性直方图。其显示氨基酸S、T、P在位置31处富集、氨基酸S、P、N在位置32处富集和氨基酸R、S、T和P在位置56处富集。

[0753] 为了能够对哺乳动物展示所选抗体进行生物物理学表征,从所选群体的基因组DNA对VH基因进行PCR扩增(门P5和P6,图7),如实例3中所述。将VH插入片段克隆到具有MEDI-1912VL的pINT17-MEDI-1912、NcoI和XhoI切割载体中,如上文所述,并且使用接合混合物转化大肠杆菌DH10B细胞。挑选188个转型体,制备质体DNA并且对这些转型体进行DNA测序以鉴定密码子在位置30、31和56的身份。基于出现频率(根据NextGen测序)(图8)所选的克隆然后选用于通过瞬时转染进行表达和亲和纯化,如实例2中所述。通过动态光散射(DLS)分析之前,利用超滤来浓缩抗体。所有抗体都能够浓缩到亲本MEDI-1912抗体的8倍与29倍之间(表4),而没有证据证明在这些浓度下沉淀,表明所选抗体的溶解度高于亲本抗体。DLS也表明所选抗体与亲本抗体MEDI-1912相比具有较低平均粒度(Z-Ave)和较小多分散性(PDI)(表4)。四种所选克隆(P5\_C06、P5\_F01、P6\_C08和P6\_F02)相较于此前报道的经改善的克隆MEDI-1912\_STT显示优良或等效的单分散度。平均来看,通过子文库随机建立和哺乳动物展示选择所选的经改善的变异数是原始疏水性残基向亲水性残基的变化。

	ID	aa30	aa31	aa56	C(mg/ml)	Z-Ave (d. nm)	PDI
[0754]	P5_C06	T	S	R	52.1	13.81	0.06
	P5_C11	P	P	N	42.3	24.77	0.135
	P5_F01	T	H	T	48.4	13.48	0.037
	P5_F07	N	T	L	43.9	18.86	0.108
	P5_F12	D	H	L	38.3	15.5	0.113
	P5_G12	H	S	L	31.8	16.44	0.103
	P6_B08	T	P	L	40.8	15.05	0.075
	P6_C08	S	T	A	30.7	12.73	0.054
	P6_C11	S	L	L	15.2	31.97	0.207

[0755]	P6_E07	R	P	L	33.9	19.54	0.177
	P6_F02	R	S	Y	39.1	12.96	0.036
	MEDI-1912_STT	S	T	T	53.1	14	0.048
	MEDI-1912	W	F	L	1.8	23.2	0.148

[0756] 表4.所选MEDI-1912变异数生物物理学特性.通过以30ml规模瞬时转染Expi293细胞(赛默飞世尔)来表达抗体,随后进行亲和纯化(蛋白质A)。在Superdex200 10/300上使用AKTA Pure系统、使用PBS(pH 7.4)操作缓冲液、通过尺寸排阻色谱进一步提纯抗体。通过离心过滤来浓缩抗体并且所得浓度以毫克/毫升(mg/ml)显示(C)。使用Nano SDLS(马尔文仪器公司,英国马尔文)对样品进行动态光散射测量并且使用zetasizer软件(马尔文仪器公司,英国马尔文)计算多分散指数(PDI)和累积量(或z-平均值)大小(Zav)。位置30、31和32以单字母码显示氨基酸一致性。

[0757] 这个实例证明通过哺乳动物展示选择能够将生物物理学特性不良的抗体转化成特性改善(就溶解度和低自身相互作用而言)的抗体。这是通过将所选残基随机诱变并且建立HEK293细胞表面上所展示的大型随机抗体变异数文库来实现。所属领域中的评估抗体可开发性特征(例如溶解度)的技术现状需要按照多毫克规模进行大规模表达和提纯,从而实

现全面的生物物理学和PK测量。在利用多肽呈现水平的差异的此实例中,如根据平均荧光强度(MFI)的差异所判断,我们证明能够建立数百万种变异数体并且选择随后显示具有改善的生物物理学特性的抗体。可以应用这种方法,其中新颖抗体选自初始文库或在另一系统(例如噬菌体展示)中预选的文库。或者,本发明还可以在抗体的亲和力成熟或人源化期间应用,其中建立变异数体文库并且在哺乳动物细胞表面上展示。

[0758] 实例5. 构建变异数体文库以及选择可开发性增强的抗PSK9克隆

[0759] 博可珠单抗是抗前蛋白转化酶枯草杆菌蛋白酶/枯草溶菌素9型(anti-proprotein convertase subtilisin/kexin type 9, PCSK9)mAb,其正由辉瑞(Pfizer)开发以减少血清中的低密度脂蛋白胆固醇(LDL-C)。博可珠单抗的作用机理是抑制PCSK9介导LDL受体(LDLR)降解且借此减少血清LDL-胆固醇(LDL-C)<sup>97</sup>。辉瑞在2016年11月退出研发这种抗体,并宣布“这种抗体不大可能给患者、医师或股东带来价值”。已经报道,博可珠单抗的生物物理学特性不是最优的<sup>2</sup>并且这可能是其临床失败的原因。举例来说,博可珠单抗在多种分析中展示自聚集与交叉相互作用<sup>2</sup>。相比之下,FDA批准的抗PCSK9阿利库单抗(Regeneron)抗体在相同分析中不展示相同水平的自聚集和交叉相互作用。

[0760] 博可珠单抗最初是通过免疫接种PCSK9基因剔除小鼠且根据其抑制PCSK9活性的能力筛选产生单克隆抗体(mAb)的融合瘤克隆而发现的<sup>98</sup>。然后通过将编码来自重链可变域(VH)和轻链可变域(VL)的互补决定区(CDR)的DNA克隆到人类构架<sup>99</sup>而使小鼠mAb 5A10(美国专利:US 8399646 B2)发生人源化,其中VH CDR1和VH CDR2中进行氨基酸取代以得到人源化mAb 5A10-i。如此前所述<sup>100</sup>,对这种人源化抗体5A10-i进一步进行亲和力成熟,以得到博可珠单抗。亲本小鼠mAb 5A10、人源化中间抗体5A10-i与博可珠单抗的VH和VL域的序列比对显示于图9中。5A10、5A10-i和博可珠单抗对PCSK9的亲和力(平衡解离常数或K<sub>D</sub>)分别是1nM、1.5nM和7,如根据表面等离子体共振(SPR)或KinExA(美国专利:US 8399646 B2)所测定。博可珠单抗Fab片段与PCSK9复合的晶体结构已经测定<sup>98</sup>且这已经显示抗体通过轻链和重链结合到PSK9的催化域,其中以VH CDR3的贡献为主。

[0761] 核酸酶介导抗体基因整合到HEK293细胞并且在细胞表面上展示之后,我们已经观察到与衍生博可珠单抗的人源化中间版本5A10-i相比,博可珠单抗的细胞表面呈现减少(图10)。这个实例的目的是证明,通过哺乳动物展示能够从变异数体文库中选择博可珠单抗变异数体,其中良好的呈现水平表明生物物理学稳定性改善、自聚集减少和交叉相互作用特性减少或“粘住”所保持的靶抗原结合。重要的是首先鉴定抗体中可能会造成其不良生物物理学特性的区域或“补丁”。举例来说,已知多肽序列内的连续疏水性氨基酸残基能够引起所述蛋白质产生不良的表达水平<sup>101</sup>。抗体上的疏水性补丁还能够导致不良生物物理学特性<sup>6,7,14</sup>。类似地,抗体表面上来自簇聚的赖氨酸或精氨酸残基的正电荷补丁还能够通过非特异性结合到新生儿Fc受体(FcRn)<sup>22</sup>[6或表达带负电分子(例如肝素硫酸盐)的细胞<sup>23</sup>]。

[0762] 使用本发明产生改善的结合剂的方法始于鉴定序列内的残基作为用于在文库内改变的候选残基。这些位置能够充当使用超过一个替代氨基酸进行随机化的位点或可以是单个氨基酸取代的位点。诱变可以使用所属领域的技术人员已知的方案进行,例如寡核苷酸定向诱变(<sup>102</sup>《分子克隆:实验指南(Molecular Cloning:aLaboratory Manual)》第3版,Russell等人,2001,冷泉港实验室出版社(Cold Spring Harbor Laboratory Press),和其中的参考文献)。在博可珠单抗的这种情况下,可获得三维结构,对其加以分析并且鉴定用

于诱变的候选氨基酸残基。结构建模可以用作有助于鉴定用于诱变的靶氨基酸的替代方案。

[0763] 在本发明内熟练产生和筛选数百万种变异数体意味着能够对序列变异数体进行充分的检索,即使在缺乏任何3D结构信息或模型的情况下,例如通过检视线性序列。这能通过分析线性序列的特征(例如疏水性或电荷簇聚)来完成。作为一个替代方案,能够进行聚焦于个别氨基酸的突变扫描以便在亲和力成熟操作期间引导较大规模的组合型诱变操作<sup>103</sup>。利用相同方案,个别氨基酸可以被替代的氨基酸组取代以鉴定具有改善生物物理学特性的潜力的个别残基。这些残基随后可以形成用于组合型诱变的基础,其中多个位置同时改变。在抗体基因的情况下,与生殖系序列进行比对可以有助于鉴定用于改善表达的最优氨基酸变化。VH上的非互补位氨基酸残基采取这种方案。对疏水性或电荷补丁有贡献的残基是VH CDR1内的Y33、VH CDR2内的F54和R57(图9)。与人类VH生殖系序列的多序列比对显示于图11中。基于这种比对,使对疏水性或电荷补丁有贡献的非互补位博可珠单抗VH氨基酸回复为下文所列的生殖系序列:

[0764] a. VH Y33A(回复为生殖系IGHV1-3\*01)

[0765] b. VH Y33D(回复为生殖系IGHV1-8\*01)

[0766] c. VH S52N,F54S,R57S(三重突变体回复为生殖系IGHV1-46\*01)

[0767] d. VH Y33A,S52N,F54S,R57S(a.和c.:组合的突变体)

[0768] e. VH Y33D,S52N,F54S,R57S(b.和c.:组合的突变体)

[0769] 对于互补位残基来说,建立随机突变体文库以能够根据呈现和保持的抗原结合进行选择(图9)。VL内存在问题的候选残基是:Y53,L94和W95。根据博可珠单抗与PCSK9的共晶体结构,这些残基直接与靶抗原相互作用或间接通过异位相互作用促成结合,例如VL CDR3残基W95似乎抵着VH CDR3残基压紧并且可以维持VH CDR3构象以便最优结合到PCSK9。通过在这些位置中构建随机文库和选择,将能探究是否存在最优的氨基酸组合以便改善生物物理学特性、同时保持抗原结合。替代的文库设计可能涉及使VH CDR1和CDR2内的所选非互补位残基发生随机诱变。

[0770] 设计合成性VH基因块并且加以合成,所述基因块编码上文(a)到(e)所列的以下构建体和原始野生型博可珠单抗(f)。编码这些合成基因的DNA序列显示于图12中。使用引物3054和3055对这些基因块进行PCR扩增(表5),以产生375bp产物。这种产物通过离心柱提纯,用NcoI/XhoI消化并且通过离心柱提纯。然后使6个经消化的VH插入片段与pINT17-VHNcoI/XhoI切割载体接合,接合用于转化大肠杆菌DH5α,挑选个别群落,制备小型制备型质体DNA且证实DNA序列。

[0771]	3052	TTTTTGCCATGGCCCAAGTG	7A2-VH-F
	3053	AAAAAAACTCGAGACGGTGACC	7A2, 107_A07-VH-R
	3054	TTTTTGCCATGGCCCAGG	博可珠单抗-VH-F
	3055	AAAAAAACTCGAGACTGTCACGG	博可珠单抗-VH、7D4-中间体-R
	3056	TTTTTGCTAGCGACATCCAGATG	博可珠单抗、7D4-中间体-VL-F
	3057	TTTTTGCCATGGCCCAGGTTC	博可珠单抗-VL-R
	3052	TTTTTGCCATGGCCCAAGTG	7A2-VH-F
	3053	AAAAAAACTCGAGACGGTGACC	7A2, 107_A07-VH-R
	3069	CTGGGCACGCCGGTGTATCTSNNNGCTGGCGCTGTAGA TCAGCAG	博可珠单抗-VL -R-Y53-随机
	3070	GTGCCCTGGCAAATGTCCGSNNNSNNAGAGTACCGCT GCTGGCAGTAG	博可珠单抗-VL -R-L94W95-随机
	3071	TTTTTGCTAGCGACATCCAGATG	博可珠单抗-VL-F1
	3072	GCTGGCGCTGTAGATCAGCAG	博可珠单抗-VL-R1
	3073	AGATACACCGGCGTGCCCAG	博可珠单抗-VL-F2
	3074	AGAGTACCGCTGCTGGCAGTAG	博可珠单抗-VL-R2
	3075	AAAAAAAGCGGCCGCGGTACGCTTGATTCCAGCTTGG TGCCCTGGCAAATGTCCG	博可珠单抗-VL-R3-延长
	3076	TTTTTGCCATGGCCCAGGTTAG	博可珠单抗-VH-F1
	3077	AAAAAAACTCGAGACTGTCACGGTGG	博可珠单抗-VH-R1

[0772] 表5.引物序列.

[0773] 使用VL基因模板、通过NNS PCR组装诱变而使VL Y53、L94、W95密码子发生随机突变,所述VL基因模板在经受诱变的位置含有终止密码子(VL基因模板序列参见图12)。进行以下PCR:

[0774] a) 使用引物3071/3047(表5)对博可珠单抗VL加终止子基因块(序列参见图12)进行PCR扩增,得到353bp产物。

[0775] b) 使用引物3071/3069(表5),使用上述模板(a)进行PCR,得到191bp产物。

[0776] c) 使用引物3073/3070,使用上述模板(a)进行PCR,得到146bp产物。

[0777] d) 使用上述PCR反应产物b和c,使用外部引物3071/3075进行PCR组装反应,得到353bp插入片段。产物用NheI/NotI消化且通过离心柱提纯。

[0778] 上述6种VH变异体a到f(参见图12)与编码恒定κ轻链(CL-κ)的“填充”片段、聚腺苷酸、CMV启动子和信号序列和VL NNS文库通过PCR组装在一起。使用引物κ填充子F4(GTACCGCGGCCGCACCTTCCG)和λ填充子R3(CAGCCATGGCGCCTGTGGAGAGAAAGG),从pINT3质体扩增填充片段(WO2015166272A2)。组装的插入片段用NheI和XhoI消化,通过离心柱提纯并且与pINT17-BSD靶向载体(100ng)接合(每个接合反应50ng插入片段),用NheI和XhoI预消化。使用mini-Elute PCR提纯试剂盒(Qiagen)提纯接合混合物(20μl)并且使提纯的接合混合物(4μl)在E.cloni 10G elite电感受态细胞(50μl,600512,Lucigen)中转化。使用0.1cm光析管对细胞进行脉冲,使用2m1回收培养基回收并且在37℃、250rpm下生长1小时。为了计算文库大小,将细胞1:1000稀释并且接种(10μl和100μl)于10cm直径的2TY-卡那霉素培养板

中。将剩余细胞离心并且接种于 $2 \times 10\text{cm}$ 直径的2TY-卡那霉素培养板中并且在37°C下培育整夜。对10 $\mu\text{l}$ 培养板中的群落进行并且文库大小经计算是 $2 \times 10^6$ 个转型体。为了代表每种变异数，所需文库大小是 $1.2 \times 10^5$ 个克隆。 $(32^3 = 3.4 \times 10^4 / \text{文库} \times 6\text{VH突变体})$ ，因此所产生的文库代表所需文库多样性的16倍过度表达。刮擦转化体培养板，测量OD600，使用2OD600接种50ml circlegrow培养基，培养基在250ml有挡板的烧瓶中在37°C生长3到4小时，收获约4000D600单位并且制备6种midiprep质体DNA，其代表6种VH博可珠单抗变异数与三个NNS密码子VL文库的组合(图12)。通过读取260nm吸光度定量6种midiprep质体DNA并且以等摩尔比混合，得到博可珠单抗靶向载体文库pINT17-BSD-Boco1-文库。

[0779] 转染前两天，经悬浮液调适的HEK293细胞以 $2.5 \times 10^5$ 个细胞/毫升接种于HEK FreeStyle 293表达培养基中。转染当天，将细胞离心并且以 $10^8$ 个细胞/毫升的最终体积再悬浮于制造商的电穿孔缓冲液(1ml, Maxcyte电穿孔缓冲液，赛默飞世尔科技，目录NC0856428)中，所述电穿孔缓冲液含有pINT17-BSD-博可珠单抗-文库(20 $\mu\text{g}$ )以及编码AAVS左和右TALEN核酸酶(TALEN，各100 $\mu\text{g}$ )的质体。将HEK293/质体DNA混合物(0.4ml)转移到单个OC-400光析管(MaxCyte，目录号OC-400R10)并且使用MaxCyte STXG2在HEK293配置下进行脉冲。对照物(减TALEN和pINT17-BSD-博可珠单抗和pINT17-BSD-5A10-i)使用OC-100光析管(MaxCyte，目录号OC-100R10)在相同配置下转染。将电穿孔的细胞转移到锥形瓶(250ml)之后，允许细胞在添加FreeStyle 293表达培养基(40ml, LifeTech, 目录号12338018)之前搁置30分钟。使细胞充分再悬浮并且置于设定为130RPM、37°C和5%CO<sub>2</sub>的回转式振荡培育箱中。

[0780] 在24小时后， $1 \times 10^6$ 个细胞用抗人类Fc PE(剑桥生物科学(Cambridge Bioscience)，目录号409304)染色以证实瞬时表达。简单来说，细胞以600xg离心2.5分钟。丢弃上清液并且将细胞再悬浮于0.1%BSA(从7.5%溶液稀释：LifeTech，目录号15260037)中。再次将这些细胞离心并且再悬浮于100 $\mu\text{l}$ 的1%BSA/PBS(其中添加有1 $\mu\text{l}$ 抗人类Fc PE)中。这些细胞在4°C下在暗处培育30分钟。细胞用1ml的0.1%BSA/PBS洗涤两次并且再悬浮于0.5ml的0.1%BSA/PBS中，所述PBS含有5 $\mu\text{l}$ 的7-AAD(eBioscience，目录号00-6993-50)。去除50 $\mu\text{l}$ 并且添加到96孔板的各孔中。使用IntelliCyt流式细胞仪分析呈现水平并且其显示用于转染的瞬时细胞表面抗体。当观察到瞬时表达时，培养物前行至进行选择，所述选择使用抗生素杀稻瘟菌素S HC1(LifeTech，目录号R21001)，浓度为7.5 $\mu\text{g}/\text{ml}$ 。细胞以 $0.25 \times 10^6$ 个细胞/毫升接种于锥形烧瓶(Erlenmeyer flasks)中。细胞还接种于10cm培养皿(康宁(Corning)，目录号353003)中的DMEM(LifeTech，目录号41965039)中，所述DMEM含有10%FBS(西格玛奥德里奇公司(Sigma Aldrich)，目录号F9665-500 ML)和1%青霉素/链霉素(西格玛奥德里奇公司，目录号P0781-100ML)，10ml中含有10,000个细胞或1000个细胞。允许这些细胞在添加7.5 $\mu\text{g}/\text{ml}$ 的杀稻瘟菌素S HC1(LifeTech，目录号R21001)之前附接24小时。转染12天之后，培养板用2%亚甲基蓝染色。通过对根据总细胞的指定输入所达成的杀稻瘟菌素群落数目计数来计算转染效率百分比。整合效率经计算是2%并且4千万细胞转染所达成的文库大小是800,000，为所需理论文库大小(120,000)的7倍大。因此，构建编码变异数的所有可能组合的哺乳动物展示文库。

[0781] 在杀稻瘟菌素S HC1(LifeTech，目录号R21001)选择5天之后，使用MACS珠粒和柱(美天旎(Miltenyi)，目录号130-048-801及目录号130-042-401)号富集细胞。文库已经扩

增到逾2亿个细胞。1亿个细胞以200xg离心并且用0.1% BSA-PBS洗涤。将这些细胞再悬浮于9.9ml的1% BSA-PBS中并且添加100μl的抗人类Fc PE抗体(剑桥生物科学,目录号409304)。剩余1亿个细胞也进行离心,洗涤并且与生物素化PCSK9抗原(10nM, PC9-H82E7, AcroBiosystems, 10ml, 在1% BSA-PBS中稀释)一起培育。两者均在4°C下、在暗处培育30分钟。从这一点看,使用自动MACS漂洗溶液(美天旎,目录号130-091-221)。细胞用自动MACS漂洗溶液(10ml, 1×PBS+2mM EDTA+0.5% BSA)洗涤,以200xg离心并且再悬浮于800μl自动MACS漂洗溶液中。添加200μl的抗PE(美天旎,目录号130-048-801)微珠或抗生蛋白链菌素(美天旎,目录号130-048-101)微珠。将这些物质在4°C下、在暗处培育10分钟,随后用10ml的自动MACS漂洗溶液洗涤并且再悬浮于5ml备用液中以施加到柱中。MACS LS柱(美天旎,目录号130-042-401)用3ml自动MACS漂洗溶液预洗涤,随后添加细胞。每组细胞使用4x柱。将1/4(约1.25ml)的细胞添加到每个柱中后,用3ml缓冲液洗柱3次。从磁力夹持器中移出LS柱并且添加5ml缓冲液。使用活塞将其通过柱推入15ml法尔康管(Falcon tube)中以洗脱所结合的细胞。为了进一步提纯群体,将此5ml添加到新制的柱(如前所述预洗涤)中并且如前所述处理。对细胞计数并且发现在5ml中是约 $1.5 \times 10^6$ 个细胞/毫升。将这些细胞离心并且再悬浮于30ml的含有7.5μg/ml杀稻瘟菌素的FreeStyle培养基FreeStyle 293表达培养基(LifeTech,目录号12338018),并且在37°C、5%CO<sub>2</sub>下培育直至备好传代。

[0782] MACS后的第48小时,针对Fc呈现和抗原结合将细胞染色(图13)。因此,基于针对抗原或Fc表达的第一轮选择来建立2个群体(选择编号分别为884和885)。随后利用流式细胞术,针对抗原与Fc表达的组合来选择这些细胞。程序与此前所述的24小时染色相同,但有如下调整:将细胞与10nM生物素化人类PCSK9、Avi-i-标签(目录号PC9-H82E7-25ug, ACRObiosystems)一起在4°C下培育30分钟,随后洗涤并且与抗人类Fc PE(1μl/1x10<sup>6</sup>个细胞)(剑桥生物科学,目录号409304)和抗strep APC(英杰公司(Invitrogen),目录号SA1005)(0.5μl/1x10<sup>6</sup>个细胞)的混合物一起培育。7-AAD(eBioscience,目录号00-6993-50)用于如前所述评估存活率。洗涤细胞并且使用IntelliCyt仪器进行分析,如此前所述(图10)。选择14天之后,使用BD Influx进行FACS。将针对抗原结合或Fc呈现分选出的MACS分选群体中的 $20 \times 10^6$ 个细胞与10nM生物素化人类PCSK9(PC9-H82E7-25ug, ACRObiosystems)一起培育(如此前),随后洗涤并且与抗人类Fc PE(1μl/1x10<sup>6</sup>个细胞)(剑桥生物科学,目录号409304)和抗strep APC(0.5μl/1x10<sup>6</sup>, SA1005, 英杰公司)的混合物一起培育。即将分选之前添加DAPI(1μl/百万个细胞)。将细胞分选到两个其它群体:较高抗原结合群体(门P5)和较低抗原结合群体(门P6),如图14所示。使FACS提纯的这些群体在缺乏杀稻瘟菌素、但存在1%青霉素/链霉素的情况下生长以免被细胞分选过程污染。培养4天之后,每个群体取1x10<sup>6</sup>个细胞用于提取基因组DNA。

[0783] 使用KOD热启动DNA聚合酶(默克密理博公司),通过巢式PCR扩增编码IgG的DNA,如实例4中所述。对PCR产物进行凝胶提纯并且用NheI和XhoI消化,克隆到哺乳动物pINT3表达载体中并且用于转化大肠杆菌DH10B细胞。

[0784] 对未选定的随机输入克隆(84)、所分选的抗原(75)和所选Fc(85)进行测序并且测定VH一致性。从哺乳动物展示选择循环所得的经测序的克隆组中都不存在原始博可珠单抗VH基因,而是强烈偏向由IGHV1-46\*01生殖系构成的变异数体(参见图15)。测定同组克隆的VL序列。计算3种突变密码子的平均pI和脂肪族指数。其显示根据哺乳动物展示选择的抗体的

pI与脂肪族指数均减小(图16),表明大大偏离CDR2和3中的原始疏水性氨基酸。

[0785] 为了表明根据哺乳动物展示选择的博可珠单抗变异体的生物物理学特性优于亲本抗体,接下来对所选抗体进行表达。根据选择将克隆挑选到96孔板中(91/群体):针对PCSK9的MACS(命名为选择884)或针对抗Fc的MACS(命名为选择885)。这些群落用于两板DNA,以便使用Qiagen Plasmid Plus 96Miniprep试剂盒(Qiagen,目录号16181)、依循制造商说明书进行转染。使用这种DNA,使用Expi293转染系统(LifeTech,目录号A14525)、依循制造商说明书转染两个96孔板的Expi293细胞。5天之后,收获这些细胞并且使上清液保持在4°C。为了测定抗体聚集倾向,使用一种称为AC-SINS(亲和力捕捉自身相互作用纳米颗粒光谱法)的方法。所用方法基本上如Liu等人,2014<sup>39</sup>所述,但有以下修改。用PEG阻断金纳米颗粒(AuNP,柠檬酸盐稳定化的20nm金纳米颗粒,15705,Ted Pella有限公司)之后,将AuNP在4°C下储存直到需要时为止(最长一周)。AuNP在4°C下以15,000RPM离心10分钟,去除95%上清液且进一步在相同条件下离心,而非使用针筒过滤器浓缩到10倍。将最终AuNP再悬浮于1/10的起始体积中。聚丙烯96孔板的每个孔(含有100μl存在于上清液中或在PBS中提纯的测试抗体)中添加10μl。培养板在室温下、在设定为700RPM的振荡平台上培育2小时。如Liu等人(2014)<sup>39</sup>所述,将内容物小心地转移到聚苯乙烯UV透明板中。使用BMG Pherastar仪器,以2nm增量收集450到650nm的吸光度数据。鉴定最大吸光度的波长并且任一侧的10个点与紧邻之前及之后的点一起平均化以减少杂讯误差。这些平均值的最高点取作最大吸光度。图17列出了AC-SINS分析的结果以及抗体CDR序列。所选变异型克隆中的大部分(86/91)显示12nm或更小的AC-SINS波长漂移,相当于人源化中间克隆5A10-i。这种波长表明在这些样品中,自缔合不会以任何大的程度发生。相比之下,博可珠单抗使AC-SINS产生26nm的波长漂移(图17),并且所选组中仅5种克隆产生大于20nm的波长漂移。因此,已经根据哺乳动物细胞展示选出自聚集倾向低于原始亲本克隆的博可珠单抗变异型克隆,如通过AC-SINS分析所判断。

[0786] 来自表达过的培养板的上清液也用于比较抗体保持结合到PCSK9的能力。这是在捕捉ELISA分析中使用单体抗原进行,已显示这是对抗体结合到其靶标进行亲和力排序的有效方式。简单来说,96孔Maxisorp培养板(Nunc,目录号437111)在4°C下用含有3μg/ml抗人类Fc抗体(Jackson ImmunoResearch,目录号209-005-098)的PBS涂布整夜。次日,培养板用1×PBS洗涤3次并且随后在室温下用300μl含有3% (w/v) 奶粉(Marvel)的1×PBS(M-PBS)阻断1小时。这些培养板用1×PBS洗涤3次,并且向每个孔中添加30μl的6% (w/v) 奶粉(Marvel)。添加30μl的各上清液并且在室温下培育1小时。培养板然后用1×PBS-Tween(0.1%)洗涤3次并且然后用1×PBS洗涤3次。向各孔中添加每孔60μl的0.1nM生物素化人类PCSK9、Avi-标签(ACROBiosystems,目录号PC9-H82E7)并且将培养板在室温下培育1小时。培养板如前所述用1×PBS-Tween洗涤,随后用1×PBS洗涤。每孔添加60μl含有抗生蛋白链菌素-铕(珀金埃尔默(Perkin Elmer),目录号1244-360)的DELFIA分析缓冲液(珀金埃尔默,目录号1244-111)(1:500稀释度)并且在室温下培育1小时。培养板最后一次用1×PBS-Tween和1×PBS洗涤,随后每孔添加50μl的DELFIA增强溶液(珀金埃尔默,目录号4001-0010)。将培养板在板振荡器上、在300RPM下放置5分钟并且在BMG Labtech PHERAStar读盘器(激发340nm,发射615nm)上读取。这表明大部分抗体对PCSK9保持结合,其中若干抗体显示的捕捉ELISA信号等效于博可珠单抗( $K_d=7\text{pM}$ )和5A10-i( $K_d=1.5\text{nM}$ )中间克隆(图17)。这

表明通过建立文库和哺乳动物展示选择,能够靶向抗体互补位残基并且选择同时具有改善的生物物理学特性和保持结合到靶抗原的能力的抗体。

[0787] 然后基于低AC-SINS波长漂移和抗原结合保持的AC-SINS培养基上清液分数来选择克隆(图17)。这些克隆连同博可珠单抗、5A10-i和阿利库单抗(批准的抗PCSK9抗体)一起然后通过瞬时转染Expi-293细胞(50ml规模)表达并且通过蛋白质A亲和色谱来提纯,随后透析,如实例3中所述。然后通过HPLC-SEC分析抗体并且这表明所选的全部抗体显示的HPLC滞留时间和峰宽与对照阿利库单抗抗体和5A10-i相等(图18)。相比之下,博可珠单抗在柱上迟滞并且显示较长的滞留时间。博可珠单抗也显示了不对称的峰,也表明其具有交叉相互作用特性并且不会特异性结合到柱基质。经提纯的抗体也通过AC-SINS分析并且其显示波长漂移与阿利库单抗和5A10-i相等( $\Delta\lambda=8$ 到 $12\text{nm}$ ),而博可珠单抗显示较长的AC-SINS波长漂移,表明其具有自身相互作用特性( $\Delta\lambda=39\text{nm}$ )。表达产量、AC-SINS波长漂移、HPLC-SEC滞留时间和HPLC-SEC峰宽概括于表6中。

[0788] 因此,这个实例已举例说明能够利用高等真核细胞上的结合剂展示来选择可开发性特征改善(包括减少的自身相互作用和减少的非特异性相互作用、同时保持对靶标的结合)的变异数体。在这个实例中,这是如下实现的:首先鉴定抗体表面上的疏水性和正电荷补丁,进行随机或靶向诱变以建立变异数体文库并且利用核酸酶介导的结合剂基因打靶来实现每个细胞的单个基因拷贝。然后基于细胞展示水平和抗原结合来分选细胞展示文库,以鉴定生物物理学特性改善的亲本抗体的变异数体。

测试克隆或对照	AC-SINS $\Delta\lambda$ (nM)	HPLC-SEC 滞留时间(分钟)	峰宽(最小值)	在Expi293中的表达(mg/L)
[0789]	884_01_G01	9	6.84	0.45
	884_01_A01	10	6.87	0.42
	884_01_A04	9	6.86	0.41
	884_01_F02	12	6.91	0.42
	884_01_E12	9	6.97	0.65
	博可珠单抗	39	7.35	1.14
	5A10-i	10	6.91	0.56
	阿利库单抗	8	6.87	0.26

[0790] 表6.比较博可珠单抗和改善型变异数体的生物物理学特性(包括AC-SINS分析中的波长漂移(nm)、HPLC-SEC滞留时间、峰宽和表达产量)。

[0791] 实例6a.通过选择增强非交叉相互作用克隆的可开发性

[0792] 具有非特异性结合到除其靶标之外的分子的特性并且具有体内不良半衰期的抗体能够引起“脱靶”结合,导致药物动力学(PK)和药效学(PD)不良。另外,在抗体制造期间,交叉相互作用或“粘着”特性会带来问题,主要导致在提纯或配制期间被柱基质迟滞的问题。

[0793] 这个实例证明能够利用哺乳动物抗体展示将存在已知“粘着”或交叉相互作用问题的抗体与表现良好且已批准用于临床用途的抗体区分开来。选择抗神经纤毛蛋白-1抗体韦森单抗(或MNRP1685A)作为“粘性”抗体的一个实例。已知这种抗体在尺寸排阻色谱期间被迟滞并且非特异性地结合到柱基质。这被认为造成其在动物模型中的半衰期不良<sup>104</sup>。另外,在观察到蛋白尿症副作用之后,停止了这种抗体的临床开发<sup>105</sup>。选择抗PD1抗体尼沃单

抗作为表现良好的抗体实例,其已批准用于临床用途<sup>106</sup>。韦森单抗在亲和力捕捉自身相互作用纳米颗粒光谱(AC-SINS)分析中也显示一些自身相互作用<sup>39</sup>。

	链	蛋白质序列
[0794]	韦森单抗 (重链)	<u>EVQLVESGGG</u> LVQPGGLRL SCAASGFTFS SYAMSWVRQA PGKGLEWVSQ <u>ISPAGGYTNY</u> ADSVKGRFTI SADTSKNTAY LQMNSLRAED TAVYYCARGE <u>LPYYRMSKVM</u> DVWGQQGTLVT
	韦森单抗 (轻链)	<u>VSSASTKGPS</u> VFPLAPSSKS TSGGTAALGC LVKDYFPEPV TVSWNSGALT SGVHTFPAVL QSSGLYSLSS VVTVPSSSLG TQTYICNVNH KPSNTKVDKK VEPKSCDKTH TCPPCPAPEL LGGPSVFLFP PKPKDTLMIS RTPEVTCVVV DVSHEDPEVK FNWYVDGVEV HNAKTKPREE QYNSTYRVVS VLTVLHQDWL NGKEYKCKVS NKALPAPIEK TISKAKGQPR EPQVYTLPPS REEMTKNQVS LTCLVKGFYP SDIAVEWESN GQPENNYKTT PPVLDSDGSF FLYSKLTVDK SRWQQGNVFS CSVMHEALHN HYTQKSLSLS PGK
[0795]	尼沃单抗 (重链)	<u>DIQMTQSPSS</u> LSASVGDRVT ITCRASQYFS SYLAWYQQKP GKAPKLLIYG <u>ASSRASGVPS</u> RFSGSGSGTD FTLTISSLQP EDFATYYCQQ YLGSPPTFGQ <u>GTKVEIKRTV</u> AAPSVFIFPP SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ ESVTEQDSKD STYSLSSTLT LSKADYEKHK VYACEVTHQG LSSPVTKSFN RGEc
	尼沃单抗 (重链)	<u>QVQLVESGGG</u> VVQPGRLRL DCKASGITFS NSGMHWVRQA PGKGLEWAV <u>IWYDGSKRYY</u> ADSVKGRFTI SRDNSKNTLF LQMNSLRAED TAVYYCATND <u>DYWGQQGTLVT</u> <u>VSSASTKGPS</u> VFPLAPCSRS TSESTAALGC LVKDYFPEPV TVSWNSGALT SGVHTFPAVL QSSGLYSLSS VVTVPSSSLG TKTYTCNVDH KPSNTKVDKR VESKYGPPCP PCPAPEFLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS QEDPEVQFNW YVDGVEVHNA KTKPREEQFN STYRVSVSLT VLHQDWLNGK EYKCKVSNKG LPSSIEKTISK AKGQPREPQV YTLPPSQEEM TKNQVSLTC LVKGFPYPSDI AVEWESNGQP ENNYKTPPPV LDSDGSFFLY SRLTVDKSRW QEGNVFSCSV MHEALHNHYT QKSLSLSLGK
	尼沃单抗 (轻链)	<u>EIVLTQSPAT</u> LSLSPGERAT LSCRASQSVS SYLAWYQQKP GQAPRLLIYD <u>ASNRATGIPA</u> RFSGSGSGTD FTLTISSLEP EDFAVYYCQQ SSNWPRTFGQ <u>GTKVEIKRTV</u> AAPSVFIFPP SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ ESVTEQDSKD STYSLSSTLT LSKADYEKHK VYACEVTHQG LSSPVTKSFNRGEC

[0796] 表7. 韦森单抗和尼沃单抗重链和轻链的蛋白质序列以单字母码显示了完整抗体重链和轻链的氨基酸序列。可变域加有下划线。

[0797] 将编码韦森单抗和尼沃单抗重链和轻链可变域的合成DNA(其序列参见表7)克隆到基于pINT3的双重启动子IgG可溶性表达载体(WO2015166272A2)中并且证实DNA序列。为了检查可溶性抗体的特性,由pINT3-韦森单抗和pINT3-尼沃单抗制备质体DNA并且使用其转染Expi293细胞(30ml最终培养体积规模),此转染使用转染试剂ExpiFectamine、根据制造商说明书(A14525,赛默飞世尔科学)进行。转染之前的第24小时,以 $2 \times 10^6$ 个细胞/毫升的密度接种存在于25.5ml Expi293表达培养基中的细胞。使质体DNA(30μg)在Opti-MEM培

养基(1.5ml)中稀释并且使ExpiFectamine 293试剂(80μl)在Opti-MEM培养基(1.5ml)中稀释并且在室温下培育5分钟。然后将稀释的质体DNA(30μg, 存在于1.5ml Opti-MEM培养基中)添加到稀释的ExpiFectamine 293试剂(80μl ExpiFectamine存在于1.5ml Opti-MEM培养基中)中并且在室温下培育20分钟。将细胞在37℃、5%CO<sub>2</sub>、5%湿度下培育且以200rpm搅拌(50mm回转摆幅, ISF1-X, Climo-Shaker, Kuhner)。表达5天之后,通过离心(2000g, 20分钟)收获培养基上清液并且如上文所述通过蛋白质A亲和色谱来提纯(实例3)。通过测量280nm吸光度和使用1.4的估测消光系数估算浓度、利用比尔-兰伯特定律计算来测定抗体产量和浓度。韦森单抗和尼沃单抗的表达产量相似(分别是95mg/L和103mg/L)。此外,增加抗体浓度时,以瞬时转染效率实现的表达产量不能解决即将发生的问题。

[0798] 尼沃单抗和韦森单抗的生物物理学特性是通过若干技术测定。如上文(实例3)所述使用Prometheus NT.4B(Nanotemper)确定解链温度( $T_m$ )和聚集起始温度( $T_{agg}$ )。两种抗体的解链温度( $T_m$ )和聚集温度( $T_{agg}$ )相似(参见表8a)。这再次证明解链温度不能预测即将发生的问题。

	抗体	$T_m$ ( °C )	$T_{agg}$ ( °C )	表达产量 (mg/L)	$V_e$ (ml)	PDI	$Z_{av}$ (nm)	$MW_{est}$ (kDa)	多分散度(%)
[0799]	韦森单抗	69.9	70.4	94.8	13.7	0.163	7.7	401	66
	尼沃单抗	67.6	67.3	103	12.0	0.053	5.1	156	0

[0800] 表8a. 韦森单抗和尼沃单抗IgG生物物理学特性。使用Prometheus NT.4B(Nanotemper), 根据制造商说明书测定解链温度( $T_m$ )和聚集起始温度( $T_{agg}$ )。通过瞬时转染30ml规模的Expi293细胞(赛默飞世尔)来测定表达产量(按照每升培养体积所表达的抗体数量(mg)), 随后进行亲和纯化(蛋白质A)并且利用280nm吸光度和1.4的估测抗体消光系数测定经提纯的抗体的产量。在Superdex 20010/300上使用AKTA Pure系统、使用PBS(pH 7.4)操作缓冲液、通过尺寸排阻色谱进一步提纯抗体。使用Nano SDLS(马尔文仪器公司, 英国马尔文)对样品进行动态光散射测量并且使用zetasizer软件(马尔文仪器公司, 英国马尔文)计算多分散指数(PDI)和累积量(或z-平均值)大小( $Z_{av}$ )。

[0801] 然而, 在制备型尺寸排阻色谱期间, 观察到显著的柱基质结合和迟滞(表8a和图19), 如此前所述。尼沃单抗和韦森单抗的洗脱体积( $V_e$ )分别是12.0ml和13.7, 表明韦森单抗被迟滞并且与柱基质发生非特异性相互作用。10.4ml的另一洗脱峰表明存在一些较高分子量的聚集抗体, 而尼沃单抗未观察到。为了研究抗体在储存期间的稳定性, 根据尺寸提纯的抗体在PBS pH7.4中、在4℃下培育2周。动态光散射(DLS)检测到韦森单抗出现较高阶聚集物质, 而尼沃单抗则未检测到(方法参见图20和图例)使用Zetasizer APS(马尔文仪器公司, 英国马尔文)对在4℃已储存2周的样品进行动态光散射测量。DLS得到的生物物理学参数-计算的多分散性百分比、多分散指数(PDI)、累积量(或z-平均值)大小和平均分子量显示于表8中并且显示与在4℃下储存2周后呈单分散性的尼沃单抗相比, 韦森单抗在相同条件下显著聚集。

[0802] 将编码韦森单抗和尼沃单抗重链和轻链可变域的合成DNA克隆到哺乳动物展示载体pINT17-BSD(载体图谱和序列参见实例1)中, 证实DNA序列并且制备转染质量的质体DNA。

HEK293细胞用TALE核酸酶转染并且如上文在实例2中所述建立稳定细胞系。转染后(dpt)14天之后,细胞在4°C下用抗人类Fc PE(409303,Biolegend)染色30分钟以测定抗体显示的展示水平(参见实例2)。展示尼沃单抗或韦森单抗的单克隆细胞系然后根据以下方案、用经标记的人类血清染色。

[0803] 将热灭活的人类AB血清(5μl,40mg/ml H4522,西格玛(Sigma))在PBS(195μl)中稀释以得到1mg/ml的最终浓度。这种稀释的人类血清然后使用Lightning-Link® Rapid Dylight® 633试剂盒(325-0000,Innova)、根据制造商说明书、用Dylight 633标记。使展示尼沃单抗或韦森单抗的HEK293细胞系或野生型HEK293细胞(一百万个细胞)集结(200g,3分钟,在微量离心管(1.5ml)中)。使集结粒再悬浮于PBS(1ml)中且集结(600g,2.5分钟)。使集结粒再悬浮于含有抗Fc PE(0.5μl,409303,Biolegend)的1%BSA/PBS(100μl)中并且用AB血清Dylight 633标记(5μl,0.5mg/ml)。混合物在4°C下避光培育30分钟。添加0.1%BSA/PBS(900μl)并且使细胞集结(600g,2.5分钟)。将细胞再悬浮于0.1%BSA/PBS(1ml)中,且重复此洗涤步骤一次。将细胞再悬浮于含有7-AAD(5μl/百万个细胞)的0.1%BSA/PBS(200μl)中。使用Intellicyte iQue筛选仪分析经标记的细胞(50μl)。流式细胞术分析(图21)显示,与分别就抗体和人类血清结合而言呈双重阳性的尼沃单抗(分别为12.3%和3.7%)相比,经标记的人类血清对HEK293细胞的结合增加。

[0804] 韦森单抗是一种开发失败而未超越1期临床试验的抗体<sup>105</sup>并且是具有已知自聚集和交叉相互作用特性的抗体<sup>39,104</sup>。我们还已经表明,这种抗体显示与尺寸排阻柱基质存在非特异性相互作用(图19)并且在储存后聚集(图20和表8),相比之下,临幊上批准的抗PD1抗体尼沃单抗在尺寸排阻色谱期间未显示交叉相互作用并且在储存之后保持单分散性。我们在此表明,当韦森单抗和尼沃单抗在交叉相互作用流式细胞术分析中展示于HEK293细胞表面上时,我们能够将他们区分开来。可能的是,能够优化这种分析以允许在“粘性”抗体与表现良好的抗体之间作出甚至更大的区分。举例来说,人类血清可以生物素化并且与荧光团标记的抗生蛋白链菌素偶联,所述抗生蛋白链菌素与增强的亲合力偶联。或者,可以修改如此前所述的流式细胞术的任何数目个交叉相互作用分析<sup>2</sup>并且包括(但非排他地)经标记的杆状病毒<sup>16,19</sup>,或蛋白质、DNA和含肝素硫酸盐分子的经标记的混合物<sup>107</sup>。

#### [0805] 实例6b.多反应性筛选的改善

[0806] 使用具有已知多反应性特征的已知抗体,我们进一步举例说明在展示结合剂的细胞克隆群体内,基于对非靶分子的结合差异(多反应性探针),将多反应性结合剂与非多反应性结合剂区分开来的可能性。我们证明未能结合多反应性探针的克隆富集。

[0807] 使用核酸酶定向整合来制备表达优特金单抗、布瑞金单抗和阿玛西单抗(amatuximab)的HEK293细胞的个别群体。所述个别群体用生物素化DNA染色。检测到表达布瑞金单抗和阿玛西单抗的细胞结合到DNA,而表达优特金单抗的细胞未被染色。使DNA结合依模式归一化。图33A.抗Fc抗体染色揭露,布瑞金单抗群体是IgG表达细胞与非表达细胞的混合物,说明所述群体内存在DNA结合细胞与DNA非结合细胞的近似50:50混合物。不同细胞群用细胞追踪染料标记并且等比例混合。阿玛西单抗细胞用Ce11Trace Far红标记(在图33B和C的Q3中、在x轴上显示),布瑞金单抗用Ce11Trace CFSE标记(在图33B和C的Q1中、在y轴上显示),而优特金单抗仍不标记(图33B和C的Q4中,双重阴性群体)。混合群体用生物素化DNA(20μg DNA/1百万细胞,存在于200μl 1%BSA中)染色并且用抗生物素微珠标记。使用

MiniMACS珠粒与MS柱的组合对所述群体进行分选。MACS分选得到的流过洗脱份使用Intellicyt流式细胞仪分析以对细胞计数。7-AAD用作活力染色剂并且将死细胞排除在分析之外。与布瑞金单抗和阿玛西单抗相比,观察到不结合DNA的优特金单抗富集(图33B、图33C、表8b)。

[0808]	样品	布瑞金单抗	混合物	阿玛西单抗	优特金单抗
[0809]	MACS 前	3120	114	3420	3350
	MACS 后	2540	0	697	6790

[0810] 表8b.MACS之前及之后,针对各种抗体的相对细胞计数(相对于优特金单抗细胞数目归一化并且相对于10,000个计数归一化)。

[0811] 与展示优特金单抗的细胞相比,展示布瑞金单抗的细胞的相对百分比减少到37%并且展示阿玛西单抗的细胞减少到10%。可能的是,由于布瑞金单抗的输入群体包括比例相对较高的非抗体表达细胞并且这些细胞保留于在此选择的“未结合群体”中,因此富集系数甚至可以更高。这种背景可以通过预分选、后分选或协同分选用于IgG表达或抗原结合的细胞来进一步减少,如早先所描述。

[0812] 在此我们成功地使用MACS,而且预期拆分和有效富集甚至大于FACS流式分选。

[0813] 还测试了其它多反应性探针将表达多反应性抗体的细胞克隆与表达非多反应性抗体的细胞克隆区分开来的能力。

[0814] 我们发现基于肝素硫酸盐结合程度可以将表达优特金单抗的细胞与表达布瑞金单抗或加尼图单抗(ganitumab)的细胞区分并分离。简单来说,使用如此前在实例5中所述的标准染色方案,用9μM肝素-FITC(Creative PEGWorks)将250,000个细胞染色。布瑞金单抗和加尼图单抗显示肝素结合。叠加图显示于图34中。这种非特异性结合可能通过布瑞金单抗和加尼图单抗的重链CDR中的带正电补丁发生。

[0815] 伴侣蛋白代表着其它多反应性探针,其可以作为非靶分子用于取消对多特异性结合剂的选择。伴侣蛋白在功能上与蛋白质折叠相关并且有助于蛋白质折叠。热休克蛋白(Hsp)在应激条件(例如高温)下过度表达。大部分伴侣蛋白在正常细胞中也受到丰富的表达,其中他们识别并且结合非原生蛋白质,从而防止聚集。

[0816] 多种治疗抗体展示于HEK293细胞上并且测试对Hsp70和Hsp90的结合。

[0817] 在我们的实验测试的抗体中,贝伦妥单抗(brentuximab)和冷自鲁单抗(lenzilumab)显示结合到Hsp70和Hsp90。贝伦妥单抗(维多汀(Vedotin))是一种抗CD30抗体-药物偶联物,其在治疗霍奇金氏淋巴瘤(Hodgkin's lymphoma)的临床试验中失败。此前显示其展现自身相互作用和交叉相互作用<sup>2</sup>。冷自鲁单抗是一种抗GM-CSF(粒细胞-巨噬细胞群落刺激因子)抗体,其在重度哮喘的II期试验中失败。

[0818] 图35显示经抗人类Fc PE和与DyLight 633偶联的热休克蛋白(Hsp70和90)双重染色的优特金单抗和布瑞金单抗的叠加图。叠加图内的门指示未显示与伴侣蛋白(Hsp70和Hsp90)发生可检测相互作用并且能够利用FACS分选以提供所选克隆群体的细胞,在所选克隆群体中,识别伴侣蛋白的结合剂已经耗竭(并且优选排除)。

[0819] 图36中描绘了使用多种不同抗体和不同多反应性探针的多项实验的混合数据。个别安瓿是在分开的独立实验中测试,但Intellicyt流式细胞仪具有固定的电压,因此预期

所有样品的荧光强度一致。另外,Hek293用作各实验的内部对照。

[0820] 下表8c概括了针对多反应性探针的结合测试的一组抗体的数据。

[0821]

抗体	DNA	肝素	伴侣蛋白	FcRn
阿利库单抗	否	否	否	否
阿玛西单抗	是	是	是	是
贝伦妥单抗	ND	ND	是	否
布瑞金单抗	是	是	是	是
加尼图单抗	是	是	否	是
冷组力单抗(Lenzulimab)	ND	是	是	否
优特金单抗	否	否	否	否
韦森单抗	ND	ND	否	是

[0822] 表8c.抗体多反应性筛选的概述.是=抗体显示与多反应性探针结合。否=抗体未显示与多反应性探针结合。ND=对多反应性探针的结合未测定。

[0823] 阿玛西单抗、贝伦妥单抗、布瑞金单抗和冷自鲁单抗显示结合到伴随蛋白。这些抗体的多反应性可能起因于抗体可变域内的疏水性氨基酸簇,从而与也具有疏水区的蛋白质(例如伴随蛋白)产生范德华力相互作用(van der Waals interactions)。多反应性的其它原因可能是存在带正电荷的氨基酸补丁(例如由精氨酸或赖氨酸残基组成),其与具有净负电荷的分子(例如DNA或乙酰肝素硫酸盐)或与表面上具有正电荷补丁的蛋白质(例如FcRn)发生相互作用<sup>22</sup>。阿玛西单抗、布瑞金单抗、加尼图单抗和冷组力单抗在我们的实验中都结合DNA和肝素。表面上具有疏水性和带正电荷补丁的抗体可能具有增强的多反应性。基于我们的数据,能够通过疏水性补丁结合伴随蛋白和在中性pH下结合DNA、乙酰肝素硫酸盐或FcRn的抗体实例包括布瑞金单抗、阿玛西单抗和冷组力单抗。这些数据与布瑞金单抗所显示的非特异性结合的早期报道一致<sup>2</sup>。

[0824] 实例7.亲本克隆和改善型克隆的展示水平的量化

[0825] 在实例2中,在核酸酶介导的转基因整合到HEK293细胞中和选择稳定细胞系之后,我们观察到高等真核细胞针对三对抗体的展示水平差异。通过用PE标记的抗Fc将细胞染色以及利用流式细胞术测量平均荧光强度所判断的展示水平与抗体自身相互作用和交叉相互作用特性相关(实例3)。在此实例7中,我们对细胞表面上的抗体展示拷贝数进行量化并且表明展示拷贝数与抗体生物物理学特性相关。使用珠粒、基于校准曲线进行定量测量,其中珠粒具有精确限定数目个Fc特异性捕捉抗体。

[0826] Quantum Simply Cellular(QSC)微球体试剂盒(815,Bangs实验室有限公司)具有5个珠粒群-一个空白珠粒群和四个具有数量递增的Fc特异性捕捉抗体(山羊抗小鼠IgG)的珠粒群。QSC珠粒用相同的荧光团偶联抗体染色,所述荧光团偶联抗体用于标记细胞并且在流式细胞仪上根据制造商说明书加以分析。简单来说,将一滴QSC微球体添加到微量离心管中并且添加50μL染色缓冲液(1%BSA)并且轻弹所述离心管。将5uL的PE抗人类IgG Fc抗体添加到QSC微球体中,轻缓地混合并且暗中培育30分钟。QSC微球体用1mL洗涤缓冲液(含有0.1%BSA的PBS)洗涤两次,以2500xG离心5分钟。将珠粒集结粒再悬浮于150uL洗涤缓冲液中。将被染色的珠粒群与空白群体合并(每个群体10uL)于单个孔中并且在Intellicyt中运作。如实例2中所述,对细胞表面上表达不同抗体的HEK293细胞进行并行细胞染色。

[0827] 通过将每个珠粒群的中值荧光强度相对于其分配的抗体结合容量作图(图26)来产生校准曲线。将抗体表达群体的荧光强度与珠粒的抗体结合容量进行比较并且使用QuickCal(Bangs实验室有限公司)计算线性回归以能够计算抗体展示拷贝数(表9)。

抗体	拷贝数(x10 <sup>3</sup> )
CNT0607	113
CNT0607-W100A	313
MEDI-1912	48
MEDI1-912_STT	433
Ang2mAb	125
血管生成素 2-C49T	570
布瑞金单抗	273
乌 坦 库 单 抗 (Ustenikumab)	706

[0828]	派 立 珠 单 抗 (Pembrolizumab)	910
--------	------------------------------	-----

[0830] 表9. 使用Quantum Simply Cellular珠粒计算的展示于HEK293细胞上的抗体拷贝数

[0831] 临幊上批准的抗PD1抗体派立珠单抗(就低自身相互作用特性而言,已知具有良好的生物物理学特征<sup>2</sup>)在这个测试集合中具有最高拷贝数。最密集的校准珠粒具有886,000个拷贝/珠粒并且这超过所述拷贝数(约910,000个拷贝/细胞)。类似地,与治疗牛皮癣的III期人类临床试验中显示不良功效的抗IL-12抗体布瑞金单抗(273,000个拷贝)相比,批准用于临幊治疗克罗恩氏疾病的抗IL12抗体优特金单抗在细胞上展示更高拷贝数(706,000)。与优特金单抗相比,布瑞金单抗被描述为具有增强的自身相互作用和交叉特性,如多种分析(包括自身相互作用分析AC-SINS和交叉相互作用分析)中使用多特异性试剂和杆状病毒颗粒所测量<sup>2,23</sup>。

[0832] 对于所有三对抗体来说,拷贝数低于对照,经再次工程改造的具有改善的生物物理学特性的子系克隆具有较高的细胞展示拷贝数。举例来说,改善的子系抗体CNT0607-W100A、MED-1912\_STT和Ang2mAb\_C49的展示拷贝数分别为31.3万、43.3万和57.0万,拷贝数相较于具有已知自身相互作用和交叉相互作用问题的原始亲本分子增加2.8倍、9倍和4.6倍。

[0833] 在这个实例中,我们表明在核酸酶介导转基因整合到宿主基因组中且选择稳定细胞系之后展示于细胞表面上的抗体拷贝数与所展示抗体的生物物理学特性之间明确相关。在为了测量自身相互作用和交叉相互作用而设计的分析中,具有自身相互作用和交叉相互作用特性的抗体展示的拷贝数低于得分不高的抗体。表现较好的抗体(具有低自身相互作用和低交叉相互作用特性的良好生物物理学特征)在高等真核细胞表面上展示高拷贝数。

[0834] 实例8a. 将针对可开发性的高水平表达与针对亲和性严格度的低水平表达组合

[0835] 高水平多肽表达,例如其中抗体重链和轻链基因由强组成型启动子驱动的抗体,已经证明适用于从群体中富集具有优良生物物理学特性(例如低自身相互作用)的抗体(如

上文在实例3、4和5中所述)。细胞上的高浓度多肽呈现能有助于检测自身相互作用并且增强的亲合力允许灵敏地检测到与其它分子发生的非期望的非特异性相互作用。然而,这种情况的一个不利结果是,高亲和力所达成的富集速率相对于低亲和力减小,即使当使用低浓度的抗原驱动严格度时。这种“低强力”富集的另一结果是,表面呈现实际上可以存在优先选择,即使当目标是富集较高亲和力结合剂时。然而通过使用细胞表面上较低密度的结合剂呈现能够增强富集速率。在以下论述中,我们将使用抗体和其抗原作为代表不同亲和力的结合剂的一般相互作用的实例。

[0836] 如果我们考虑用100微升体积的 $10^6$ 个细胞展示 $6 \times 10^5$ 个单价结合位点/细胞,则我们具有 $6 \times 10^{11}$ 个分子/100毫升或 $6 \times 10^{15}$ 个分子/升。这个体积中的这种数目个结合位点相当于10nM的浓度。在这种情况下使用低于10nM的抗原浓度意味着抗体相对于抗原是过量的并且相对较高的抗体浓度将有助于驱动缔合,即使是较低亲和力抗体。举例来说,如果试图将表达具有 $K_D$  0.1nM的抗体的稀少细胞与表达具有 $K_D$  10nM的较低亲和力的抗体的过量细胞分离,则在上文概述的条件(即,10nM抗体浓度的等效条件)下,将有显著比例的抗原与较低亲和力抗体复合。

[0837] 质量作用定律涉及两种相互作用分子形成复合物。这些公式旨在涵盖自由溶液中的相互作用并且我们正在研究悬浮细胞上固定的抗体。下文的选择率计算考虑了单一抗体在溶液中在各种情况下所形成的复合物的浓度。然而,我们能够利用质量作用定律得到的溶液中的分子行为的这种知识更好地理解亲和力、抗体浓度、抗原浓度和复合物形成之间的关系,原因是其可能影响表面呈现的抗体。

[0838] 如果混合两种相互作用分子(例如单价抗体(A)和抗原(B)),则他们能够形成复合物(A:B)且最终达到平衡。平衡位置依赖于抗体和抗原的浓度,但能够通过解离常数 $K_D$ 描述如下(其中[A]、[B]和[AB]表示处于平衡时的浓度) :

$$K_D = \frac{[A][B]}{[AB]}$$

[0839]

$$\frac{[A][B]}{[AB]}$$

[0840] 这个方程式能够重新排列以计算在 $K_D$ 、A浓度和B浓度的不同条件下形成的复合物(AB)的浓度。

[0841] 如果结合反应处于平衡状态,

[0842] (1)  $A + B \leftrightarrow AB$

[0843] 则解离常数( $K_D$ )定义为:

$$(2) K_D = \frac{[A][B]}{[AB]}$$

[0845] 其中[A]、[B]和[AB]是反应物处于平衡时的浓度。反应物的总浓度( $A_T$ 和 $B_T$ ,其是添加到“试管”中的浓度)如下:

[0846] (3)  $[A_T] = [A] + [AB]$ ,其能够重新排列为 $[A] = [A_T] - [AB]$

[0847] (4)  $[B_T] = [B] + [AB]$ ,其能够重新排列为 $[B] = [B_T] - [AB]$

[0848] 将方程式3和方程式4代入方程式2:

$$(9) K_D = \frac{([A_T] - [AB])([B_T] - [AB])}{[AB]}$$

[0850] 重新排列方程式:

[0851] (10)  $K_D[AB] = ([A_T] - [AB])([B_T] - [AB])$

[0852] 将其相乘并且重新排列(为了清晰起见去除浓度括号):

[0853] (11)  $AB^2 - (A_T + B_T + K_D)(AB) + (A_T B_T) = 0$

[0854] 形成如下形式

[0855] (12)  $ax^2 + bx + c = 0$

[0856] 其中,

[0857] (14)  $a = 1$

[0858] (15)  $b = -(A_T + B_T + K_D)$

[0859] (16)  $c = (A_T B_T)$

[0860] 从而允许通过二次项方程式求解:

$$[0861] (16) x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$[0862] (17) AB = \frac{(A_T + B_T + K_D) - \sqrt{(A_T + B_T + K_D)^2 - 4(A_T B_T)}}{2}$$

[0863] 假设使用0.1nM的抗原浓度,图22a显示抗体(A)的不同浓度对不同亲和力的2种抗体的影响(Ab1的 $K_D$ 等于10nM或Ab2的 $K_D$ 等于0.1nM)。图22b显示2种抗体之间的相对“选择率”,其为当抗体个别地与0.1nM抗原一起培育时所形成的复合物的浓度与不同抗体浓度的比率。

[0864] 这表明尽管使用的抗原浓度远低于相互作用的 $K_D$ ,但当Ab1以10nM的浓度存在时,复合物浓度是50pM,代表着复合物中存在50%抗原。(10nM是在以上实例中根据 $6 \times 10^5$ 个拷贝/细胞的细胞表面展示所计算的浓度)。在使用高亲和力Ab2抗体的相同条件下,复合物浓度是99pM(代表99%的抗原),因此选择率仅存在2倍差异(表10)。然而,如果抗体浓度减少到0.1nM(在这个实例中相当于展示水平减小100倍),则复合物浓度就Ab1而言降低到1pM,而38.2pM的复合物浓度将通过Ab2实现(表10),代表着38倍的选择率。减小的密度还具有使靶标再结合的潜力减小的优势。在基于表面的亲和力测量(例如表面等离子体共振(BIAcore手册))时,在高密度的固定结合剂存在下再结合的问题已得到充分的认知和记载。

[0865] 在流式分选的情况下,通过检测荧光标记的抗原(直接或间接标记)在细胞表面上的存在来测量复合物的相对浓度。在流式细胞仪内检测的细胞上的荧光信号因此指示着在所用条件下在细胞上所形成的复合物的浓度。因此,在抗体呈现的更多限制性条件下,在呈现高亲和力抗体与低亲和力抗体的克隆之间实现较大的分离。

	[Ab(nM)]	Ab1 的 [ 复合物 (pM) ] ( $K_D$ 10 nM)	Ab2 的 \[ 复合物 (pM) ] ( $K_D$ 0.1 nM)	选择率
[0866]	10	49.9	99.0	2
	1	9.0	90.1	10
	0.1	1.0	38.2	39

[0867] 表10. 使用0.1nM抗原和不同浓度的高亲和力( $K_D$  0.1nM)和低亲和力( $K_D$  10nM)抗体在溶液中形成的复合物的浓度

[0868] 在抗体发现操作期间,在选择最优抗体的过程中,可能希望选择对其靶标具有较高亲和力的抗体。在抗体噬菌体展示选择期间增加严格度和富集对其靶标具有改善的亲和

力的克隆的方法是在选择期间减小靶抗原浓度<sup>57</sup>(Fellouse FA,Sidhu SS:通过细菌制备抗体(Making antibodies in bacteria).在:《制备和使用抗体:实用手册(Making and Using Antibodies:A Practical Handbook)》Howard GC,Kaser MR编辑:CRC出版社;2007:157-180<sup>108</sup>)。在哺乳动物展示选择时,为了能够选择性富集亲和力改善的克隆,还能够在FACS或MACS期间减小经标记的抗原浓度。然而,在抗体表达由强组成型启动子驱动的情况下所实现的高展示水平可以等同于抗体浓度高于所期望的靶标亲和力。因此,可能希望将HEK293细胞上的抗体细胞展示水平降低到例如低于60,000/细胞,从而在以上实例中得到10nM与0.1nM之间的10倍选择率。这能够使亲和力改善的抗体富集优于较低亲和力克隆。

[0869] 可以在抗体的转录、转录后、翻译或翻译后阶段通过若干不同方法降低抗体展示水平。举例来说,弱启动子可以用于减小初始mRNA转录物的产生速率,可以合并非最优的剪接/受体位点以减小成熟mRNA的生产效率和速率以及从细胞核输出到细胞质。可以减少mRNA的稳定性,从而减小转录物半衰期和有效浓度。表达的翻译控制可以通过改变Kozak共同序列以影响核糖体对mRNA的结合来达成。翻译后控制的一个实例可以是利用非最优前导序列降低输送到内质网的效率。

[0870] 在这个实例中,我们表明利用剪接/受体位点工程改造来降低抗体展示水平。这种降低的展示系统显示能够对HEK293细胞系混合物进行更高效的分离,所述细胞系混合物展示对其靶标具有不同亲和力的抗体。

[0871] 我们已设计了一种将表面展示与抗体分泌组合的方式,其基于B细胞中所用的天然系统。在B细胞成熟期间,抗体以膜结合形式表达并且随着浆细胞成熟将此切换到主要分泌形式。这符合真核细胞展示的要求,其中影响细胞表面展示的能力依赖于跨膜形式的表达。或者,克隆已选择后能够以分泌形式表达抗体将允许即刻产生游离的可溶性抗体用于进一步表征。

[0872] B细胞的表面展示与分泌之间的平衡大部分是由聚腺苷酸添加(引起IgG分泌)与剪接(引起膜系栓)之间的平衡驱动<sup>109,110</sup>。“近端”聚腺苷酸化位点发现于CH3域末端之后100-200bp处,并且这产生了mRNA,其在CH3域的末端终止翻译,从而产生分泌型产物。在CH3域末端附近,还存在潜在的剪接供体位点,其能够与下游外显子(M1)剪接以与“铰链”和跨膜域产生同框融合体。(M1外显子继而与编码胞内域的M2外显子剪接)。分泌型IgG呈现相对于膜结合型IgG呈现的平衡依赖于近端聚腺苷酸位点处的聚腺苷酸化与和M1外显子的剪接之间的平衡。这不同于最近公开的方法,其中使用2种替代外显子中的一种在分泌型与膜结合型之间切换<sup>111</sup>。

[0873] 剪接通常通过U1小核RNA(snRNA)发生,是起始剪接体组装、从而去除内含子所必需的。相较于共同剪接供体位点,CH3末端的剪接供体位点是次优的。非最优剪接供体在整个进化期间是保守的并且预期会与U1 snRNA形成非最优的碱基成对。实际上,已经显示,非最优剪接供体位点突变为共同剪接供体序列使编码主要分泌形式的经处理RNA与编码主要膜结合形式的RNA的平衡发生变化作为剪接增强的结果<sup>86</sup>。这代表着修饰剪接供体的一个早期实例,其改变了剪接相对于聚腺苷酸化之间的平衡,从而使分泌与聚腺苷酸化之间的平衡实现显著变化。基于Peterson等人的著作<sup>109,110</sup>,预期剪接供体优化程度会影响剪接与聚腺苷酸化之间的平衡,因此影响所展示抗体的比例。为了在膜形式与分泌形式之间找到最优平衡,在CH3外显子的末端建立许多替代的剪接供体位点。

[0874] 涉及剪接起始的U1 snRNA的序列在由IgG2 CH3域产生的mRNA序列上方显示。错配位置加有下划线：

[0875] U1 snRNA uc/cauuca

[0876] IgG2 CH3剪接供体 gg/guaaau

[0877] 围绕剪接供体设计四种变异数，包括野生型 (J9-GG/GTAAT)、部分优化 (J10-AG/GTAAA)、部分优化 (J29-GG/GTAAG) 和全部优化 (J30-AG GTAAG)，如图23所示。根据与U1snRNA的杂交，预期J30变异数会允许最有效的剪接，从而在细胞表面上产生更大比例的膜系拴抗体，相比之下，在“野生型”序列J9中，不太有效的剪接产生较低水平的膜系拴抗体和较大比例的分泌。J10和J29(部分优化的变异数)预期会使抗体展示水平在J9与J30之间的中间。

[0878] 构建展示抗体的靶向载体pINT17-BSD的变异数，其中将嵌入式HindIII限制位点添加到编码IgG1 CH3域C端和人类IgG内含子和M1外显子的DNA中，从而编码置换PDGFR跨膜域的跨膜域，所述PDGFR跨膜域由pINT17-BSD编码(实例1)。通过将合成基因合成、PCR组装和限制酶克隆进行组合来构建剪接供体变异数J9、10、29和30。pINT17-J30载体的标注DNA序列从XhoI到SbfI限制酶位点显示于图24中。XhoI-SbfI插入片段外部并且在图24中未示出的载体主链与pINT17-BSD一致(图1)。

[0879] 将编码抗PD1抗体尼沃单抗的VH和VL链的DNA克隆到四种剪接供体变异数靶向载体pINT17-J9、pINT17-J10、pINT17-J29和pINT17-J30中并且用于通过如实例2中所述的核酸酶介导的基因整合来建立稳定细胞系。为了比较，还将尼沃单抗克隆到图1所示的标准pINT17\_BSD载体中。这种构建体使抗体CH3域与PDGFr跨膜域直接融合而无需剪接并且在这个实例中称为pINT17-PDGFR。杀稻瘟菌素选择27天之后，细胞用PE标记的抗Fc染色并且如上文所述通过流式细胞术分析(图25)。这表明，与具有原生IgG跨膜域的构建体相比(其中内含子置于CH3与跨膜域之间)(图25a到d)，pINT17-PDGFR(由IgG1 CH3域与PDGFR跨膜域之间的直接融合表达的抗体)的抗体展示水平更大(图25e)。J30与U1 snRNA完全互补并且因此预期与编码跨膜域的M1外显子的剪接比J9、J10或J29变异数更高效。图25d表明显示最高水平的J30变异数实际上就是如此。然而其还表明，与整个早先实例中所用的pINT17-PDGFr构建体(图25d)上所发现的表达水平相比，J30变异数的表达水平甚至显著地更低。

[0880] 为了证明抗体展示拷贝数的减少能够有助于区分和分离具有不同亲和力的抗体，选择对其靶标具有不同亲和力的一对测试抗PD1抗体。所选抗体是尼沃单抗和另一种PD1抗体(337\_1\_C08)，通过表面等离子体共振(SPR)所测定，其分别具有3nM<sup>112</sup>和74nM的平衡解离常数( $K_D$ )亲和力。将编码抗PD1抗体的VH和VL链的DNA克隆到靶向载体pINT17-J30和pINT17-PDGFr(也称为pINT17-BSD)中并且用于通过核酸酶介导的基因整合如实例2中所述建立稳定细胞系。通过定量流式细胞术分析对抗体展示水平的量化如此前<sup>113</sup>和制造商说明书(Quantum Simply Cellular抗小鼠IgG珠粒，目录号815，Bangs实验室有限公司)所述，在杀稻瘟菌素选择后的第21天针对细胞进行(图26)。细胞上的拷贝数是通过测量经抗Fc-PE染色的细胞的中值荧光强度、与参考珠粒集合(目录号815，Bangs实验室有限公司)比较来计算(图26)。如表11A所示，对于两种抗体而言，与pINT17-BSD表达盒(CH3与跨膜域之间无内含子)相比，pINT17-J30(加内含子)表达盒的细胞展示拷贝数减少。

		拷贝数(x10 <sup>3</sup> )
[0881]	抗体	pINT17-BSD
	尼沃单抗	61
	337_1_C08	607
		18
		11

[0882] 表11A.拷贝数计算

[0883] 随着抗体拷贝数和抗原浓度减小,通过流式细胞术观察到的信号强度也减小(图25)。可以使用磁珠分选富集低于流式细胞术的灵敏度极限的经标记的细胞并且本文中使用这种磁珠分选。为了测试细胞表面上的减少的拷贝数是否会有助于分离具有不同亲和力的抗体,表达尼沃单抗或337\_1\_C08的细胞系各自用不同荧光团染料染色。这样可以观察到一种细胞相对于另一种细胞的相对富集。将经标记的细胞混合,与不同浓度的生物素化PD1一起培育并且通过MACS分离。然后进行流式细胞术分析以确定表达对PD1具有较高亲和力的抗体的细胞是否富集。在这个实例中测试四种HEK293细胞系,其初始经以下转染:

[0884] a.pINT17-BSD-尼沃单抗

[0885] b.pINT17-J30-尼沃单抗

[0886] c.pINT17-BSD-337\_1\_C08

[0887] d.pINT17-J30-337\_1\_C08

[0888] 上方显示的pINT17-BSD表达盒(抗体CH3与跨膜域之间的直接融合体)表达的抗体展示水平高于pINT17-J30表达盒。,根据制造商说明书,使用细胞追踪绿(50nM,C7025,赛默飞世尔)将表达尼沃单抗的细胞( $5 \times 10^6$ )染色,并且使用细胞追踪深红(50nM,C34565,赛默飞世尔)将表达337\_1\_C08的细胞染色。简单来说,细胞用PBS洗涤,与含有追踪染料(50nM)的PBS一起培育并且在37°C下培育10分钟。然后用PBS洗涤细胞。由表达尼沃单抗的预染色细胞或衍生自pINT17-J30表达盒的337\_1\_C08以1:1比率( $5 \times 10^6$ 个细胞,各自体积为10ml)混合。使混合的细胞集结(100g,3分钟)并且将每个细胞集结粒再悬浮于含有0、0.1、1或10nM生物素化PD1(PD1-H82E4,AcroBiosystems)的PBS(1ml)中并且在4°C下培育30分钟。细胞用0.1%BSA、PBS和抗生素链菌素珠粒(10μl)洗涤并且将90μl的1%BSA添加到每个样品中且混合。样品在冰箱(4°C)中培育15分钟。使用2ml的0.1%BSA洗涤细胞并且以200×G离心4分钟。将集结粒再悬浮于500μl分离缓冲液(MACS漂洗缓冲液,其由1×PBS+2mM EDTA+0.5%BSA组成)。LS柱用3ml分离缓冲液洗涤。将细胞悬浮液添加到柱中,每个柱一个样品。未捕获的细胞通过用3ml分离缓冲液洗柱三次来收集并且收集流过物。将柱置于一个新收集管中并且将5ml分离缓冲液添加到每个柱中并且使用所提供的柱活塞立即清空。使用细胞计数器和锥虫蓝对洗脱样品和流过样品进行计数,以测定细胞回收率。使每个样品 $1 \times 10^6$ 个细胞在500μl 0.1%BSA/PBS中稀释。通过流式细胞术分析50μl经稀释的细胞。

[0889] 当使用低密度展示载体pINT17-J30时,可以选择性地将展示低亲和力抗PD1抗体337\_1\_C08的细胞系与表达较高亲和力抗PD1抗体尼沃单抗的细胞分离。这显示于图27a中,其中展现了MACS洗脱或流过群体的点阵图流式细胞术结果。每个图显示点阵图,其中与0、0.1、1或10nM生物素化PD1一起预培育的洗脱和流过洗脱份的绿色荧光强度在x轴上绘制(FL1,表达尼沃单抗的细胞)且红色荧光强度在y轴上绘制(FL4,表达337\_C08的细胞)。举例来说,所述细胞系与1nM生物素化PD1混合培育、随后进行MACS分离之后,与红色染料染色的较低亲和力抗PD1337\_1\_C08抗体相比,绿色染料染色的高亲和力抗PD1尼沃单抗展示细胞

发生95%富集。相比之下,流过物中的未结合细胞主要含有较低亲和力的红色染色抗PD1337\_1\_C08抗体展示细胞。此外,与10nM PD1培育相比,观察到1nM PD1存在良好的分离。然而,当使用衍生自pINT17-BSD盒的较高拷贝数细胞系进行相同实验时,观察到尼沃单抗与337\_1\_C08表达细胞之间不存在优先富集(图27b)。

[0890] 这个实例因此已经说明,降低细胞表面上的抗体展示水平能够增加高亲和力抗体相对于较低亲和力的富集,尤其是在所述群体内实现高亲和力时。高亲和力抗体与对其靶标亲和力较低的抗体的最优细胞展示分离是通过将抗体展示水平的降低与选择期间所用的经标记靶抗原的浓度的减小组合来实现。根据所期望抗体的所需候选靶标特征,在真核细胞展示选择期间根据亲和力分离抗体的能力在抗体发现操作期间是重要的。在亲和力成熟期间富集对其靶标的亲和力改善的抗体也是重要的。这个实例已经表明细胞展示拷贝数帮助区分对其靶标具有高亲和力的抗体的重要性。

[0891] 实例8b. 使用诱导型tet启动子的改善

[0892] 建立能够使哺乳动物细胞展示的抗体得到诱导型表达的载体是有利的。这促进了选择步骤在不同表面呈现水平的组合,例如严格选择所需的低展示水平选择之后可以是选择具有改善的生物物理学特性的抗体所需的高展示水平诱导,如实例4和5中举例说明。

[0893] 通过将合成基因合成、PCR组装和限制酶介导的克隆组合来构建展示诱导型抗体的靶向载体(pINT18-Tet1)(图28)。pINT18-Tet1含有与pINT17-BSD相同的载体主链、AAVS同源臂和无启动子杀稻瘟菌素抗性基因。图28显示AAVS同源臂之间的pINT18-Tet1的标注核酸序列。这种载体的关键特征包括驱动反向Tet活化因子(rtTA)蛋白质表达的CMV启动子<sup>114</sup>、Tet操纵子(TetO)四联体,随后为最小CMV启动子、BM40前导序列与轻链基因(VL和CL)的融合、能够实现核糖体“跳过”的P2A肽<sup>115</sup>,随后为抗体重链编码序列。将派立珠单抗、尼沃单抗和337\_1\_C03的重链和轻链可变基因(VH和VL)克隆到诱导型靶向载体中并且如上文所述通过核酸酶介导的基因整合来建立稳定的HEK293细胞系。图29显示当细胞用抗Fc-PE染色并且通过流式细胞术分析时,在缺乏多西环素的情况下低基本抗体展示。然而,在添加20ng/ml多西环素之后的第24小时,观察到细胞表面上的抗体表达。当多西环素浓度滴定到2ng/ml时,能够控制这种展示水平,如根据Fc染色减少所示。因此,显示了诱导型细胞所展示的结合剂的范例和通过改变诱导剂浓度来控制展示水平的能力。

[0894] 随后构建改善的(“第三代”)诱导型靶向载体以能够实现改善的抗体展示水平范围。通过将合成基因合成、PCR组装和限制酶介导的克隆组合来构建pINT17-Tet。pINT17-Tet含有与pINT17-BSD相同的载体主链、AAVS同源臂和无启动子杀稻瘟菌素抗性基因。图37显示AAVS同源臂之间的pINT17-Tet的标注核酸序列。这种质体包括驱动反向Tet活化因子(rtTA-3G)表达的延伸因子启动子。rtTA-3G是原始rtTA蛋白质的经修饰形式<sup>114</sup>,其进化而对诱导剂多西环素更敏感<sup>116</sup>。pINT17-Tet也含有驱动抗体重链和轻链表达的双向诱导型启动子(pTRE3G)。pTRE3G经优化而在诱导之后拓宽低基础表达与高最大表达之间的窗口<sup>117</sup>。其由以下组成:19bp tet操纵子(TetO)序列的7个重复,和两个侧接的最小CMV启动子。

[0895] 将抗PD-1抗体:1549\_02\_D06、1535\_01\_E03和337\_1\_C08以及抗PCSK9抗体博可珠单抗、884\_01\_G01、5A10i和阿利库单抗的VH和VL基因克隆到pTet17-Tet中。实例8中描述了对PD1具有74nM亲和力的抗PD1抗体337\_1\_C08。抗体1549\_02\_D06和1535\_01\_E03是亲本抗体337\_1\_C08的亲和力成熟子系克隆,其对PD-1分别具有2.9nM和17nM的亲和力( $K_p$ )。还将

抗PCSK9抗体博可珠单抗、884\_01\_G01、5A10i. 和阿利库单抗(此前都在实例5中描述)的VH和VL基因克隆到pINT17-Tet中。

[0896] 使用具有抗PD1和抗PCSK9抗体基因的pINT17-Tet靶向载体,使用编码AAVS TALE核酸酶的质体如上文所述转染HEK293细胞。杀稻瘟菌素药物选择25天后,用0、2、4或100ng/ml多西环素诱导稳定的细胞系。诱导后24小时,将诱导的细胞系用抗Fc-PE染色,并通过流式细胞术分析细胞。通过将细胞数目相对于荧光强度作图来产生细胞系直方图。

[0897] 在缺乏多西环素的情况下,细胞系显示出非常低的基础表达(图38a)。诱导后的第24小时,通过添加浓度为100ng/ml的多西环素观察到充分的诱导(图38d)。通过添加浓度为2和4ng/ml的多西环素来实现中等诱导(图38b和38c)。这使得平均展示水平低于充分诱导,但诱导的展度类似于此前使用七个TetO重复所观察到的双峰基因表达<sup>118</sup>。

[0898] 利用实例7中所述的方法定量抗体的平均展示拷贝数。在此,测定细胞系被不同量的多西环素诱导的平均荧光强度平均值(MFI)并且使用校准图将其换算为拷贝数,所述校准图使用经小鼠IgG-PE标记染色的Quantum Simply Cellular抗小鼠IgG珠粒(目录号815, Bangs实验室有限公司)产生。表11c和11d显示诱导后(hpi)第24或48小时的所计算展示拷贝数。在缺乏多西环素的情况下观察到抗体的极低基础表达,在一些情况下,当用抗Fc-PE染色时,所述极低基础表达下降而低于检测极限。所有抗体在诱导后48小时期间存在可检测的基础表达并且这可能是因为与诱导后24小时相比,细胞已达到平稳生长期。随着多西环素浓度从2ng/ml到4ng/ml增加到100ng/ml多西环素,所有抗体的平均展示水平增加。多西环素浓度增加而高于100ng/ml未引起展示水平增加并且因此以100ng/ml多西环素的浓度实现最大诱导。

[0899] 在倾向于自身相互作用和多反应性的抗PCSK9 IgG博可珠单抗<sup>2</sup>与表现良好的亲本抗体5A10i之间观察到展示水平差异。当细胞用100ng/ml多西环素诱导并且在诱导后(hpi)24小时测定展示水平时,亲本抗体5A10i以235,000的拷贝数展示于HEK293细胞表面上,而博可珠单抗以36,000的拷贝数展示于HEK293细胞表面上。这表示博可珠单抗的展示水平相较于亲本抗体5A10i降低6.5倍。当抗体表达用组成型启动子驱动时,与5A10i相比,博可珠单抗的细胞展示拷贝数的这种所观察的减少在此前已观察到(实例5,图10)。因此,这种诱导型系统在充分诱导下能够区分具有不同可开发性特征(按照自身相互作用和多反应性)的抗体,如我们此前使用组成型启动子展示系统所证明。

[0900] 我们还展示了通过哺乳动物展示文库筛选所鉴定的博可珠单抗变体,其命名为884\_01\_G01(实例5)。AC-SINS分析和HPLC-SEC表明这种抗体在生物物理学方面表现良好(表6、实例5和图18、实例5)。诱导后24小时,这种抗体在HEK293细胞表面上高水平展示(100ng/ml多西环素),展示拷贝数为768,000(表11B)。在充分诱导条件下诱导后24小时,表现良好的抗体阿利库单抗<sup>2</sup>在HEK293细胞表面上也高水平展示(表11B)。

[0901] 多西环素已知在培养基上清液中降解,半衰期为约一天,此视培养条件而定。在不补充多西环素的情况下在诱导后的不同时间点检查抗体展示水平提供了关于抗体在细胞表面上动态代谢的一些证据。或者,诱导可以在多西环素存在下进行24小时时间段,随后更换整个培养基,从而使细胞不再暴露于多西环素。在此,我们简单地重新检查诱导后48小时的展示水平(表11C)。诱导后48小时,表现良好的若干抗体(例如阿利库单抗)维持或增加其细胞表面展示水平。与诱导后24小时相比,其它抗体(例如博可珠单抗)在诱导后48小时的

细胞表面展示水平降低大于2倍(图39)。细胞上所展示的抗体的代谢、降解或内化速率(其中所展示的抗体未经新表达的抗体连续补充)是基于其可开发性特征(按照自身相互作用和稳定性)选择抗体的另一方式(还参见实例11)。

	拷贝数(x10 <sup>3</sup> )				
	Dox(0 ng/ml)	Dox(2 ng/ml)	Dox(4 ng/ml)	Dox(100 ng/ml)	
[0902]	1549_02_D06	Nd	4.5	21.9	356
	1535_01_E03	Nd	3.3	16.5	137
	337_1_C08	Nd	5.6	31.0	251
	博可珠单抗	Nd	2.4	5.1	36.1
	884_01_G01	Nd	26.7	115	768
	5A10i	0.1	14.0	48	235
	阿利库单抗	0.4	32.8	144	557

[0903] 表11B.在0、2、4或100ng/ml多西环素(Dox)诱导的情况下诱导后24小时,使用Quantum Simply Cellular珠粒计算的HEK293细胞上所展示的抗体拷贝数.nd:不可检测。

	拷贝数(x10 <sup>3</sup> )				
	Dox(0 ng/ml)	Dox(2 ng/ml)	Dox(4 ng/ml)	Dox(100 ng/ml)	
[0904]	1549_02_D06	0.4	11.1	51.5	317
	1535_01_E03	0.7	6.1	46.8	238
	337_1_C08	0.4	20.4	108	294
	博可珠单抗	0.7	2.0	5	17.5
	884_01_G01	0.6	26.4	112	417
	5A10i	5.2	20.7	40.2	73.5
	阿利库单抗	1.7	101	339	668

[0905] 表11C.在0、2、4或100ng/ml多西环素(Dox)诱导的情况下诱导后48小时,使用Quantum Simply Cellular珠粒计算的HEK293细胞上所展示的抗体拷贝数。

[0907] 因此,pINT17-Tet代表着诱导型哺乳动物展示靶向载体的一个实例,其能够用于通过核酸酶介导的基因整合来建立单克隆细胞系,所述基因整合能够切换为高表达、充分诱导的模式以实现上文关于抗体自身相互作用和多反应性筛选所述的可开发性筛选。在缺乏诱导剂或添加有限浓度的多西环素的情况下,使用pINT17-Tet建立的细胞系还能够切换为基础展示水平的低拷贝展示模式。这种低拷贝数展示模式能够使具有不同亲和力的抗体克隆实现更高效的分离,如实例8a中所述。

[0908] 为了证明诱导型哺乳动物展示细胞系用于分离细胞表面上所展示的对其靶标具有不同亲和力的抗体的效用,通过核酸酶介导的通过pINT17-Tet靶向载体发生的基因整合来建立展示高亲和力抗PD1抗体1549\_02\_D06(对于PD-1而言,K<sub>D</sub>=2.9nM)或低亲和力抗PD-1抗体337\_1\_C08(对于PD-1而言,K<sub>D</sub>=74nM)的HEK293细胞。展示1549\_02\_D06抗体的Hek293细胞用CellTrace CFSE细胞增殖试剂盒(Thermo目录号C34554-激发/发射波长-492nm/517nm)标记并且展示337\_1\_C08抗体的细胞未经标记。用0、2、4或100ng/ml多西环素诱导细胞。诱导后(hpi)48小时,展示1549\_02\_D06的细胞用CellTrace CFSE细胞增殖试剂盒

(C34554, 赛默飞世尔科学) 染色并且与展示337\_1\_C08的未标记细胞混合。经标记的细胞与未标记的细胞等量混合并且然后使用0.1、1或10nM PD-1-生物素对混合的IgG展示细胞群进行MACS提纯, 如实例8a中所详述。然后通过流式细胞术分析细胞(图40)以测定高亲和力抗PD1 IgG相较于低亲和力抗PD1 IgG的相对富集。

[0909] 当细胞用2ng/ml有限浓度的多西环素诱导以降低抗体展示水平并且使用0.1nM PD-1-生物素进行MACS提纯时, 展示高亲和力抗PD-1抗体的细胞与展示低亲和力抗PD-1抗体的细胞发生最有效的分离。在此, 将群体中的高亲和力克隆富集到群体中的96% (图40bi)。当用2ng/ml多西环素诱导细胞时, 诱导后48小时的细胞拷贝数就高亲和力和低亲和力抗PD-1克隆而言分别是11,000和20,000。相比之下, 当细胞系被充分诱导时, 仅使用0.1nM PD-1的MACS实现高亲和力克隆的77%富集(图40di)。当细胞用100ng/ml多西环素充分诱导时, 诱导后48小时的细胞拷贝数就高亲和力和低亲和力抗PD-1克隆而言分别是317,000和294,000。除实例8中的数据之外, 这进一步证明了细胞表面抗体展示水平(拷贝数)的降低将使对其靶标具有不同亲和力的展示抗体的分离效率增加。随着用于MACS的PD-1浓度增加, 这使得高亲和力克隆的富集效率减小, 如根据实例8中的理论分析所预期(图40bii和40iii)。

[0910] 在这个实例中, 我们已经证明诱导型启动子系统用于高等真核细胞展示的效用能够根据自身相互作用和聚集倾向筛选实现高拷贝数展示并且根据高亲和力抗体的严格选择实现低拷贝数展示。在此, 我们使用有限浓度的诱导剂得到低展示水平, 以便将对其靶标具有2.9nM亲和力的克隆与对相同靶标具有74nM亲和力的第二克隆有效分离。设想通过控制细胞表面上所展示的抗体拷贝数和用于MACS分离的靶抗原浓度, 会使对其靶标的亲和力( $K_D$ ) 小于1nM的克隆与具有个位数nM亲和力的克隆的分离效率增加。

[0911] 实例9. 根据与Fc受体的最优相互作用、通过选择来增强可开发性

[0912] 实例6证明能够将展示具有不同交叉相互作用特性的抗体的真核细胞区分开来。这是通过将细胞与经标记的人类血清一起培育并且利用流式细胞术检测结合来实现的。与临幊上批准的抗体尼沃单抗(被视为表现良好的抗体)相比, 韦森单抗(一种具有已知交叉相互作用特性<sup>39</sup>的抗体, 其未进一步通过1期临床试验<sup>82</sup>)对人类血清的结合更大。韦森单抗的体内半衰期短<sup>105</sup>。抗体的“粘着”或交叉相互作用特性通常与不良药物动力学相关并且体内短半衰期被认为是因非特异性结合到弥散性组织而引起的全身性清除所致<sup>23,122,123</sup>。

[0913] 体内抗体半衰期也关键地依赖于IgG(通过其CH2和CH3域)与新生儿Fc受体之间的相互作用<sup>124</sup>。循环中的抗体长半衰期是通过IgG的细胞内化、通过非特异性胞饮或Fc受体介导的吸收来实现。一旦内化, 则IgG以高亲和力结合到核内体(pH5-6)内的FcRn, 借此保护IgG以免溶酶体降解。由于IgG与FcRn之间的亲和力在生理pH 7.4下极弱, 因此IgG最终被输送到细胞表面并且释放回至循环中。抗IL12布瑞金单抗<sup>20</sup>在治疗牛皮癣的III期临床试验中未显示功效并且已知在体内具有短半衰期。布瑞金单抗在其可变域内具有由精氨酸和赖氨酸残基组成的正电荷补丁, 所述正电荷补丁能够结合到FcRn上的负电荷补丁, 从而增加在pH7.4下结合到FcRn的亲和力。布瑞金单抗在pH7.4下对FcRn的结合已经显示与小鼠体内的不良半衰期相关。此外, 临幊上批准治疗克罗恩氏病的抗IL12抗体优特金单抗在其可变域内不具有正电荷补丁, 已知在pH7.4下弱结合到FcRn并且体内半衰期优于布瑞金单抗。在这个实例中, 我们表明能够区分在pH7.4对FcRn具有已知的不同亲和力的由高等真核细胞展

示的抗体并且通过流式细胞术检测FcRn结合。

[0914] 将编码布瑞金单抗和优特金单抗重链和轻链可变域(序列参见表12)克隆到pINT17-BSD靶向载体(载体图谱和序列参见实例1)中并且证实DNA序列。

[0915]	布瑞金单抗 VH	QVQLVESGGVVQPGRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAFIRY DGSNKYYADSVKGRTFTISRDNSKNTLYLQMNSLRAEDTAVYYCKTHGSHDNWG QGTMVTVS
	布瑞金单抗 VL	QSVLTQPPSVSGAPGQRVTISCSGSRSNIGSNTVKWYQQLPGTAPKLLIYYND QRPSGVPDFSGSKSGTSASLAITGLQAEDAEADYYCQSYDRYTHPALLFGTGT KVTVLGQP
	优特金单抗 VH	EVQLVQSGAEVKKPGESLKISCKGSYSFTTYWLGWVRQMPGKGLDWIGIMSP VDSDIRYSPSFQGQVTMSVDKSITTAYLQWNSLKASDTAMYYCARRRPGQGYF DFWGQGTLVTVS
	优特金单抗 VL	DIQMTQSPSSLSASVGDRVТИTCRASQGISSWLAWYQQKPEAKPKSLIYAASS LQSGVPSRFSGSGSGTDFLTISLQPEDFATYYCQQYNIYPYTFGQGTKLEI KRTA

[0916] 表12.布瑞金单抗和优特金单抗VH和VL链的序列

[0917] 转染前一天,以 $5 \times 10^5$ 个细胞/毫升接种经调适的HEK293细胞于HEK FreeStyle 293表达培养基中的悬浮液。转染当天,将细胞离心并且以 $10^8$ 个细胞/毫升的最终体积再悬浮于制造商的电穿孔缓冲液(Maxcyte电穿孔缓冲液,赛默飞世尔科技,目录号NC0856428)。使用OC-100光析管(Maxcyte)对100 $\mu$ l进行脉冲。转染之后的两天,以 $2.5 \times 10^5$ 个细胞/毫升接种细胞并且添加7.5ug/ml杀稻瘟菌素。转染之后的57天,细胞用与抗生蛋白链菌素PE(11nM)预偶联的生物素化FcRn(50nM)标记。在不同pH的染色缓冲液(含有1%BSA的PBS, pH7.4,或含有140mM NaCl和1%BSA的20mM MES, pH6.0)中进行预偶联。简单来说,将 $1 \times 10^6$ 个细胞离心,用PBS pH 7.4或20mM MES缓冲液pH 6.0洗涤一次。将细胞集结粒再悬浮于100ul染色缓冲液中并且用1.0ml洗涤缓冲液(含有0.1%BSA的PBS pH7.4,或含有140mM NaCl和0.1%BSA的20mM MES pH6.0)洗涤两次。将细胞集结粒再悬浮于含有活力染料7AAD的500ul洗涤缓冲液(5ul/百万个细胞)中。使用Intellicyt流式细胞仪分析细胞。分析期间排除死细胞。使用FlowJo软件(图30)产生直方图。

[0918] 当抗体展示于HEK293细胞表面上时,布瑞金单抗与优特金单抗在pH6下均结合到FcRn(图30)。不同于FcRn结合未观察到的优特金单抗,只有布瑞金单抗展示在pH 7.4显著结合到FcRn。这个实例证明能够区分在pH7.4下对FcRn具有不同结合亲和力的由高等真核细胞展示的克隆以及通过流式细胞术分析FcRn结合。这将能够预测抗体药物动力学(PK)特征。此外,当从细胞展示的结合剂文库中选择时,这种技术可以用于排除在pH7.4下结合到FcRn的克隆且因此选择预期具有有利PK特征的克隆。举例来说,细胞展示的抗体对MACS珠粒的结合可以在pH6实现。在pH7.4洗涤珠粒应该洗脱在pH7.4具有低亲和力的克隆并且预期这些克隆展示出在体内的半衰期比在pH7.4保持结合的结合克隆长。

[0919] FcRn对IgG的结合已知部分地依赖于IgG的翻译后糖基化。因此,使用高等真核细胞执行所述选择的优势在于,不同于在未修饰的低等真核细胞展示系统(例如酵母)中展示,发生的可靠糖基化能够实现FcRn结合。

[0920] 我们进一步证明能够从混合抗体群体中分离出对FcRn的结合减少的抗体。将分别

表达不同抗IL-12抗体(布瑞金单抗和优特金单抗)的高等真核细胞克隆的两种群体共培养。表达布瑞金单抗的细胞显示在pH 7.4(图30)结合到FcRn。结合可能归因于其可变域内的其正电荷补丁结合到FcRn的带负电侧链(参见实例6,其中布瑞金单抗显示结合带负电的多反应性探针)。优特金单抗在pH 7.4下未显示对FcRn的结合(图30)。混合抗体表达群体经双重染色且通过FACS富集展示优特金单抗的细胞。图41.我们因此证明从也含有FcRn结合剂的混合群体中能够富集FcRn非结合剂。相同技术可以应用于更多样化的克隆混合物,例如含有数百万种不同抗体的文库。

[0921] 抗体Fc区已知诱发多种效应功能,包括Fc受体<sup>139</sup>和补体结合<sup>140</sup>。通过建立Fc域文库,随后使用已知效应分子(例如Fc  $\gamma$  受体、NK受体、FcRn,或补体级联成员)进行选择,能够利用高等真核细胞展示来选择对效应功能分子的结合增强或减小的变异数。这会选出可以增强、减小或静默抗体依赖性细胞的细胞毒性(ADCC)、补体依赖性细胞毒性(CDC)或延长体内抗体半衰期的Fc变异数。还可以选择结合到Fc受体的Fc变异数以增加其对靶标受体的促效作用且因此增加其效能<sup>141,142</sup>。Fc变异数能够减小抗体稳定性并且对其生物物理学特性<sup>134,143</sup>产生不利影响。通过高等真核细胞展示选择Fc变异数的一个优势在于,能够选出高水平展示于细胞表面上的变异数并且预期这些变异数具有优良的生物物理学特性。

[0922] 实例10.通过哺乳动物展示来选择所选群体内的最优可开发性克隆(间皮素选择)

[0923] 在此,我们证明利用高等真核(哺乳动物)细胞展示能够从一组高度多样化的输入抗体中选择具有最优可开发性潜力的抗体。分离不同生殖系序列的抗体的复杂混合物,从而允许基于抗体的自身相互作用特性将抗体分组且选择。

[0924] 对此我们如下证明:首先通过抗体噬菌体展示选择来产生一组富集的抗间皮素抗体,然后将抗体克隆到哺乳动物细胞展示文库中并且对来自分别展现高、中或低表面呈现的克隆的结合剂的特征进行比较。

[0925] 间皮素是过度表达于若干人类肿瘤中的抗原并且因此作为治疗癌症的治疗抗体靶标而受到关注。为了产生抗人类间皮素的富集群体,进行两轮抗体噬菌体展示选择,其中抗原直接固定于聚苯乙烯表面上,如此前所述<sup>8</sup>。使用初始人类抗体文库(WO2015166272A2),其如此前所述并且以与此前描述的人类初始抗体文库相似的方式构建<sup>8</sup>。分别使用VL $\lambda$ 和VL $\kappa$ 生殖系进行选择并且第1轮输出数目是 $2 \times 10^5$ (平均值)。如此前所述(WO2015166272A2),转化为IgG形式和克隆到pINT17-BSD靶向载体中都“一同”进行。如实例4和5中所述将 $\lambda$ 和 $\kappa$ 富集群体中的大肠杆菌转型体接种到琼脂板上,以建立 $2.4 \times 10^6$ 个克隆(12倍过量的输入抗体群体)的组合文库大小。制备转染质量的质体DNA,其编码pINT17-BSD载体中的抗间皮素抗体。使用如上文(实例2、4和5)所述的AAVS TALE核酸酶,使用单个OC-400光析管(总计 $10^8$ 个细胞,MaxCyte,目录号OC-400R10),在MaxCyte STXG2上的HEK293配置下转染文库。对照组是使用OC-100光析管(MaxCyte,目录OC-100R10)在相同配置下转染。将电穿孔的细胞转移到适当尺寸的锥形瓶之后,允许细胞在添加适量FreeStyle 293表达培养基(LifeTech,目录号12338018)之前搁置30分钟。使细胞再悬浮并且置于设定为130rpm、37°C和5%CO<sub>2</sub>的回转式振荡培育箱中。稳定的转染效率如上文所述计算(实例4和5)为4.2%,得到 $4.2 \times 10^6$ 个稳定抗体展示HEK293细胞系的总文库大小。杀稻瘟菌素选择19天之后,如上文所述(实例4和5)使用BD Influx细胞分选仪相应地进行FACS。将 $100 \times 10^6$ 个细胞与抗人类Fc PE( $1\mu\text{l}/1 \times 10^6$ 个细胞)(剑桥生物科学,目录号409304)一起培育并且即将

分选之前添加DAPI (1 $\mu$ l/百万细胞)。绘制了三个门,从而获得展示水平低的群体 (P4)、展示水平中等的群体 (P6) 和展示水平高的群体 (P5)。Influx分选仪仅能一次分选两种群体,因此门P4和P5分选 $50 \times 10^6$ 个细胞,并且P6分选其它 $50 \times 10^6$ 个细胞(图31)。如上文所述(实例4和5)从三个群体中提取基因组DNA,将抗体基因克隆到pINT3载体中(WO2015166272A2)并且对VH和VL基因测序。

	输入	低展示	中等展示	高展示
测序总数	79	89	92	85
独特的 VH CDR3	56 (71%)	38 (49%)	62 (67%)	63 (74%)
独特的 VL CDR3	49 (62%)	32 (36%)	57 (62%)	48 (56%)
独特的 VH + VL CDR3	69 (87%)	43 (48%)	79 (89%)	71 (84%)
VH 生殖系	14	10	10	12
VL $\kappa$ 生殖系	11	12	10	11
VL $\lambda$ 生殖系	3	2	5	5

[0926] [0927] 表13. 抗间皮素哺乳动物展示所选克隆群体的序列分析。对FACS中根据低、中和高展示水平门控的群体(图30)进行DNA序列分析并且与用于产生起始抗间皮素展示文库的输入抗体产生进行比较。生殖系以及VL CDR1、2和3以及VH CDR1和2遵循命名法,而VH CDR3是根据Kabat编号系统定义。

[0928] 序列分析揭露所有亚群内存在较大多样性(表13)。为了确保这不归因于取样问题,我们将分析聚焦于出现超过一次的克隆以确定不同克隆组是否“分组”成不同展示组。因此,对抗体进行重叠分析,其中在展示水平低、中和高的群体中已经鉴定出双重复克隆(基于VH和VL CDR3序列)(表14)。展示水平高的群体中存在13种此类克隆,并且28种克隆在展示水平低的群体中出现多次,其间无重叠。鉴定为“中等表达”的克隆组中有11种序列出现超过一次并且另外,这些序列在其它组中未发现。“中等”组中出现4种克隆,其在3种情况下与低表达组重叠并且在一种情况下与高表达组重叠。这种结果是基于哺乳动物细胞的展示水平进行的分选选择该组所独有的序列的特定亚群。

克隆编号	VH 生殖系	VL 生殖系	高展示组中出现的次数	中等展示组中出现的次数	低展示组中出现的次数
1	IGHV1-69	IGKV2-28	3	0	0
2	IGHV4-34	IGLV2-8	2	0	0
3	IGHV1-46	IGKV1-5	2	0	0
4	IGHV1-46	IGKV6-21	2	0	0
5	IGHV3-9	IGLV3-19	2	0	0
6	IGHV3-9	IGLV3-19	2	0	0

[0930]	7	IGHV1-46	IGKV3-15	3	0	0
	8	IGHV3-9	IGLV3-19	1	1	0
	9	IGHV3-9	IGLV3-19	1	1	0
	10	IGHV3-9	IGLV3-19	2	0	0
	11	IGHV1-2	IGKV3-15	2	0	0
	12	IGHV3-9	IGLV3-19	3	0	0
	13	IGHV3-9	IGLV3-19	1	1	0
	14	IGHV3-9	IGLV3-19	0	0	2
	15	IGHV5-51	IGKV1-39	0	0	6
	16	IGHV1-3	IGKV1-39	0	0	10
	17	<b>IGHV1-2</b>	<b>IGKV4-1</b>	<b>0</b>	<b>0</b>	<b>4</b>
	18	IGHV1-18	IGKV1-33	0	0	3
	19	IGHV1-46	IGKV2-30	0	0	3
	20	IGHV1-69	IGKV1-39	0	0	3
	21	IGHV1-69	IGKV1-39	0	0	2
	22	<b>IGHV1-69</b>	<b>IGKV4-1</b>	<b>0</b>	<b>0</b>	<b>5</b>
	23	<b>IGHV3-30/33rn</b>	<b>IGKV4-1</b>	<b>0</b>	<b>0</b>	<b>4</b>
	24	IGHV3-9	IGLV3-19	0	0	3
	25	IGHV3-9	IGLV3-19	0	0	2
	26	IGHV1-46	IGKV1-39	0	0	2
	27	IGHV1-69	IGKV1-39	0	1	3
	28	<b>IGHV5-51</b>	<b>IGKV4-1</b>	<b>0</b>	<b>0</b>	<b>2</b>
	29	IGHV3-30/33rn	IGKV2-28	0	0	2
	30	<b>IGHV1-69</b>	<b>IGKV4-1</b>	<b>0</b>	<b>0</b>	<b>6</b>
	31	IGHV3-9	IGKV1-39	0	2	0
	32	IGHV3-9	IGLV3-19	0	2	0
	33	IGHV3-9	IGLV3-19	0	2	0
	34	IGHV1-69	IGKV2-28	0	2	0
	35	IGHV3-9	IGLV3-19	0	2	0
	36	IGHV3-9	IGKV3-20	0	2	0
	37	IGHV3-9	IGLV3-19	0	3	0
	38	IGHV3-9	IGLV3-19	0	2	0
	39	IGHV3-9	IGLV3-19	0	2	0
	40	IGHV3-9	IGLV3-19	0	2	0
	41	IGHV3-9	IGLV3-19	0	2	0

[0931] 表14. 根据HEK293细胞高、中等和低展示水平混合的抗间皮素抗体的抗体克隆重叠分析。根据细胞表面展示水平混合根据哺乳动物细胞展示所选的抗间皮素抗体(图31)。对来自展示水平高、中等和低的群体的克隆进行DNA测序并且根据高、中等或低展示组中的出现次数对出现超过一次的独特克隆(基于VH和VL CDR3序列)进行表征。表中的数字表示出现的次数。

[0932] 使用下一代测序(NGS),通过深入分析抗体群体来证实基于表面呈现所选的细胞群的独特性。简单来说,扩增来自展示水平高、中和低的群体和来自输入群体的抗体的成对可变域基因,测序并且使用不对称的条形码<sup>125,126</sup>进行多路解编并且然后将输出的BAM文件

转化成FASTQ文件格式以能够用Geneious Biologics软件进行抗体序列分析。

[0933] 总共19792个成对VL和VH基因读段由这个分析产生。对所述群体进行多路解编之后,输入群体和展示水平高、中和低的群体由此分别产生2998、1516、6180和9098个CCS读段。对抗体序列进行标注以定位其构架和互补决定区(CDR)并且根据IMGT数据库指定重链和轻链生殖系<sup>127</sup>。不具有终止密码子的VH和VL CDR3序列已标注的克隆的这种分析的结果概述于表15中。通过使用初始抗体噬菌体展示文库<sup>8</sup>针对间皮素执行两轮抗体噬菌体展示选择而预选的抗体输入群体具有高度多样性,其中51%克隆是VH CDR3序列独有的并且88%克隆是VH和VL CDR3序列独有的。这种输入群体也具有抗体生殖系多样性,其中测序的每1000种克隆分别鉴定19、18和10VH、VL $\kappa$ 和VL $\lambda$ 生殖系。当根据哺乳动物高、中和低展示门控群体进行分析时,就CDR3序列和生殖系而言的这种多样性减少。展示水平低的群体的多样性明显减少,其中VH CDR3独特抗体的数目下降到小于4%并且所测序的每1000种克隆的VH、VL $\kappa$ 和VL $\lambda$ 生殖系的数目分别下降到4、2和3。与输入群体相比,抗体多样性的这种减少指示某些克隆富集根据展示水平门控抗体群体,从而增加群体的冗余。

[0934] 也检查VL $\kappa$ 和VL $\lambda$ 生殖系在不同群体中的比率。虽然起始输入群体的VL $\lambda$ 几乎2倍过量于VL $\kappa$ 抗体,但对于根据低展示水平门控群体而言,这种比率是反的,其VL $\kappa$ 抗体过量逾四倍。这些结果表明低展示组富集具有特定轻链生殖系序列的抗体的特定子集。

	输入	低展示	中等展示	高展示
测序总数 总数	1132	7378	3837	968
独特的 VH CDR3	575 (51%)	284 (3.8%)	547 (14%)	236 (24%)
独特的 VL CDR3	535 (47%)	196 (2.7%)	446 (11.6%)	185 (19%)
独特的 VH + VL CDR3	995 (88%)	1170 (15.9%)	2204 (57%)	637 (66%)
VH 生殖系总数	22	14	18	16
VL $\kappa$ 生殖系总数	20	19	19	17
VL $\lambda$ 生殖系总数	11	4	17	9
每 1000 个克隆的 VH 生殖系	19.4	1.9	4.7	16.5
每 1000 个克隆的 VL $\kappa$ 生殖系	17.7	2.6	5.0	17.6
每 1000 个克隆的 VL $\lambda$ 生殖系	9.7	0.5	4.4	9.3
VL $\kappa$ /VL $\lambda$ 比率	0.54	4.4	0.82	1.3

[0935] 表15.通过PacBio测序对根据哺乳动物抗间皮素展示所选的所产生克隆群体进行的序列分析。对FACS中根据低、中和高展示水平门控群体(图30)进行DNA序列分析并且与用于产生起始抗间皮素展示文库的输入抗体产生进行比较。生殖系以及VL CDR1、2和3以及VH CDR1和2遵循命名法,而VH CDR3是根据Kabat编号系统定义。

[0936] 为了对输入群体和基于低、中和高展示水平、利用用哺乳动物高等真核细胞展示所选的克隆进行较详细的生殖系分析,向抗体指定其最接近的匹配VH、VL $\kappa$ 和VL $\lambda$ 生殖系并且计算每组的生殖系出现频率。结果以表格式(表16、17和18)和直方图(图42、43和44)显

示。这表明在测序的群体中,展示水平低或高的组富集或缺少若干生殖系,表明基于抗体展示水平的哺乳动物展示选择能够选择性富集抗体生殖系。举例来说,与低展示组相比,在高展示组中,IGHV3-23和IGHV1-2生殖系分别富集105倍和9倍。IGHV3-53、IGHV3-21和IGHV4-34生殖系分别以2.7%、1.1%和0.9%的频率出现于高展示组,但在低展示组中完全缺乏。IGHV3-23生殖系被认为是表现良好的生殖系并且体现于若干已批准用于临床用途的治疗抗体,包括阿利库单抗、贝伐单抗、赛妥珠单抗(certolizumab)、达土木单抗(daratumumab)、地诺单抗(denosumab)、杜匹鲁单抗(dupilumab)、艾密次单抗(emicizumab)、奥克雷珠单抗(ocrelizumab)、兰尼单抗(ranibizumab)和思图昔单抗(siltuximab)。被认为表现良好的IGHV3-21和IGHV3-53生殖系已纳入人类合成抗体文库设计<sup>12</sup>。发现低展示组中富集了五种VH生殖系:IGHV1-69D、IGH3-30、IGHV5-51、IGHV1-58和IGH3-64D。此前已经观察到,IGH3-30生殖系当与某些轻链成对时倾向于自身相互作用,如通过动态光散射测量所测定<sup>12</sup>。

[0939] 对抗体轻链生殖系的出现频率的分析也表明与低展示组相比,某些生殖系在高展示组中富集。举例来说,IGKV1D-13和IGKV3-19分别以3.5%和6%频率出现于高展示组中,但在低展示组中完全不存在。与低展示组相比,IGKV1-12、IGKV1-17、IGKV2-28、IGKV2D-29和IGKV6-21生殖系在高展示组中分别富集11倍、22倍、6倍、11倍和132倍。IGKV2-28和相关生殖系IGKV2D-29被认为是表现良好的并且体现于若干已批准用于临床用途或正经历临床评估的抗体,包括坎妥珠单抗(cantuzumab)、杜匹鲁单抗(dupilumab)、鲁卡木单抗(lucatumumab)、莫加珠单抗(mogamulizumab)、奥比珠单抗(obinutuzumab)、司韦单抗(sevirimab)、替妥莫单抗(tenatumomab)和杂图昔单抗(zatuximab)。发现低展示组中富集四种VL $\kappa$ 生殖系:IGKV1-16、IGKV1-33、IGKV2-30和IGKV4-1。由于单一VH $\lambda$ 生殖系在输入群体中占主导(IGLV3-19),因此VH $\lambda$ 生殖系分析不太明显。尽管如此,高展示组中富集以下VH $\lambda$ 生殖系:IGLV1-47、IGLV2-11、IGLV2-14、IGLV2-23、IGLV2-8、IGLV3-10和IGLV6-57。已知V $\lambda$ 1-47和V $\lambda$ 2-14轻链表现良好并且已经作为支架纳入合成文库设计<sup>12,128</sup>。IGLV2-14是临幊上已批准的抗PCSK9抗体依伏库单抗(Evolocumab)中存在的轻链,所述抗体是表现良好的抗体<sup>2</sup>。

[0940] 这种分析能够将人类抗体生殖系“分组”成哺乳动物高展示组或低展示组。先前已表明高水平展示与表现良好的生物物理学特性(就蛋白质自身相互作用倾向低而言)相关(实例3)。因此,观察到某些生殖系在哺乳动物展示高展示水平组中富集是归因于其固有的生物物理学特性。因此,哺乳动物展示能够实现具有表现良好的生物物理学特性(就自身相互作用倾向低而言)的抗体生殖系的富集和鉴定。

VH	出现频率(%)				高/低比率	
	输入	低	中	高		
[0941]	IGHV1-18	1. 9	8. 1	1. 8	3. 5	0. 43
	IGHV1-2	5. 5	0. 42	6. 7	3. 9	9. 34
	IGHV1-3	3. 0	13	5. 6	4. 3	0. 34
	IGHV1-46	5. 8	5. 6	9. 7	13	2. 40
	IGHV1-58		1. 4	5. 9	0	0. 00
	IGHV1-69	9. 3	28	2. 1	12	0. 44
	IGHV1-69D	5. 4	5. 8	0	0. 83	0. 14
	IGHV2-70D	0. 18	0	0	0	-
	IGHV3-11	0. 18	0	0	0. 31	-
	IGHV3-20	61	17	54	47	2. 75
	IGHV3-21	0. 18	0	0	1. 1	-
	IGHV3-23	1. 1	0. 04	0. 57	4. 2	104
	IGHV3-30	1. 5	4. 5	3. 3	0. 21	0. 05
	IGHV3-33	0. 71	1. 1	2. 4	4. 2	3. 9
	IGHV3-43	0. 18	0. 11	0. 57	0	-
[0942]	IGHV3-48	0. 09	0	0	0	-
	IGHV3-53	0. 80	0	2. 19	2. 7	-
	IGHV3-64D	0	0. 84	0	0	-
	IGHV3-66	0	0	0. 03	0	-
	IGHV3-7	0. 09	0	0. 03	0	-
	IGHV3-72	0. 09	0	0	0	-
	IGHV4-34	0. 09	0	0. 16	0. 93	-
	IGHV5-51	2. 5	15	2. 50	1. 34	-
	IGHV6-1	0. 18	0	2. 32	0. 21	-
	IGHV7-81	0. 09	0	0. 08	0	0. 43

[0943] 表16. 对根据哺乳动物展示所选的抗间皮素克隆群体进行的可变重链(VH)生殖系分析。向抗体指定最接近的匹配人类VH生殖系序列并且显示了输入群体以及根据哺乳动物低、中和高展示门控群体中的出现频率。最后一栏显示根据高和低展示水平门控群体中的出现比率, 其中两个群体中均发现生殖系。

[0944]

VL κ	出现频率 (%)				高/低比率
	输入	低	中	高	
IGKV1-12	0.76	0.02	1.85	0.18	11
IGKV1-13	0.25	0	0.75	0	-
IGKV1-16	1.76	3.47	1.45	0.18	0.05
IGKV1-17	0.50	0.02	0.41	0.36	22
IGKV1-27	0.76	1.16	0.87	0.73	0.6
IGKV1-33	1.76	7.14	2.72	1.28	0.18
IGKV1-39	56.2	49.2	42.2	61.2	1.2
IGKV1-5	3.02	0.70	4.05	0.18	0.26
IGKV1-6	0	0.03	0	0	-
IGKV1-8	0	0	0	0.18	-
IGKV1-9	2.02	0.56	0.29	1.28	2.3
IGKV1D-13	1.76	0	4.05	3.46	-
IGKV1D-16	0.25	0.05	0.12	0	-
IGKV2-24	0.25	0.07	0	0	-
IGKV2-28	6.30	2.19	11.93	13.66	6.2
IGKV2-30	3.02	4.19	5.50	0	-
IGKV2D-29	0.25	0.05	0	0.55	11
IGKV3-11	0.50	0	0.12	0	-
IGKV3-15	6.30	4.39	10.19	3.64	0.83
IGKV3-19	0	0	0	6.01	-
IGKV3-20	1.26	0.40	4.00	1.09	2.7
IGKV3D-15	0	0.02	0.06	0	-

[0945]

IGKV4-1	12.85	26.36	8.63	3.83	0.15
IGKV6-21	0.25	0.02	0.87	2.19	131

[0946] 表17. 对根据哺乳动物展示所选的抗间皮素克隆群体进行的可变重链(VLκ)生殖系分析。向抗体指定最接近的匹配人类VLκ生殖系序列并且显示了输入群体以及根据哺乳动物低、中和高展示门控群体中的出现频率。最后一栏显示根据高和低展示水平门控群体中的出现比率, 其中两个群体中均发现生殖系。

[0947]

VL λ	出现频率(%)				高/低比率
	输入	低	中	高	
IGLV1-40	0. 14	0	0	0	-
IGLV1-44	0. 14	0	0. 33	0	-
IGLV1-47	0	0	0. 28	0. 72	-
IGLV1-50	0	0	0. 24	0	-
IGLV2-11	1. 09	0. 07	0. 24	2. 63	36
IGLV2-14	0. 41	0	0. 14	0. 48	-
IGLV2-18	0	0	0. 09	0	-
IGLV2-23	0	0	0. 85	0. 24	-
IGLV2-8	0. 14	0	0. 19	1. 67	-
IGLV3-1	1. 50	0. 29	0. 95	0	-
IGLV3-10	0. 14	0	0. 47	1. 43	-
IGLV3-19	94. 55	97. 64	90. 57	90. 21	0. 9
IGLV3-21	0	0	0. 09	0	-
IGLV3-25	1. 09	1. 99	2. 13	1. 19	0. 60
IGLV3-27	0	0	0. 24	0	-
IGLV3-9	0. 14	0	0. 71	0	-
IGLV5-45	0	0	0. 05	0	-
IGLV6-57	0. 68	0	2. 42	1. 43	-

[0948] 表18. 对根据哺乳动物展示所选的抗间皮素克隆群体进行的可变重链(VLλ)生殖系分析。向抗体指定最接近的匹配人类VLλ生殖系序列并且显示了输入群体以及根据哺乳动物低、中和高展示门控群体中的出现频率。最后一栏显示根据高和低展示水平门控群体中的出现比率, 其中两个群体中均发现生殖系。

克隆	克隆名称	展示组	VH-VL CDR3	重链生殖系	轻链生殖系	高展示 (%)	中等展示 (%)	低展示 (%)	输入 (%)	Δ AC-SINS (nm)	HPLC-SEC 滞留体积 (ml)
1	932_01_D10	H	DSRPPYYGMDV-MQALQTPPT	IGHV1-69	IGKV2-28	3.17	0.62	0.00	0.47	1	nd
2	932_01_C12	H	DGRGGFDY-QQFNSYHLLT	IGHV3-43	IGKV1-39	1.69	0.00	0.00	0.00	2.5	1.53
3	932_01_A03	H	GRSSVIDYGMDV-NSRDSSGNHV	IGHV3-20	IGLV3-19	1.19	0.40	0.00	0.00	2	1.54
4	N/A	H	DRAVAATHYYYYGMDV-QQS YGSPFT	IGHV3-30	IGKV1-39	1.09	0.20	0.00	0.00	nd	nd
5	N/A	H	GILLEKGAFDI-QQS YSTPQT	IGHV1-46	IGKV1-39	0.99	0.00	0.00	0.00	nd	nd
6	N/A	H	DSRPPYYGMDV-QQS YNSRPYT	IGHV1-69	IGKV2-28	0.79	0.07	0.00	0.00	nd	nd
7	932_01_A06	H	TSPYSGSYNN-SSYGGNYKYL	IGHV4-34	IGLV2-8	0.79	0.00	0.00	0.00	3	1.59
8	932_01_F02	H	GLSSTWAGGAFDI-NSRDSSGNHV	IGHV3-20	IGLV3-19	0.69	1.52	0.02	1.41	2	1.54
9	N/A	H	GAHSGYDSDFDY-QQYNSYPLT	IGHV1-18	IGKV1-39	0.69	0.00	0.02	0.00	nd	nd
10	932_01_B01	H	VSGSSNHAFDI-HQSSSFPLT	IGHV1-2	IGKV1-2	0.69	0.00	0.02	0.00	1	1.55
11	931_01_F01	M	DTSSRYAGGAFDI-NSRDSSGNHV	IGHV1-69D	IGLV1-50	0.30	0.62	0.00	0.71	6	1.58
12	932_01_C04	M	DTGSSARGGDFDY-NSRDSSGNHV	IGHV3-20	IGLV3-19	0.30	0.75	0.20	0.55	2	1.54
13	931_01_B10	M	DERYYGDMDV-MQQLQTPRT	IGHV1-69	IGKV2-28	0.10	0.47	0.00	0.31	nd	nd
14	930_01_A01	M	DYSSGWSGDAFDI-NSRDSSGNHV	IGHV1-3	IGKV1-39	0.10	0.40	0.39	0.08	nd	nd
15	931_01_B02	M	RGTRGYYYGMDV-QQS YNSRPYT	IGHV1-69D	IGKV1-39	0.10	0.65	3.63	2.36	23	1.61
16	N/A	M	TNNGFIDY-SQASHWPYT	IGHV6-1	IGKV2-30	0.00	1.12	0.00	0.08	nd	nd
17	N/A	M	GRGRRYFDL-QQS YSTPRT	IGHV1-2	IGKV1-39	0.00	0.55	0.00	0.00	nd	nd
18	931_01_E03	M	DGDGEELGAFDI-QQS YGSPFT	IGHV1-46	IGKV1-39	0.00	0.42	0.00	0.00	nd	nd
19	930_01_D09	L	GDNNYYFDY-QQYNSRPIT	IGHV1-69	IGKV4-1	0.00	0.00	1.88	0.00	5	1.55
20	930_01_A09	L	ETGEGRWELLGY-MQGTHWPRT	IGHV1-46	IGKV2-30	0.00	0.00	2.22	0.00	4	nd
21	N/A	L	LSHTAPLVDY-QQYSKWPLT	IGHV5-51	IGKV3-15	0.00	0.12	2.26	0.63	nd	nd
22	930_01_C02	L	AIA PRRYYGMDV-QQS YSTPRT	IGHV1-69	IGKV1-39	0.00	0.22	2.65	1.18	nd	nd
23	930_01_A12	L	AIA PRRYYGMDV-QQYNSYPLT	IGHV1-69	IGKV1-39	0.00	0.00	2.90	0.08	20	1.61
24	930_01_A08	L	DGYNSDY-QQYYSKPLT	IGHV1-39	IGKV1-33	0.00	0.00	3.70	0.00	1	1.55
25	930_01_C07	L	RRYNW DYDV DV-QQS YSTPRT	IGHV5-51	IGKV1-39	0.00	0.00	5.79	0.08	25	1.56
26	930_01_A05	L	DKPVGSSGWYPFDY-QQS YSTPYT	IGHV1-3	IGKV1-39	0.00	0.00	10.50	0.00	nd	nd
27	930_01_B02	L	QINWGYYFDY-MQALQTPPT	IGHV1-69	IGKV4-1	0.00	0.00	11.68	0.00	22	1.73

[0949] 表19. 根据HEK293细胞高、中和低展示水平、针对抗间皮素抗体进行的抗体克隆聚类分析。根据细胞表面展示水平混合根据哺乳动物细胞展示所选的抗间皮素抗体(图31)。利用PacBio NGS对展示水平高(H)、中(M)和低(L)的群体进行DNA测序并且鉴定每组中的前10种最丰裕克隆并且这些最丰裕克隆描绘于这个表中,其中VH和VH CDR3序列以通过虚线分隔的单字母氨基酸码显示。向此前通过对个别克隆进行桑格测序(Sanger sequencing)鉴定的抗体指定克隆名称,显示于第二栏中。显示了每种克隆在高、中和低展示组中的丰度百分比。如果抗体存在于超过一个展示水平组中,则细胞加阴影。所选克隆在AC-SINS分析<sup>40</sup>中表达、提纯且测试并且显示了波长漂移(Δ AC-SINS) (nm)。HPLC-SEC滞留体积显示于最后一栏中(nd:未进行)。

[0950] 进行聚类分析以产生四种抗间皮素抗体群体中的每一种的序列独特VH和VL CDR3克隆清单以及其出现频率。根据出现频率得到的前10种克隆显示于表19中,展示水平高、中和低的抗间皮素抗体如图31中所描绘进行FACS门控。为了比较,还显示了输入和替代群体中的克隆的出现频率。如此前以较有限的克隆集合所示,高水平展示组中的前7种最丰裕克隆与低展示组之间不存在重叠。高水平展示组中的7种最丰裕克隆在低水平展示组中完全不存在。低水平展示组中的9种最丰裕的克隆在高水平展示组中完全不存在。对于最丰裕的、中水平展示组来说,高水平展示组与低水平展示组存在一些重叠:中水平展示组中仅存在三种克隆,中和高水平展示组中仅存在两种克隆并且所有三个水平展示组中存在三种克隆。对于来源于与低水平展示组存在一些重叠的高水平展示组的克隆(例如克隆8、9和10)来说,这些克隆在高水平展示组中富集逾41倍。

[0951] 抗体根据其序列分组成高、中和低水平展示组表明基于真核细胞抗体展示水平的分离归因于所述抗体的特性,如根据其多肽序列所测定。如通过PacBio NGS分析所测定的抗体的出现频率与通过桑格测序对一组更有限克隆测序的结果非常吻合。此前已经利用

PacBio NGS、通过桑格测序鉴定出低、中和高水平展示组中的前10种最丰裕克隆中的9、8和6种。

[0953] 然后表达从高和低水平展示组中选出的抗体并且提纯并且通过AC-SINS<sup>41</sup>或HPLC-SEC确定其生物物理学特性。AC-SINS结果显示于表19的倒数第二栏中。来源于高或低展示组的克隆的AC-SINS平均波长漂移分别是1.9和15.2。与高水平展示组相比,从来源于低水平展示组的克隆观察到的显著较高AC-SINS平均波长漂移表明低水平展示组抗体比高水平展示组抗体更倾向于自身相互作用和聚集。所选抗体也通过HPLC-SEC检查。图45显示被表达且通过HPLC-SEC提纯和分析的四种克隆的实例:一种来自高展示组(932\_01\_A03)且三种来自低展示组(930\_01\_A12、930\_01\_B02和930\_01\_C12)。这表明从高水平展示组中分离出的抗体932\_01\_A03是96%单体。相比之下,抗体930\_01\_C12是82%单体和18%二聚体并且在柱上还显示一些迟滞,随后的洗脱图表明与柱存在一些非特异性相互作用。这种克隆还包括小片段(16kDa分子量),表明其在表达或提纯期间可能已片段化。来自低展示组的两种其它克隆(930\_01\_A12、930\_01\_B02)在柱上也展示延迟的滞留时间,表明抗体非特异性吸收到柱基质。HPLC-SEC滞留体积(表19最后一栏)是关于与柱基质发生相互作用的非特异性抗体的度量。在此,较大滞留体积能够引起较宽的HPLC-SEC峰且指示与HPLC-SEC柱基质发生非特异性相互作用。高展示组克隆的HPLC-SEC平均滞留体积是1.55ml,而低展示组克隆的HPLC-SEC平均滞留体积是1.6ml,从而证明低展示组克隆比高展示组克隆更倾向于与非特异性HPLC-SEC柱基质发生相互作用。在尝试通过超滤浓缩之后,低水平展示组中的一种克隆(930\_01\_C07)沉淀。低水平展示组中的第二克隆的表达和提纯失败。来源于高水平展示组的克隆无一者证明在其表达、提纯和后续分析期间存在任何问题。这包括在PBS(pH7.4)中通过超滤能够浓缩到大于12mg/ml。

[0954] 因此,这个实例已举例说明能利用哺乳动物高等真核细胞展示来分离复杂混合物内存在的抗体,包括具有不同生殖系序列的抗体,此分离是基于其自身相互作用和交叉相互作用特性。

[0955] 实例11. 基于细胞表面上的快速积聚来选择可开发性克隆

[0956] 在上述实例中,我们已经注意到稳定的高等真核细胞群的表面上的组成型抗体展示水平与所述抗体的生物物理学特性之间的关系。在此,我们还注意到在转染之后的24小时观察到展示水平的差异。在实例5、图10中,我们在转染后的第1天(1dpt)时间点观察到博可珠单抗与亲本抗体5A10i之间的展示水平差异(图10,左图)。所展示抗体的积聚差异的这种初始速率与稳定细胞系在平衡状态下的展示水平差异(图10,右图,21dpt)和抗体的生物物理学特性相关。实例5中显示与亲本抗体5A10-i相比,博可珠单抗具有高自身相互作用和交叉相互作用特性,如AC-SINS分析和HPLC-SEC中所测量。与亲本抗体5A10-i相比,在转染后24小时短暂阶段和稳定细胞系阶段、转染后21天(图10),博可珠单抗以更低水平展示。因此,初始短暂抗体细胞展示水平预测在药物选择若干天之后建立的稳定细胞系的抗体细胞展示水平。

[0957] 因此,我们已经表明高等真核细胞表面的表面上的抗体展示速率与抗体的生物物理学特性之间的关系。抗体的初始瞬时展示与其生物物理学特性之间的经证实的关系适用于在更早阶段(与产生稳定细胞系相比)从抗体群体或文库中富集生物物理学特征优良的抗体。其也适用于通过个别单克隆转染来筛选个别抗体集合。细胞表面展示抗体的初始速

率的差异也可以是适用的,其中通过诱导型启动子调控抗体表达并且在诱导后开始基于细胞表面展示水平对群体进行选择。

[0958] 实例12. 基于CHO细胞文库中的表面呈现水平,根据IgG的可开发性进行选择

[0959] IgG抗体能够展示于CHO细胞表面上,其展示水平与个别抗体的生物物理学特性(包括其自身相互作用和聚集的倾向)相关。在这个实例中,我们表明在执行两种替代的核酸酶介导抗体基因整合方法之后,抗体展示于CHO细胞表面上,并且我们证明展示水平与所展示抗体的生物物理学特性相关。

[0960] 基因靶向载体被设计成且用于将抗体基因表达盒整合到CHO AAVS基因的内含子1中,所述CHO AAVS基因与实例1中所述的人类AAVS基因组基因座直系同源<sup>93</sup>。TALEN核酸酶靶向CHO AAVS基因的基因座的内含子1,在这个位置建立双股断裂。为了比较,还用并行方法测试CRISPR/Cas9核酸酶。

[0961] 使用TALEN定向和CRISPR/Cas9定向整合方法成功地实现CHO细胞上的抗体展示。使用TALEN核酸酶对或CRISPR 1设计来实现最佳展示水平。替代的CRISPR设计(CRISPR 2和CRISPR 3)也是成功的,但是观察到稍微较低的抗体展示水平。

[0962] 图48.

[0963] 我们进一步证明CHO细胞上所展示的抗体的细胞表面水平与所展示抗体的生物物理学特征相关,与HEK293细胞上所发现的结果一致。将所选抗体克隆到pINT17-BSD-CHO靶向载体:MEDI-1912、MEDI-1912-STT、博可珠单抗、884\_01\_G01。

[0964] MEDI-1912(抗NGF)就其自身聚集的倾向而言具有不良的生物物理学特性<sup>7</sup>。与溶解度增强的子系克隆MEDI1012-STT相比,观察到MEDI-1912在CHO细胞上的细胞表面展示水平降低。

[0965] 884\_01\_G01是博可珠单抗衍生物并且通过诱变和哺乳动物展示选择、基于展示水平来鉴定并且保持结合到PCSK9(实例5)。就AC-SINS得到的低分数和HPLC-SEC得到的单体峰而言,884\_01\_G01在此前表明是“表现良好的”。与改善的子系克隆884\_01\_G01相比,还观察到博可珠单抗的CHO细胞表面展示水平降低。

[0966] 图49.

[0967] 所述实例因此证实CHO细胞用于展示抗体的效用和基于其细胞表面展示水平来区分具有不同生物物理学特性的抗体的效用。具有不同生物物理学特性的抗体能够通过流式细胞术很好地分离,从而能够根据CHO细胞展示从复杂混合物内分离出生物物理学特性较好的抗体。

[0968] 材料与方法

[0969] 从悬浮液调适的CHO细胞系基因组DNA,对编码CHO AAVS左和右同源臂的DNA进行PCR扩增,使用引物对3165/3166和3169/3170(表20)从CHO细胞中分离以分别产生大小为1.2kb和1.35kb的PCR产物。这些PCR产物然后用作模板以产生CHO AAVS左和右同源臂,同时敲除后续抗体基因克隆所需的多个限制位点。建立现有NcoI、NheI、NsiI和DraIII位点发生突变的CHO AAVS左同源臂,其通过使用上述CHO AAVS左臂模板、使用PCR引物对3195/3196、3197/3198、3199/3200进行PCR扩增以建立大小分别为241bp、576bp和142bp的三种PCR产物。然后使用引物3199和3196(表20)组装片段以产生大小为893bp的产物。进行额外的组装PCR反应,其例外之处在于使用引物3123替代3196以得到稍微较短的AAVS-左同源臂

(880bp)。产生这种稍微较短的左同源臂的原因是为了避免靶向载体内的CRISPR 2和3设计识别位点(图46)。使用AsiSI和NsI限制酶消化所组装的AAVS左同源臂并且克隆到用AsiSI和NsI预切割的pINT17-BSD载体中。为了对CHO AAVS同源臂进行PCR扩增并且敲除BcI限制位点,使用引物对3201/3202和3203/3204进行两次单独的PCR反应以建立大小分别为707bp和208bp的PCR产物,然后使用引物3201和3204进行PCR组装以产生大小900bp的PCR产物。组装的AAVS右同源臂用BstZ171和SbfI限制酶消化并且克隆到用BstZ171和SbfI预切割的pINT17-BSD载体中并且然后克隆到BstZ171和SbfI切割的pINT17-BSD切割载体中,所述载体已经具有CHO AAVS左同源臂。然后建立CHO靶向载体pINT17-BSD-CHO,其与pINT17-BSD相同(图1),例外之处在于人类AAVS同源臂已被CHO AAVS同源臂置换。

名称	引物序列	描述
3165	GGTGCTCGACTCCACCAA	CHO-AAVS-左-F
3166	GATGGAAGTTGCCATGAAAGA	CHO-AAVS-左-R
3169	TCTTGTATTGCCGGGATCCTTC	CHO-AAVS-右-F
3170	TAACCTCCAGCCCTACCTACTC	CHO-AAVS-右-R
3195	CTCCACCTACCACCTCATGGACTATATTG	CHO-AAVS-左-F1-exNcoI
3196	TTTTTT <u>ATGCAT</u> CTTATGCCAGCTTTGGATGACGG	CHO-AAVS-左-R1-exNsI
3197	CTCCTCTGAG <u>TCTAG</u> CCAGGCC	CHO-AAVS-左-F2-exNheI
3198	CAAATATA <u>GTCCATG</u> AGGTGGTAGGTGGAG	CHO-AAVS-左-R2-exNcoI
3199	TTTTTT <u>GCGATCGCGATGGCTTACATCCC</u> GTGCCTT	CHO-AAVS-左-F3+AsiS1-exDraIII
3200	GGCCTGGCTAG <u>ACTCAGAGGAG</u>	CHO-AAVS-左-R3-exNheI
[0970] 3201	TATATT <u>GTATA</u> CGGCCGCCTGTCAGGGACAAGATTAGTCACAG	CHO-AAVS-右-F4-exBstZ171+AscI
3202	GACTTTGGTGATAATGTGAGCAGC	CHO-AAVS-右-R4-ex-BcI
3203	GCTGCTCACATTATCACCAAAGTC	CHO-AAVS-右-F5-ex-BcI
3204	TATATT <u>CCTGCAGGCTCCTG</u> CAAAGGCCTGAAGAG	CHO-AAVS-右-R5+SbfI
3213	TTTTTT <u>ATGCAT</u> CTTGATGACGGGGAGATAAAAGCAT	CHO-AAVS-左-R1+NsI-CRISPR2+3
3201	TATATT <u>GTATA</u> CGGCCGCCTGTCAGGGACAAGATTAGTCACAG	CHO-AAVS-右-F4+BstZ171+AscI
3202	GACTTTGGTGATAATGTGAGCAGC	CHO-AAVS-右-R4-ex-BcI
3203	GCTGCTCACATTATCACCAAAGTC	CHO-AAVS-右-F5-ex-BcI
3204	TATATT <u>CCTGCAGGCTCCTG</u> CAAAGGCCTGAAGAG	CHO-AAVS-右-R5+SbfI
3213	TTTTTT <u>ATGCAT</u> CTTGATGACGGGGAGATAAAAGCAT	CHO-AAVS-左
	C	-R1+NsI-CRISPR2+3
[0971] 3220	GGAATCATGGGAAATAGGCCCT	CRISPR1-R-载体
3221	CGCTCACAA <u>TTCCACACAACAT</u>	CRISPR1-F-载体
3222	AGGGCCTAT <u>TTCCC</u> ATGATTCC	CRISPR-seqF
3223	ATGTTGTGTGGAATTGTGAGCG	CRISPR-seqR

[0972] 表20.能够扩增CHO AAVS左和右同源臂以及敲除限制位点的引物序列和CRISPR向导RNA引物。限制位点加有下划线并且敲除NheI、NcoI、BcI和DraIII限制位点的突变以粗体显示。

[0973] TALEN对设计成识别CHO AAVS基因座,从而识别TAL靶序列:TCCCCGTCATCCAAAAGC和TCTGCTGTGACTAATCTT,如图46所示。为了与TALEN引起的位点特异性DNA裂解进行比较,我们还测试了替代核酸酶(导向核酸的核酸酶CRISPR/Cas9)的性能。设计靶向CHO AAVS基因座的三种CRISPR/Cas9向导RNA,并且识别序列显示于图46中。

[0974] 克隆到靶向载体中的CHO AAVS同源臂的核酸序列显示于图47中。设计编码U6 RNA聚合酶启动子向导RNA和tracrRNA的合成基因块并且用引物3222和3223进行PCR扩增(表21)。使用引物3220/3221对CRISPR/Cas9载体(A21177, Geneart)进行PCR扩增以产生9375bp产物。然后使用NEB Builder(新英格兰生物实验室(New England Biolabs))、使用三种CRISPR基因块(表20)组装这种产物,以产生编码图46中所描绘的三种向导RNA的三种CRISPR/Cas9载体。

CHO CRISPR	基因块序列
[0975]	<p>1</p> <pre> <u>AGGGCCTATTC</u>CCATGATT<u>C</u>TTCATATTGCATATA<u>C</u>GATACAG<u>A</u>GGCTGTTAGAGAGA <u>T</u>AATTAGAATT<u>A</u>TTGACTGTAA<u>A</u>CAACAA<u>A</u>AGATATTAGTACAA<u>A</u>ACGTGACGTAGAA <u>A</u>GTAA<u>A</u>TTTCTTGGTAGTTGCAG<u>T</u>TTAAAATTATGTTAA<u>A</u>ATGGACTATCATA <u>T</u>GCTTACCGTA<u>A</u>CTTGAA<u>A</u>GTATT<u>C</u>GATT<u>C</u>TTGG<u>C</u>TTATATATCTTG<u>G</u>AA<u>A</u>GG<u>A</u> <u>G</u>AA<u>A</u>CC<u>G</u><b>ATCCAAAAGCTGGCATTGTCGTTT</b>AG<u>G</u>C<u>A</u>AA<u>A</u>TAG<u>C</u>A<u>G</u>TT<u>A</u>AA<u>A</u> <u>A</u>GG<u>C</u>TAG<u>T</u>CC<u>G</u>TT<u>T</u>CA<u>A</u>CT<u>T</u>G<u>A</u>AA<u>A</u>AG<u>T</u>GG<u>C</u>ACC<u>G</u>AG<u>T</u>CG<u>G</u>T<u>G</u>CT<u>T</u>TT<u>T</u>CT<u>A</u>GT<u>T</u>AC <u>C</u>GT<u>G</u>AC<u>C</u>CT<u>T</u>AG<u>C</u>TAG<u>A</u>G<u>G</u>CT<u>T</u>GG<u>C</u>TA<u>A</u>T<u>C</u>AT<u>GG</u>TC<u>A</u>T<u>AG</u>CT<u>G</u>TT<u>C</u>CT<u>T</u>GT<u>G</u>AA<u>A</u>TT <u>G</u>TT<u>A</u>TC<u>CG</u>CT<u>C</u>ACA<u>A</u>TT<u>CC</u>AC<u>AC</u>AC<u>AT</u> </pre>
	<p>2</p> <pre> <u>AGGGCCTATTC</u>CCATGATT<u>C</u>TTCATATTGCATATA<u>C</u>GATACAG<u>A</u>GGCTGTTAGAGAGA <u>T</u>AATTAGAATT<u>A</u>TTGACTGTAA<u>A</u>CAACAA<u>A</u>AGATATTAGTACAA<u>A</u>ACGTGACGTAGAA </pre>

	<pre> AGTAATAATTCTGGTAGTTGCAGTTAAAATTATGTTAAAATGGACTATCATA TGCTTACCGTAACTGAAAGTATTCGATTCTGGCTTATATCTGTGGAAAGGAC <u>GAAACACCGTCTCCCCGTCACTCCAAAAGCG</u> TTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGCTTTCTAGTATAC CGTCGACCTCTAGCTAGAGCTGGCGTAATCATGGCATAGCTGTTCCGTGAAATT GTTATCCGCTACAATTCCACACACAT </pre>
[0976]	<pre> AGGCCTATTCCCAGATTCTCATATTGCATATACGATACAAGGCTTAGAGAGA TAATTAGAATTAAATTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAA <u>AGTAATAATTCTGGTAGTTGCAGTTAAAATTATGTTAAAATGGACTATCATA</u> TGCTTACCGTAACTGAAAGTATTCGATTCTGGCTTATATCTGTGGAAAGGAC <u>GAAACACCGATGCCAGCTTGGATGACGG</u> TTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGCTTTCTAGTATAC CGTCGACCTCTAGCTAGAGCTGGCGTAATCATGGCATAGCTGTTCCGTGAAATT GTTATCCGCTACAATTCCACACACAT </pre>
3	

[0977] 表21. 编码CHO CRISPR序列的基因块序列。CHO CRISPR向导和tracrRNA序列分别以粗体和斜体显示。U6 RNA聚合酶启动子加下划线显示。

[0978] 在编码CHO AAVS TAL-1F和CHO AAVS TAL-1R TALEN对(M3770, 赛默飞世尔科技)或CRISPR/Cas9和CHO AAVS特异性向导RNA的质体不存在或存在下, 使用其中893bp或880bp AAVS左同源臂(分别命名为V1或V2)具有尼沃单抗抗体重链和轻链基因的靶向载体pINT17-BSD-CHO转染Freestyle CHO-S细胞(赛默飞世尔科技)并且经历杀稻瘟菌素药物选择(7.5μg/ml)。转染方法如实例2中针对TALEN转染所述。在CRISPR/Cas9转染中, 使用10:1的CRISPR/Cas9、pINT17-BSD-CHO质体DNA比率。细胞在转染后(dpt)的第14天用抗Fc PE染色。荧光强度(抗Fc)相对于细胞计数作图。

[0979] 将两对抗体的重链和轻链基因克隆到CHO靶向载体pINT17-BSD-CHO中, 所述靶向载体具有893bp AAVS左同源臂(图47): 抗NGF抗体MEDI-1912或MED-1912-STT<sup>7</sup>或抗PCSK9博可珠单抗或博可珠单抗衍生物克隆884\_01\_G01(参见实例5)。所得靶向载体在编码CHO AAVS TAL-1F和CHO AAVS TAL-1R TALEN对(M3770, 赛默飞世尔科技)的质体存在下用于转染Freestyle CHO-S细胞(赛默飞世尔科技)并且经历杀稻瘟菌素药物选择(7.5μg/ml)。转染后第15天的细胞用抗Fc-PE染色, 如实例2中所述。转染后第15天的展示抗体的CHO细胞通过抗Fc-PE MACS提纯并且用抗Fc PE染色。荧光强度(抗Fc)相对于CHO细胞计数作图。

[0980] 实例13. 利用哺乳动物展示来建立“可开发性增强的”群体用于随后的结合选择

[0981] 实例10证明针对抗原预选的克隆群体(在这个实例中, 得自噬菌体展示选择)能够基于不同呈现水平进一步拆分, 继而与生物物理学特性关联。如在别处所述, 对多反应性探针的结合还可以用于鉴定和去除具有多反应性特性的克隆。从而能够排除存在问题的克隆, 其可能以其它方式被选用于进一步表征或开发。

[0982] 在实例10中, 初始基于结合特性(在这个实例中使用噬菌体展示)选择抗体群体,

随后基于生物物理学特性进行后续选择。也可以颠倒选择次序并且基于高等真核细胞的呈现水平或对多反应性探针的结合来产生已经根据最优生物物理学特性选择的克隆群体。举例来说,结合剂展示文库可以在哺乳动物细胞或其它高等真核细胞中制备并且使用结合到结合剂的恒定区的试剂、根据呈现水平选择,以便所有结合剂都能够独立于其序列而被同等标记。举例来说,在结合剂是IgG抗体的情况下,可以使细胞与可检测试剂接触,所述可检测试剂结合到IgG Fc区,例如经标记的抗IgG抗体。然后能够选择呈现水平高于峰值或中值的群体部分,例如通过流式细胞术来选择。举例来说,使用本文所述的方法,在一轮或多轮分选中能够基于呈现水平选择前5%、10%或25%的克隆。这将建立“可开发性增强”的结合剂群体,其能够用于随后基于对不同抗原的结合进行选择。

[0983] 使用实例10的方案,选择 $4 \times 10^{10}$ 个克隆的大型噬菌体展示文库以产生结合剂的亚群,然后并入哺乳动物展示文库中以便基于呈现水平进行选择。颠倒次序(即,根据高等真核细胞中的生物物理学特性预选,随后选择结合剂)的一个潜在缺点是,建立哺乳动物细胞的大型文库比建立用于体外系统(例如核糖体展示或细菌系统(例如噬菌体展示))的文库更费力且代价更高。因此,如果使用例如 $10^7$ 种克隆的初始哺乳动物展示文库并且基于呈现水平而选择前10%,则用于随后根据抗原结合进行选择的文库潜在多样性将减少,例如减少到 $10^6$ 种克隆,使得选择高亲和力结合剂的可能性减小。一种替代策略是预选哺乳动物展示文库中的组分,例如在抗体的情况下,分开选择最优VH和最优VL组分,然后将其重组以从组合多样性中受益。

[0984] 因此,在一个分支中,可以建立哺乳动物展示文库,其中呈IgG或scFv形式的抗体含有单次或有限次选择的VL和较大多样性的VH链以选择最优VH基因。为了选择VH基因以拯救搭配物链的不良生物物理学特性,可以随机选择或基于不良生物物理学特性选择单一或有限搭配物VL基因。或者,可以选择生物物理学特性良好的VL基因以鉴定和去除损害呈现水平的那些VH基因。使用与上述实例相同的数目,可以建立具有VH多样性的 $10^7$ 种克隆的文库并且基于呈现水平而选择的前10%VH基因潜在地产生 $10^6$ 种经选择的VH克隆。在一个并行分支中,可以将单个或有限数目个VH基因与VL基因的多样性组合以选择最优VL基因。将所选不同VH和VL基因结合在一起的组合多样性显著高于个别选择的链的数目。因此,每个分支中选择前 $10^6$ 个VH和前 $10^6$ 个VL基因建立了 $10^{12}$ 种变异体的潜在组合多样性。所选VH和VL域可以在允许根据靶向结合进行选择的任何展示系统内呈现,包括噬菌体展示、核糖体展示、酵母展示和高等真核细胞上的展示。以此方式,使用本发明根据最优生物物理学特性进行的选择可以用于在其它展示系统内产生“可开发性增强”的文库。举例来说,预处置以产生具有最优生物物理学特性的克隆的噬菌体展示文库。

[0985] 在上述方案的另一实例中,根据最优生物物理学特性进行的选择可以定向于个别结构域内的区域。举例来说,VH域包括3种不同的互补决定区(CDR1、CDR2和CDR3),并且生物物理学可靠性可以由个别CDR或CDR组合内的序列促成。举例来说,可以回收非免疫抗体谱系中的VH基因的CDR1和2区域,与单个或有限数目个VH CDR3区域和VL搭配物链组合,且具有多样性的所得抗体群体聚焦于根据最优生物物理学特性所选的CDR1和2内。然后能够将所选CDR1和2区域与CDR3区段和VL区段的多样性组合,所述区段可以未选定或可能已根据最优生物物理学特性以类似方式选择。潜在的组合多样性因这种“改组”方案而增加。上述实例描述了根据最优生物物理学特性针对CDR 1和CDR2区段进行的选择,但所述相同方案

如何能够结合单独或组合的不同CDR区域使用将是显而易见的。在一些情况下,根据最优生物物理学特性选择的输入文库可以是来自非免疫来源或来自多样化初始支架的初始文库。或者,为了潜在地改善生物物理学特性以及亲和力,输入文库可以是基于初始输入克隆或克隆选择的多样化文库。

[0986] 在各种情况下,上述方案将利用所属领域的技术人员已知的标准分子生物学技术。另外,W02015/166272(Iontas有限公司)还阐述了文库(包括链改组文库)的构建和使用方法,其内容并入本文中供参考。将VH或VL域内的不同区域组合将受益于使用最优PCR引物扩增VH和VL域内的个别区域。抗体VH和VL基因的生殖系序列容易获得,例如允许设计此类引物的IMGT数据库<sup>132</sup>。另外,W02017/118761(Iontas有限公司)中的实例描述了引入多样性以及将不同CDR区域组合以重构完整VH:VL组合的方法。

[0987] 下面的实例用于说明。将多样性引入初始PD1结合抗体克隆的VH CDR3和VL CDR3中并且所得群体根据哺乳动物细胞上的最优呈现进行选择。可以采用相同方案将多样性引入其它初始构架和其它CDR,包括生殖系编码基因的CDR 1和2。举例来说,表16鉴定了针对哺乳动物细胞上的高水平呈现所选的群体中频繁出现的生殖系基因,例如IGHV1-2、IGHV3-23和IGKV6-21,并且这些生殖系基因可以选作初始构架用于进一步的多样化和可开发性选择。类似地,可以鉴定出除抗体之外的结合域亚区域以便根据最优生物物理学特性进行单独选择。

[0988] 举例来说,根据噬菌体展示来鉴定阻断PD-1与PD-L1相互作用的抗PD1抗体(337\_1\_C08)(图50a)。337\_1\_C08对PD-1的亲和力( $K_D$ )是74nM。这用于建立多样化聚焦于VH CDR3和VL CDR3的未诱变文库。所属领域的技术人员已知用于多样化的多种方法。举例来说,多样化可能涉及通过饱和诱变来建立诱变文库,其中使用随机化密码子,例如NNS(编码32个密码子内的所有20种氨基酸)。一种替代方案是保持序列信息,同时探究周围序列空间。为了实现这种方案,寡核苷酸设计为VH和VL的CDR3中的8个氨基酸的连续片段,以便包括每种可能2聚体和1聚体变异体(使用所有氨基酸(除了半胱氨酸之外))(图50a)。

[0989] 因此,文库设计成每个CDR3区域中保持原始8个氨基酸中的至少6个。这可以通过合成靶向VH CDR3或VL CDR3中的每一个的9216种寡核苷酸来实现。通过TWIST Biosciences合成VH和VL的基因块,其中每种2聚体氨基酸变异都涵盖于8个氨基酸窗口内(9216种变异体)。通过高通量测序(Twist Bioscience)来证实每一组中存在所有9216种寡核苷酸。使用引物对扩增VH基因:正向引物:5' -CTTCTCTCCACAGGCGCCATGGCCGAAGTGCAG CC-3' 和反向引物:5' -TTTTTCTCGAGACGGTGACCAGGGTTC-3'。使用引物对扩增VL基因:正向引物:5' -TTTTTGCTAGCTCCTATGAGCTGACTC-3' 和反向引物:5' -GTCACGCTGGTGCGGCCGCGGG CTGACCTAG-3'。使用以下引物对,利用pINT17-BSD载体对编码恒定轻链(CL)和CMV启动子的“填充”片段进行PCR扩增:正向引物:5' -GGCCGCACCAAGCGTGAC-3' 和反向引物:5' -GGCGCCTGTGGAGAGAAAG-3'。首先通过“模拟”PCR而不使用扩增引物来组装三个基因片段,以确保部分组装和随后扩增不引入偏差。通过将上述三种PCR产物按照等摩尔比率(各90nM)混合且使用KOD热启动预混液(西格玛,71842)、根据制造商说明书、使用60°C的粘接温度和68°C的延长温度(45秒)执行PCR反应。随后使用引物对扩增组装产物:正向引物:5' -TTTTTGCTAGCTCCTATGAGCTGACTC-3' 和反向引物:5' -TTTTTCTCGAGACGGTGACCAGGGTTC-3'。组装的抗PD-1 VH和VL CDR3文库用NheI和XhoI限制酶消化并且克隆到pINT17-BSD靶向载体

中，并且通过对E.coli 10G SUPREME电感受态细胞(Lucigen目录号60081-1)进行电穿孔，建立 $1.1 \times 10^8$ 的文库大小，如通过对涂铺转化混合物的稀释液的琼脂板上的个别卡那霉素抗性群落进行计数所测定。制备转染质量的质体DNA且通过Maxcyte电穿孔用于与TALE AAVS左和右臂核酸酶共转染HEK293悬浮细胞( $1.35 \times 10^9$ 个细胞)以实现单个拷贝抗体基因整合。如通过对转染后的稀释板中的杀稻瘟菌素抗性群落形成单位(CFU)进行计数所测定，基因靶向效率是0.8%，以产生10.8百万的哺乳动物展示文库大小。文库在杀稻瘟菌素选择下繁殖7天且然后通过抗Fc MACS选择以去除不表达IgG的克隆。

[0990] 如实例10中所述，通过流式细胞术、利用荧光标记的抗Fc抗体、基于表达水平来分离Fc阳性细胞群(图50b)。将分离出的群体培养若干天并且针对Fc表达进行再分析。图50b显示用于表达的具有不同峰值的不同亚群已经产生。这包括呈现水平较差的群体以及呈现水平相当于亲本的群体。如实例10中所述，能够从这些细胞中回收抗体基因以产生生物物理学特性最优的亚群。在这个实例中，细胞仅基于生物物理学特性分选，但也可以基于与抗原结合一起或依序执行选择。

[0991] 实例14. 在哺乳动物细胞表面上展示双特异性抗体

[0992] 双特异性抗体或替代形式(评述于Spiess等人, 2015<sup>129</sup>)已能够达成此前的单一特异性抗体或蛋白质所不能达成的新治疗作用机理(MOA)。双特异性治疗剂的用途实例包括作为癌症治疗剂使T细胞的细胞毒性活性重定向、能够穿越血脑屏障、同时阻断两种信号传导路径以及抗体的组织特异性递送或活性<sup>130</sup>。双特异性形式就其倾向于自身或交叉相互作用特性而言，与传统的单特异性抗体一样都面临着相同的可开发性障碍。由于每一个结合区能具有补偿其自身相互作用或交叉相互作用特性或与之混配的不同特性，因此筛选双特异性部分的最终形式是一个优势。当筛选双特异性的个别结合“臂”时，这些特性可能不明显，原因是亲合力能够使针对脱靶分子(例如肝素硫酸盐)的亲和力增加若干数量级。由于总体表面特性(例如疏水性)可以改变，因此双特异性分子相较于其个别结合区而言，自身相互作用特性也可以不同。双特异性抗体(bsAb)形式不仅能以多种不同形式存在<sup>129</sup>，而且能够广泛分成：IgG样bsAb，例如κ/λ体<sup>131</sup>、共同轻链、臼包杵<sup>132</sup>、电荷对和互换单抗形式<sup>130</sup>；基于片段的bsAb，例如BiTE形式；附接的IgG，例如DVD-IgG，或经工程改造以在其恒定域中具有额外结合区的抗体，例如Fcab或mAb2形式<sup>133</sup>。然而，发现Fcab的CH3域的结构构架的改变使其热稳定性减小<sup>134</sup>且需要对Fcab分子进行额外的工程改造来增强其可开发性<sup>134</sup>。在这个实例中，我们证明哺乳动物高等真核展示能够应用于双特异性抗体的展示，这通过双特异性抗体艾密次单抗展示于HEK293细胞表面上和表明这能够结合抗原FIXa和FX。

[0993] 艾密次单抗是一种双特异性抗体，其利用重链“臼包杵”技术、使用此前所述的共同轻链<sup>132</sup>产生，其为了治疗血友病而开发且充当因子VIII模拟物<sup>135</sup>。艾密次单抗的一个臂特异性针对因子IXa(FIXa)且第二个臂特异性针对因子X(FX)。通过将艾密次单抗抗FIXa重链、抗FX重链和共同轻链基因(表22)克隆到诱导型靶向载体pINT17-Tet中来构建三顺反子靶向载体。将包括下游聚腺苷酸化(pA)位点的共同轻链基因克隆到pINT17-Tet的BgIII和NheI限制位点。将包括PDGFR跨膜域的抗FIXa重链基因克隆于pINT17-Tet的NcoI与HindIII限制位点之间。将包括信号肽、PDGFR跨膜域和SV40 pA的抗FX重链基因克隆于pINT17-Tet的EcoRI与BstZ171限制位点之间。最终载体(pINT17-Bi-CMV-艾密次单抗)含有介于AAVS同源臂之间的编码序列。图51. 还将艾密次单抗抗FIXa和VL基因克隆到pINT17-BSD靶向载体

内以能够展示单特异性抗FIXa IgG抗体。

[0994] 为了证明HEK293细胞表面展示双特异性抗体艾密次单抗,使用pINT17-Bi-CMV-艾密次单抗或pINT17-BSD-抗FIXa在如上文所述的编码AAVS TALEN的质体存在下转染HEK293细胞。转染后24小时,使用抗Fc-APC分析细胞上的抗体展示(图52)。这表明经单特异性抗FIXa臂的IgG形式转染或经双特异性抗体艾密次单抗转染的HEK293细胞在细胞表面上具有可检测到的抗体表达。然而,与标准单特异性抗体形式相比,具有“白包杆”重链的双特异性艾密次单抗的表达减少。设想通过建立Fc变异体文库、随后根据哺乳动物展示的高水平表达进行选择,可以选择能够实现更高效异源重链配对和细胞展示水平增加的Fc变异体。以此方式能够选择新抗体CH2或CH3变异体或CH2与CH3变异体的组合,其不仅在异源重链配对以建立双特异性抗体方面更高效,而且具有优良的可开发特性,包括自身聚集的倾向低。还能够根据与其它分子的交叉相互作用倾向低来筛选Fc变异体文库。

[0995] 还通过流式细胞术证明所展示的双特异性艾密次单抗结合其靶抗原的能力。FIXa和FX(Complement技术有限公司)以化学方式生物素化(EZ-连接碘基-NHS-生物素,赛默飞世尔科学)。原因是艾密次单抗对FIXa和FX的亲和力相对较低( $K_d$ 在微摩尔浓度范围内<sup>136</sup>)。细胞染色之前,使抗原在100nM的抗原复合物浓度下与四聚体抗生蛋白链菌素-PE预偶联以增加结合亲合力。展示双特异性艾密次单抗的HEK293细胞显示结合FIXa和FX,而不结合到未偶联的抗生蛋白链菌素-PE(图52a)。这证明双特异性抗体在HEK293细胞表面上的功能展示。在HEK293细胞表面上展示的单特异性抗FIXa臂的IgG形式还结合到FIXa(图52b)。抗FIXa抗体对FX也显示一些结合,然而结合水平低于艾密次单抗,这可能归因于共同轻链或抗FIXa重链与FX交叉反应。

[0996] 实例6b描述了肝素硫酸盐与HEK293细胞表面上所展示的抗IL12抗体布瑞金单抗和优特金单抗的差异结合。布瑞金单抗在其可变域内具有正电荷补丁,所述正电荷补丁可能促成其结合FcRn上的负电荷补丁<sup>22</sup>。布瑞金单抗上的正电荷补丁也可能是其与负电荷肝素硫酸盐发生交叉相互作用的原因。抗体结合到肝素硫酸盐能够增加体内非特异性清除,从而缩短半衰期<sup>137,138</sup>。因此,抗体或治疗蛋白质(包括双特异性分子)结合到肝素是非期望的特性。实例6b证明能够区分结合肝素硫酸盐的抗体并且因此能够通过哺乳动物高等真核展示选择来分离和排除肝素硫酸盐结合蛋白质。根据这个实例中呈现的数据,基于结合肝素硫酸盐的能力将一系列抗FIXa/抗FX双特异性抗体区分开来的证明是能够设想的。

[0997] 在艾密次单抗开发过程中,发现了前体人源化抗体,命名为hBS106<sup>136</sup>。发现这种分子在小鼠中具有不良药物动力学,与人类IgG4相比,其清除快且体内半衰期短。这种体内快速清除可归因于VH和VL抗FIXa臂上的正电荷补丁。对正电荷补丁有贡献的共同轻链上的氨基酸包括VL CDR1内的K24、R27和R31以及VL CDR2内的R53和R54以及FW3内的R61(图53)。对正电荷补丁有贡献的VH氨基酸包括R60和R95。对正电荷补丁有贡献的赖氨酸和精氨酸残基是FIXa结合必不可少的互补位残基。因此,引入带负电的氨基酸残基(谷氨酸或天冬氨酸)会破坏正电荷补丁并且还降低抗体等电点(pI)。带负电的氨基酸变化引入共同轻链中使得小鼠清除速率减小8倍且使最大体内血浆浓度(Cmax)增加近6倍。图53显示艾密次单抗的共同VL链与一系列前体VL的比对(US2016/0222129),与最终艾密次单抗VL相比,所述前体带负电荷的氨基酸少三个到一个。艾密次单抗亲本抗体相对于艾密次单抗命名。举例来说,相对于艾密次单抗,克隆:E30Y\_E55Y\_D93S在位置30、55和93分别具有酪氨酸、酪氨酸和

丝氨酸残基以替代谷氨酸、谷氨酸和天冬氨酸。

[0998] 可以将艾密次单抗前体VL抗FIXa基因(表23)克隆到pINT17-Bi-CMV-艾密次单抗靶向载体中,并且用于与如上文所述的人类AAVS TALEN对共转染HEK293细胞以能够实现核酸酶介导的抗体基因整合。转染后的第14天之后,HEK293细胞可以用肝素硫酸盐(FITC标记)和抗Fc-PE染色,如实例B中所述。流式细胞术分析将表明,与抗FIXa艾密次单抗和E30Y VL克隆相比,哺乳动物细胞展示的抗FIXa艾密次单抗克隆E30Y\_K54R\_E55Y\_D93S和E30Y\_E55Y\_D93S对肝素硫酸盐展示增强的结合。E30Y\_K54R\_E55Y\_D93S和E30Y\_E55Y\_D93S克隆具有最完整的正电荷补丁并且这些克隆会引起可检测的肝素硫酸盐结合到展示这些抗体的细胞。这个实例已经表明如何能利用哺乳动物高等真核展示来区分结合到肝素硫酸盐的能力有差异的具有点突变的克隆。这将能够排除复杂文库内的结合肝素硫酸盐的双特异性克隆,且因此去除其可变域内具有正电荷补丁的克隆,所述正电荷补丁可能造成体内不良药物动力学。

[0999] 这个实例已经说明通过哺乳动物展示来筛选双特异性抗体或替代双特异性形式的潜力。与单一靶抗体发现项目相比,将双特异性发现操作时需要筛选的变异体数目乘以若干个数量级。举例来说,在艾密次单抗发现操作中,为了鉴定主要的初始分子,使200种抗FIXa抗体与200种抗FX抗体互换以建立40,000个双特异性分子,然后通过费力的板基筛选进行筛选。主要的双特异性分子然后必须经历若干次迭代,包括人源化、亲和力和特异性优化以及可开发性增强,其关于通过破坏可变域正电荷补丁来减少与肝素硫酸盐发生的交叉相互作用<sup>136</sup>。通过哺乳动物展示对数百万个双特异性抗体克隆进行多维FACS的能力,包括对靶标的亲和力、特异性、展示水平(根据自聚集的倾向低来筛选)和交叉相互作用,能够以较大的筛选深度对较大多目个克隆执行更快、更高效的双特异性抗体筛选过程。

[1000]

基因	基因序列
艾密次 单 抗 VL-CL	TAATAAGCTAGCAGAGGA GACATCCAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGC GTGGGAGACAGAGTGACCATCACATGCAAGGCCAGCCGGAACATCGAGAGACAGCTGGCC TGGTATCAGCAGAACGCTGGACAGGCTCCTGAGCTGCTGATCTATCAGGCCAGCAGAAAA GAAAGCGCGTGCCCGATAGATTAGCAGCGGCAGCAGATA CGGCACCGACTTCACCCCTGACA ATATCCAGCCTCCAGCCTGAGGATATCGCCACCTACTACTGCCAGCAGTACAGCGACCC CCACTGACATTGGCGGAGGCACCAAGGTGGAAATCAAGCGGACAGCGGCCGCCCCTAGC GTGTTCATCTTCCACCTAGCGACGAGCAGCTGAAGTCTGGCACAGCCTCTGCGTGTGC CTGCTGAACAACCTCTACCCAGAGAAGCCAAGGTGCAGTGGAAAGGTGGACAACGCCCTC CAGAGCGCAATAGCCAAGAGAGCGTGACCGAGCAGGACAGCAAGGACTTACCTACAGC CTGAGCAGCACACTGACCCTGAGCAAGGCCGACTACGAGAAGCACAAGTGTACGCCCTGC GAAGTGACCCACCAGGGCCTTCTAGCCCTGTGACCAAGAGCTTCAACCAGGGCGAATGT

[1001]

	TAATAATCTAGAGCCTCGACTGTGCCCTAGTTGCCAGCCATCTGTTGTTGCCCTCC CCCCTGCCCTCCTTGACCCTGGAAGGTGCCACTCCCCTGCTCTTCTAATAAAATGAG GAAATTGCATCGCATTGCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGTGGGCAG GACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGA <u>AAGATCTTAATAA</u>
艾密次 单抗抗 FIXa VH-CH1 -CH2-C H3	TTATTAG <u>CCATGCC</u> CAGGTGCA <u>GCTGGTGA</u> ATCTGGCGGAGGGACTGGTT <u>CAGCCTGGC</u> <u>GGATCTCTGAGACTGTCTTGCC</u> GCCAGCGGCTTCACCTTCAGCTACTACGATATCCAG TGGGT <u>CCGACAGGCC</u> CTGGCAAAGGACTTGAATGGGTGTC <u>CCAGCATCAGCC</u> CCCTCTGGC CAGTCCACCTACTACCGCGAGAAGTGAAGGGCAGATT <u>CACCATCAGCC</u> GGGACAACAGC AAGAACACCC <u>CTGTACCTGCAG</u> ATGAACAGCCTGAGAG <u>CCGAGGACACCGCC</u> GTACTAC TGCGCCAGAAGAACCGGAGAGAGTACGGCGGAGGCTGGTACTTTGATTACTGGGCCAG GGCACCC <u>CTGGTCA</u> CAG <u>CTCGAGCGC</u> CTACAAAGGGCCCCAGCGTTTCCACTGGCT CCCTGTAGCAGAACGACCAGCGAATCTACAGCCGCTCTGGGCTGCCTGGTCAAGGACTAC TTCCCTGAGCCTGTGACC <u>GTCTGGA</u> ACTCTGGCGCTCTGACATCTGGCGTGCACACC TTTCCAGCCGTGCTGCAAAGCAGCGGCTGTACAGTCTGAGCAGCGTGTGACAGTGCCT AGCAGCTCTCTGGGACCCAGACCTACACCTGA <u>ATGTGGACC</u> ACAAGCCTAGCAACACC AAGGTGGACAAGCGCGTGGAA <u>CTAAGTACGGCC</u> CTCCTGTCC <u>CTCCATGT</u> CC <u>TGAC</u> CT GAGTTCTCGGCGGACCC <u>CTCCGTG</u> TT <u>CCCTG</u> TT <u>CCAAAGC</u> CTAAGGACACC <u>CTGATG</u> ATCTCCAGAACACCC <u>GAAGT</u> GAC <u>CTGC</u> GTGGTGG <u>ACGTT</u> ACAAGAGGACCC <u>GAG</u> GTGCAGTTAATTGGTACGTGGACGGCGTGGAA <u>GTG</u> CAC <u>ACGCCA</u> AG <u>ACCA</u> AGCCTAGA GAGGAACAGTACAACAGCAC <u>CTACAGAGTGGT</u> GTCC <u>GTG</u> CTGACAGT <u>GTG</u> CAC <u>CCAGG</u> TGGCTGAACGGCAAAGAGTACAAGTGA <u>AGGT</u> GT <u>CCAACA</u> AGGGCC <u>CTGCCA</u> AG <u>CGC</u> ATC GAGAAAACC <u>ATCAGCA</u> AG <u>GCCA</u> AGGG <u>CC</u> <u>AGC</u> <u>CTAGGG</u> <u>ACCC</u> <u>CAGG</u> <u>TT</u> <u>ACAC</u> <u>ACT</u> <u>GCCT</u> CCAAG <u>CCAGAA</u> AG <u>AGAT</u> GAC <u>CAAA</u> AG <u>ACC</u> <u>AGGT</u> <u>GTCC</u> <u>CTGAC</u> <u>CTG</u> <u>GAAGGG</u> <u>CTTC</u> TACCC <u>TTCCG</u> <u>ATATCG</u> <u>CCGT</u> <u>GGAA</u> <u>ATGGGAGAGCA</u> <u>ATGGCC</u> <u>AGCC</u> <u>AGAGA</u> <u>AAACT</u> <u>ACAAG</u> ACCAC <u>ACCTCTGTG</u> <u>CTGG</u> <u>ACAGCG</u> <u>ACGG</u> <u>GCTC</u> <u>ATTCTT</u> <u>CCGT</u> <u>TACAGCA</u> <u>AGCTG</u> <u>ACCG</u> <u>GT</u> GACAAGAGCAG <u>ATGG</u> <u>CAAGAGGG</u> <u>CAAC</u> <u>GTG</u> <u>TT</u> <u>CAG</u> <u>CTG</u> <u>AC</u> <u>CG</u> <u>GT</u> <u>ATG</u> <u>GCAC</u> <u>GAGG</u> <u>CC</u> <u>GT</u> ACAACAGATA <u>ACCC</u> <u>AGA</u> <u>AGTCC</u> <u>CTGT</u> <u>CTGAG</u> <u>CCCC</u> <u>GAAC</u> <u>AAA</u> <u>ACT</u> <u>CAT</u> <u>CTCAG</u> <u>AA</u> GAGGAT <u>CTG</u> <u>AATG</u> <u>CTGTGG</u> <u>CCAGG</u> <u>ACAC</u> <u>GCAGGAGG</u> <u>TCAT</u> <u>GTGG</u> <u>TGCC</u> <u>ACACT</u> <u>CC</u> <u>TTG</u> <u>CCCTTAAAGTGGTGTG</u> <u>ATCTCAGC</u> <u>ATCCTGG</u> <u>CCCTGG</u> <u>GTG</u> <u>CTAC</u> <u>CCAT</u> <u>GT</u> <u>AC</u> <u>CC</u> <u>AGG</u> <u>CTTATCATCCTCATCATG</u> <u>CTTGG</u> <u>CAGAAGAAGCCACGTTAGTAA</u> <u>AAGCTTT</u> <u>ATTA</u>
艾密次 单抗抗 FX VH-CH1 -CH2-C H3	TTATTAG <u>AAATTCAAC</u> <u>ATGG</u> <u>ACTGG</u> <u>AC</u> <u>CTGG</u> <u>AGGG</u> <u>CTTCTG</u> <u>CTG</u> <u>CTGG</u> <u>GCTG</u> <u>TA</u> <u>GCTCCA</u> GGTAAAGGG <u>CCAA</u> <u>CTGG</u> <u>TTCC</u> <u>AGGG</u> <u>CTG</u> <u>AGGA</u> <u>AGGG</u> <u>ATTT</u> <u>TTCC</u> <u>AG</u> <u>TT</u> <u>AGAGG</u> <u>ACTG</u> CATT <u>CTCTACTG</u> <u>TG</u> <u>TCCTCT</u> <u>CCCG</u> <u>CAGG</u> <u>TG</u> <u>CTC</u> <u>ACTCC</u> <u>CAGG</u> <u>TT</u> <u>CAG</u> <u>CTGG</u> <u>TG</u> <u>CA</u> <u>GT</u> <u>CC</u> <u>AGCG</u> <u>ACTG</u> <u>AAAAA</u> <u>ACCTGG</u> <u>CC</u> <u>CTGG</u> <u>GA</u> <u>AGGT</u> <u>GT</u> <u>CC</u> <u>TG</u> <u>CA</u> <u>AGG</u> <u>CT</u> <u>CTGG</u> <u>CT</u> <u>ACAC</u> <u>CTTTACCG</u> <u>ACA</u> <u>ACAC</u> <u>ATGG</u> <u>ACTGG</u> <u>GT</u> <u>CC</u> <u>ACAGG</u> <u>CC</u> <u>CTGG</u> <u>ACA</u> <u>AGG</u> <u>ACTG</u> <u>AGTGG</u> <u>AT</u> <u>GGGC</u> <u>GAC</u> <u>ATCA</u> <u>ACAC</u> <u>CCAGA</u> <u>AGCG</u> <u>CC</u> <u>GGC</u> <u>AGCAT</u> <u>CT</u> <u>ACA</u> <u>ACGA</u> <u>AGAGT</u> <u>CC</u> <u>AGG</u> <u>ACAGAGT</u> <u>CAT</u> <u>CATG</u> <u>ACCG</u> <u>GTG</u> <u>ACA</u> <u>AGAG</u> <u>GCAC</u> <u>CC</u> <u>GCT</u> <u>AC</u> <u>ATGG</u> <u>AA</u> <u>CTG</u> <u>AG</u> <u>GC</u> <u>AG</u> <u>CTG</u> <u>GAGA</u> CGAGGACACC <u>GCC</u> <u>AC</u> <u>CT</u> <u>ATG</u> <u>CG</u> <u>CC</u> <u>AGA</u> <u>AGA</u> <u>AGAG</u> <u>CTAC</u> <u>GG</u> <u>G</u> <u>CT</u> <u>AC</u> <u>CT</u> <u>GG</u> <u>AC</u> GTGGGG <u>CGAGGG</u> <u>AA</u> <u>CTG</u> <u>GT</u> <u>AC</u> <u>AG</u> <u>GT</u> <u>CT</u> <u>AG</u> <u>CG</u> <u>CC</u> <u>AG</u> <u>CA</u> <u>AA</u> <u>AGGG</u> <u>CC</u> <u>CT</u> <u>AG</u> <u>CG</u> <u>TT</u> CCC <u>ACTGG</u> <u>CTCC</u> <u>GT</u> <u>AG</u> <u>CAG</u> <u>CC</u> <u>GT</u> <u>AC</u> <u>AG</u> <u>CC</u> <u>GT</u> <u>CT</u> <u>GGG</u> <u>GT</u> <u>GC</u> <u>CT</u> <u>GT</u> GAAGGACT <u>ACTT</u> <u>CC</u> <u>GTGAG</u> <u>CC</u> <u>GTGAC</u> <u>CGT</u> <u>AG</u> <u>CTG</u> <u>GA</u> <u>AC</u> <u>AGCG</u> <u>GGAGC</u> <u>ACTG</u> <u>ACA</u> <u>AGCG</u> CGTG <u>CAC</u> <u>AC</u> <u>TT</u> <u>CC</u> <u>AGCC</u> <u>GT</u> <u>CTG</u> <u>CAA</u> <u>AGC</u> <u>AGCG</u> <u>GGC</u> <u>CTG</u> <u>ACT</u> <u>CT</u> <u>GTGAG</u> <u>CG</u> <u>CGT</u> <u>GT</u>

[1002]

[1002]	<pre> GACAGTGCCTAGCAGCTCTGGCACCCAGACCTACACCTGTAATGTGGACCACAAGCC TAGCAACACCAAGGTGGACAAGCGCGTGGAACTAAGTACGGCCCTCCTGTCCTCCATG TCCTGCTCCAGAGTTCTGGCGGACCCCTCCGTGTTCTGTTCTCAAAGCCTAAGGA CACCTGATGATCTCCAGAACACCCGAAGTGACCTGCGTGGTGGACGTTACAAGA GGACCCCGAGGTGCGAGTTCAATTGGTACGTGGACGGCGTGGAACTGACAAACGCCAAGAC CAAGCCTAGAGAGGAACAGTACAACAGCACCTACAGAGTGGTGTCCGTGCTGACAGTGCT GCACCAAGGATTGGCTGAACGGCAAAGAGTACAAGTGCAAGGTGTCACAAGGGCCTGCC AAGCAGCATCGAGAAAACCATCAGCAAGGCCAAGGGCAGCCTAGGGAACCCCAGGTTA CACACTGCCTCCAAGCCAAGAGGAATGACCAAGAACCCAGGTGTCCTGACCTGCCCTGGT CAAGGGCTTCTACCCCTCCGATATGCCGTGGAATGGGAGAGCAATGCCAGCCAGAGAA CAACTACAAGACCACACCTCCTGTGCTGGACAGCGACGGCTATTCTCCTGTACAGCAA GCTGACTGTGGATAAGAGCCGGTGGCAAGAGGGCAACGTGTTAGCTGTAGCGTGATGCA CGAGGCCCTGCACAACCACTACACCCAAAGAGAGCCTGCTCTGAGCCCTGAACAAAAACT CATCTCAGAAGAGGATCTG<b>AATGCTGTGGGCCAGGACACGCAGGAGGT</b>CATCGTGGTGCC <b>ACACTCCTTGCCCTTAAGGTGGTGGT</b>GATCTCAGCCATCCTGGCCCTGGTGGTGCTCAC <b>CATCATCTCCCTTATCATCCTCATCATGCTTGGCAGAAGAAGGCCACGTTAGTAAC</b>TAG TCGACATCCAGACATGATAAGATACTTGATGAGTTGGACAAACCACAACAGATAATGCA GTGAAAAAAATGCTTATTGTGAAATTGTGATGCTATTGCTTATTGTAACCATATT AAGCTGCAATAAACAAAGTTAACAAACAATTGCATTCAATTGCTTATGTTCAGGTTAGGG GGAGGTGTGGAGGTTTTAAAGCAAGTAAACACAAATTGCTTACAAATGTGGTATGGCTGATTA TGATCCTGCAAGCCTCGTCGTCCTGGCCGGACCACGCTATCTGTGCAAGGTCCCCGGCC CGGACGCGCGCTCCATGAGCAGAGCGCCCGCCGAGGCGAAGACTCGGGCGCGCCCT GCCCGTCCCACCAGGTCAACAGGCGTAACCGGCCTCTTCATCGGGAATGCGCGCGACCT TCAGCATCGCCGGCATGTCCTGGCGACGGGAAGTATGTATACTTATTA </pre>
--------	---

[1003] 表22. 艾密次单抗重链和轻链基因。可变和PDGFR跨膜域编码区分别以斜体和粗体突出显示。限制位点加有下划线。

[1004]

基因	基因序列
艾密次 单抗抗 FIXa VH	<pre> TTATTAGCCATGGCCCAGGTGCAGCTGGTGAATCTGGCGGAGGA<u>CTGGTT</u>CAGCCTGGC GGATCTCTGAGACTGTCTTGCCGCCAGCGGCTTCACCTCAGCTACTACGATATCCAG TGGGTCCGACAGGCCCTGGCAAAGGACTTGAATGGGTGTCCAGCATCAGCCCTCTGGC CAGTCCACCTACTACCGCGAGAAGTGAAGGGCAGATTACCATCAGCCGGACAACAGC AAGAACACCCCTGTACCTGCAGATGAACAGCCTGAGAGCCGAGGACACCGCCGTGTACTAC TGCGCCAGAAGAACCGGCAGAGAGTACGGCGGAGGCTGGTACTTGATTACTGGGCCAG GGCACCCCTGGTACAGTCTCGAGTTATTA </pre>
艾密次 单抗 VL	<pre> TTATTAG<u>CTAGCGACATCC</u>CAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGCGTGGGA GACAGAGTGACCATCACATGCAAGGCCAGCGGAACATCGAGAGACAGCTGGCTGGTAT CAGCAGAACGCCCTGGACAGGCTCCTGAGCTGCTGATCTACGGCCAGCAGAAAAGAAC GGCGTGCCCGATAGATTCAAGCGGCAGCAGATACGGCACCGACTTCACCCCTGACAATATCC AGCCTCCAGCCTGAGGATATGCCACCTACTACTGCCCAGCAGTACAGCGACCCCTCCACTG ACATTGGCGGAGGCACCAAGGTGGAA<u>ATCAAGCGGACAGCGGCCG</u>CTTATTA </pre>

[1005]	艾密次 单抗 VL E30Y	TTATTAG <u>CTAGCGACATCCAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGCGTGGGA</u> GACAGAGTGACCACATGCAAGGCCAGCCGGAACAT <u>CTATAGACAGCTGGCTGGTAT</u> CAGCAGAAAGCCTGGACAGGCTCCTGAGCTGCTGATCTATCAGGCCAGCAGAAA <u>TATAGC</u> GGCGTGCCCCGATAGATT <u>CAGCGGCAGCAGATA</u> CGGCACCGACTTCACCCGTACAATATCC AGCCTCCAGCCTGAGGATATGCCACCTACTACTGCCAGCAGTACAGCGACCCCTCCACTG ACATTGGCGGAGGCACCAAGGTGAAATCAAGCGGACAG <u>CGGCCGCTTATTA</u>
	艾密次 单抗 VL E30Y E55Y	TTATTAG <u>CTAGCGACATCCAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGCGTGGGA</u> GACAGAGTGACCACATGCAAGGCCAGCCGGAACAT <u>CTATAGACAGCTGGCTGGTAT</u> CAGCAGAAAGCCTGGACAGGCTCCTGAGCTGCTGATCTATCAGGCCAGCAGAAA <u>AGAACG</u> GGCGTGCCCCGATAGATT <u>CAGCGGCAGCAGATA</u> CGGCACCGACTTCACCCGTACAATATCC AGCCTCCAGCCTGAGGATATGCCACCTACTACTGCCAGCAGTACAGC <u>AGCCCTCCACTG</u> ACATTGGCGGAGGCACCAAGGTGAAATCAAGCGGACAG <u>CGGCCGCTTATTA</u>
	艾密次 单抗 VL E30Y E55Y D93S	TTATTAG <u>CTAGCGACATCCAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGCGTGGGA</u> GACAGAGTGACCACATGCAAGGCCAGCCGGAACAT <u>CTATAGACAGCTGGCTGGTAT</u> CAGCAGAAAGCCTGGACAGGCTCCTGAGCTGCTGATCTATCAGGCCAGCAGAAA <u>TATAGC</u> GGCGTGCCCCGATAGATT <u>CAGCGGCAGCAGATA</u> CGGCACCGACTTCACCCGTACAATATCC AGCCTCCAGCCTGAGGATATGCCACCTACTACTGCCAGCAGTACAGC <u>AGCCCTCCACTG</u> ACATTGGCGGAGGCACCAAGGTGAAATCAAGCGGACAG <u>CGGCCGCTTATTA</u>
	艾密次 单抗 VL E30Y K54R E55Y D93S	TTATTAG <u>CTAGCGACATCCAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGCGTGGGA</u> GACAGAGTGACCACATGCAAGGCCAGCCGGAACAT <u>CTATAGACAGCTGGCTGGTAT</u> CAGCAGAAAGCCTGGACAGGCTCCTGAGCTGCTGATCTATCAGGCCAGCAGA <u>AGATATAGC</u> GGCGTGCCCCGATAGATT <u>CAGCGGCAGCAGATA</u> CGGCACCGACTTCACCCGTACAATATCC AGCCTCCAGCCTGAGGATATGCCACCTACTACTGCCAGCAGTACAGC <u>AGCCCTCCACTG</u> ACATTGGCGGAGGCACCAAGGTGAAATCAAGCGGACAG <u>CGGCCGCTTATTA</u>

[1006] 表23. 抗FIXa艾密次单抗和艾密次单抗亲本抗体的VL和VH合成基因块DNA序列。艾密次单抗亲本抗体是根据相对于艾密次单抗的单字母氨基酸取代进行命名。举例来说，亲本抗体艾密次单抗VL E30Y指示艾密次单抗谷氨酸在位置30处被酪氨酸置换。相对于最终艾密次单抗VL的变化以粗体突出显示。侧翼的限制位点加有下划线。

[1007] 实例15. 基于表面呈现水平、根据可开发性选择打结抗体

[1008] 实例14证明HEK293细胞表面上能够展示“臼包杵”双特异性抗体艾密次单抗。还能够通过使多肽与抗体重链或轻链融合或对重链或轻链进行工程改造以赋予新颖结合特异性来建立双特异性分子<sup>133</sup>。哺乳动物细胞表面上展示双特异性分子的能力能够根据对两个靶标的结合进行筛选以及选择生物物理学特性良好的双特异性分子，如上文所述。打结抗体是新颖的抗体融合体形式，其中富半胱氨酸肽（打结素）并入抗体域的外周CDR环中（WO2017/118761）并且VH或VL域可用于结合相同抗原或不同抗原上的第二抗原决定基。这个实例的目的是证明哺乳动物展示技术能够区分和分离具有不同特性（关于其自身相互作用和多反应性特性）的打结抗体。

[1009] 打结抗体专利WO2017/118761描述了通过将胰蛋白酶结合打结素EETI-II插入抗体的VL CDR2位置来产生两种胰蛋白酶结合打结抗体（KB\_A07和KB\_A12）。这个实例中测试的打结抗体是KB\_A12\_ProTx-III\_2M（下文称为KB\_A12\_ProTx-III）和KB\_A12\_HsTx1。通过用离子通道阻断打结素或毒素肽ProTx-III\_2M（PCT/EP2018/068855, 2018年7月11日提交）和HsTx1（PCT/EP2018/068856, 2018年7月11日提交）置换KB\_A12打结抗体的VL CDR2位置的

EETI-II打结素来建立这些打结抗体。KB\_A12 EETI-II、KB\_A12 ProTx-III和KB\_A12 HsTxI的VL序列显示于表24中。

打结抗体	打结抗体 VL 序列
[1010]	QSVLTQPPSVSEAPRQRVTITCSGSSNIGNNAVNWYQQLPGKAPKLLIYAAGR <b>CPRI</b> <b>LMRCKQDSDCLAGCVCGPNGFCGANSGVSDRFSAAKSGTSASLAINGLRSEDEADYYC</b> AAWDDSLNGLYVFVGTGKLTVLG
	QSVLTQPPSVSEAPRQRVTITCSGSSNIGNNAVNWYQQLPGKAPKLLIYAAGR <b>GCLK</b> <b>FGWKCNPRNDKCCSGLKCGSNHNWCKWHIGANSGVSDRFSAAKSGTSASLAINGLRSE</b> DEADYYCAAWDDSLNGLYVFVGTGKLTVLG
	QSVLTQPPSVSEAPRQRVTITCSGSSNIGNNAVNWYQQLPGKAPKLLIYAAGR <b>ASCR</b> <b>TPKD CADPCR KET GCPY GKCMNRKCKCNRCANSGVSDRFSAAKSGTSASLAINGLRSE</b> DEADYYCAAWDDSLNGLYVFVGTGKLTVLG

[1011] 表24. 打结抗体VL序列。VL CDR2中插有接头氨基酸的打结素以粗体突出显示。

[1012] 将编码打结抗体KB\_A12 EETI-II、ProTx-III和HsTxI的基因克隆到靶向载体pINT17-BSD中并且用于通过核酸酶介导的基因整合到HEK293细胞中来建立稳定细胞系,如上文所述。图54a表明KB\_A12 EETI-II在HEK293细胞表面上的展示水平高于KB\_A12 ProTx-III或KB\_A12 HsTxI。HEK293细胞表面上以相对较低水平展示的KB\_A12 ProTx-III或KB\_A12 HsTxI与细胞计数相对于Fc表达的流式细胞术直方图(图54a)是等效的。打结抗体是通过Expi293细胞的瞬时转染来表达,通过蛋白质A亲和色谱法提纯且通过HPLC-SEC分析。图54b表明KB\_A12 EETI-II显示了单体峰,其滞留时间和体积与表现良好的抗HER2抗体曲妥珠单抗相当(图54c)。相比之下,KB\_A12 ProTx-III显示了二聚体和多聚体形成的证据,如根据较早的滞留时间所说明(图54d)。KB\_A12 ProTx-III的洗脱体积大于曲妥珠单抗和KB\_A12 EETI-II是KB\_A12 ProTx-III异源多聚体和聚集体形成的证据。在AC-SINS分析(表25)<sup>39</sup>中,与KB\_A12 EETI-II相比,KB\_A12 ProTx-III和KB\_A12 HsTxI还显示了自身相互作用增加的倾向,其中KB\_A12 ProTx-III和KB\_A12 HsTxI显示的AC-SINS波长漂移比KB\_A12 EETI-II长。因此,如此前利用IgG展示所发现,高等真核细胞展示打结抗体的水平与其发生自身相互作用和聚集倾向的生物物理学特性之间有关。打结抗体是双特异性抗体的实例且因此,所展示分子的高等真核生物细胞展示水平与生物物理学特性之间的这种关系可能转移到替代的双特异性模式<sup>129</sup>。

打结抗体	AC-SINS分数
KB_A12 EETI-II	3
KB_A12ProTx III	10
KB_A12HsTx I	20

[1014] 表25. 经提纯的打结抗体的AC-分数<sup>39</sup>。

[1015] 为了改善这些打结抗体的生物物理学特性,通过靶向诱发打结素序列中的疏水性或带正电残基突变来建立三种打结抗体哺乳动物文库。靶向诱发每种打结素突变的疏水性或带正电残基以粗体加下划线突出显示(参见下文)。使用编码VNS密码子(氨基酸序列中以X表示)或NSG密码子(氨基酸序列中以Z表示)的引物使这些残基突变。VNS密码子(V=A/C/G,N=A/G/C/T且S=G/C)通过23种密码子组合来编码16种氨基酸(排除半胱氨酸、酪氨酸、色氨酸、苯丙氨酸和终止密码子),而NSG密码子(N=A/G/C/T且S=G/C)通过8种密码子组合

来编码7种氨基酸(精氨酸、色氨酸、甘氨酸、苏氨酸、丝氨酸、丙氨酸和脯氨酸)。NSG密码子用于其中野生型色氨酸残基可能涉及对离子通道的结合接点的位置。

[1016] 野生型ProTx-III打结素序列,其中靶向诱变的残基呈粗体加下划线:

[1017] **GCLKFGWKCNPRNDKCCSGLKCGSNHNWCKWHI**

[1018] KB\_A12 ProTx-III A集合文库中的ProTx-III序列:

[1019] **GCXXXXZKCNPRNDKCCSGLKCGSNHNWCKWHI**

[1020] KB\_A12 ProTx-III B集合文库中的ProTx-III序列:

[1021] **GCXXXXZKCNPRNDKCCSGLXCGSNHNWCKZXX**

[1022] 野生型HsTx1打结素序列,其中靶向诱变的残基呈粗体加下划线:

[1023] ASCRTPKDCADPCRKETGCPYGKCMNR**KCKCNRC**

[1024] KB\_A12 HsTx-I文库中的HsTx1序列:

[1025] ASCRTPKDCADPCXXETGCPYGKCXNR**XCKCNRC**

[1026] 通过对打结抗体VL基因进行两片段组装PCR来产生文库。使用表26中所述的引物和模板扩增每个文库中的个别片段。使用引物pINT BM40前导序列正向引物和pINT CλNot反向引物组装与每一个文库对应的扩增片段。所有引物序列提供于表27中。使用NheI和NotI限制酶将组装的PCR片段消化并且接入编码D1A12或D12 VH(用于打结抗体构建体的重链,描述于WO2017/118761中)的pINT17-杀稻瘟菌素载体中。使用MinElute PCR提纯试剂盒(Qiagen,目录号28004)提纯接合产物。通过电穿孔将3x2.5μl的经提纯的接合混合物送入3x50μl的大肠杆菌细胞(E.cloni 10G Supreme, Lucigen, 目录号60080-2)。简单来说,使用0.1cm光析管对50μl细胞进行脉冲,使用2ml回收培养基进行回收并且在37°C、250rpm下生长1小时,并且通过稀释接种来估测文库大小。KB\_A12 ProTx-III A集合、KB\_A12 ProTx-III B集合和KB\_A12 HsTx1文库所得的文库大小分别是9x10<sup>7</sup>、7x10<sup>7</sup>和2x10<sup>7</sup>。为了建立哺乳动物展示文库,使用Macherey Nagel Midi制备型试剂盒(Macherey Nagel, 目录号740410.10)、依循制造商说明书,由这些文库储备液制备转染质量的DNA。

文库	片段1正向引物	片段1反向引物	片段2正向引物	片段2反向引物
[1027]	KB_A12 HsTx1 pINT BM40 前导 序列正向引物	HsTx1 第 1 部分 反向引物	HsTx1 第 2 部分正 向引物	pINT C λ Not 反向引物
	KB_A12 ProTx-I II A 集合 pINT BM40 前导 序列正向引物	ProTx-III 第 1 部分反向引物	ProTx-III 第 2 部 分正向引物	pINT C λ Not 反向引物
	KB_A12 ProTx-I II B 集合 pINT BM40 前导 序列正向引物	ProTx-III B 集 合第 1 部分反向 引物	ProTx-III B 集合 第 2 部分正向引 物	pINT C λ Not 反向引物

[1028] 表26. 用于扩增文库的片段1和片段2的引物。

引物名称	引物序列(5' -3')
[1029]	HsTx1 第 1 部分 反向引物 ACAGGGGTCCGCGCAGTCTTAGGAGTTCG
	HsTx1 第 2 部分 正向引物 CCTAAAGACTGCGCGACCCCTGTVNSVNSGAGACTGGATGTCCATACGGTAA GTGCVNSAATAGAVNSTGCAAATGTAACCGATGCGCAAACAGT
	ProTx-III 第 1 部分反向引物 CATCCCCTTCCCGCTGCGTAAATGAGAAG
	ProTx-III 第 2 部分正向引物 CTCATTACGCAGCGGAAGGGATGCVNSVNSVNSNSGAAATGCAACCC AAGAACGATAAA
	ProTx-III B 集 GAGTCCTGAGCAGCATTATCGTTCTGGTTGCATTCNSNBSNBSNBSN
[1030]	合第 1 部分反向 引物 BGCATCCCCCTTCCCGCTGCGTAAATGAG
	ProTx-III B 集 合第 2 部分正向 引物 CCAAGAACGATAATGCTGCTCAGGACTCVNSTGCGGCAGCAACCACAAC GTGCAAANSGVNSVNSGGCGCAAACAGTGGCGTCAGTGAC
	pINT BM40 前导 序列正向引物 GTTTGCCTGGCCGGGAGGGCTCTGGC
	pINT C λ Not 反 向引物 AGTCACGCTTGGTGC GGCGC

[1031] 表27. 所用引物的序列。

[1032] 使用MaxCyte STXG2、通过电穿孔将每一个文库的DNA送入经悬浮液调适的HEK293F细胞中(参见实例5)。类似地,克隆野生型KB\_A12 ProTx-III和KB\_A12HsTx1构建体且通过电穿孔送入HEK293F细胞中作为对照。每个文库使用1亿个细胞并且每个对照构建体使用1千万个细胞。转染后的第二天,以7.5μg/ml的浓度添加抗生素杀稻瘟菌素S HC1(LifeTech,目录号R21001)。细胞以0.25x10<sup>6</sup>个细胞/毫升接种于锥形烧瓶(Erlenmeyer flasks)中。为了计算基因整合效率,还将细胞接种到10cm培养皿中,如实例5中所述。KB\_A12 ProTx-III文库A集合、KB\_A12 ProTx-III文库B集合和KB\_A12 HsTx1文库的整合效率经计算分别是3%、2.5%和2%。因此,1亿细胞转染所达成的文库大小分别是3x10<sup>6</sup>、2.5x10<sup>6</sup>和2x10<sup>6</sup>。5天的杀稻瘟菌素S HC1(LifeTech,目录号R21001)选择之后,使用抗Fc MACS珠粒富集细胞,如实例5中所述。14天的选择之后,通过流式细胞术分析文库中的打结抗体表达水平。如图55所示,与HEK293细胞上所展示亲本打结抗体相比,所有三种打结抗体文库展示改善的平均表达水平。这证明了已经产生打结抗体变异数,与亲本打结抗体分子相比,所述打结抗体变异数将具有改善的生物物理学特性。

[1033] 接着,使用BioRad S3e细胞分选仪对打结抗体文库进行FACS。将MACS分选群体中的30x10<sup>6</sup>个细胞与抗人类Fc PE(1μl/1x10<sup>6</sup>个细胞)(剑桥生物科学,目录号409304)一起培育。在三种文库的针对高Fc表达的直方图上绘制门。对于KB\_A12ProTxIII A集合文库而言,由此获取8.16%的门内细胞,而在所述区域中发现0.15%的对照KB\_A12 ProTxIII。对于KB\_A12 ProTxIII B集合文库而言,获取7.25%的门内细胞,而在这个区域中发现0.10%的对照。获取12.6%的KB\_A12 HsTxI文库,而在这个区域中发现0.03%的KB\_A12 HsTxI对照。每个群体取0.5x10<sup>6</sup>个细胞用于提取基因组DNA。使用KOD热启动DNA聚合酶(默克密理博公

司),通过巢式PCR扩增编码IgG的DNA,如实例4中所述。对PCR产物进行凝胶提纯并且用NheI和XhoI消化,克隆到哺乳动物pINT3表达载体中并且用于转化大肠杆菌DH10B细胞。

[1034] 通过利用如实例4和5中所述的相同方法(其中建立哺乳动物细胞展示突变体文库,然后基于表面展示水平进行选择),鉴定出生物物理学特性与其亲本分子相比改善的个别打结抗体克隆。因此,这个实例已经描述了利用哺乳动物高等真核展示来改善打结抗体或任何替代双特异性形式的生物物理学特性。在此所述的方法还可以在打结抗体或双特异性发现项目期间进行,其中能够利用多参数FACS、基于高展示水平和对靶分子的特异性结合进行选择。

[1035] 参考文献

[1036] 1 Jarasch, A. 等人, 在新颖治疗抗体选择期间进行的可开发性评估 (Developability assessment during the selection of novel therapeutic antibodies),《药学杂志 (J Pharm Sci)》104, 1885-1898, doi:10.1002/jps.24430 (2015) .

[1037] 2 Jain, T. 等人, 临床阶段抗体概貌的生物物理学特性 (Biophysical properties of the clinical-stage antibody landscape),《美国国家科学院院刊 (Proceedings of the National Academy of Sciences of the United States of America)》114, 944-949, doi:10.1073/pnas.1616408114 (2017) .

[1038] 3 Yang, X. 等人, 工艺开发起始之前的可开发性研究:改善单克隆抗体的可制造性 (Developability studies before initiation of process development: improving manufacturability of monoclonal antibodies), mAbs 5, 787-794, doi:10.4161/mabs.25269 (2013) .

[1039] 4 Ratanji, K.D., Derrick, J.P., Dearman, R.J. 和 Kimber, I. 治疗蛋白的免疫原性:聚集的影响 (Immunogenicity of therapeutic proteins: influence of aggregation),《免疫毒理学杂志 (J Immunotoxicol)》11, 99-109, doi:10.3109/1547691X.2013.821564 (2014) .

[1040] 5 Moussa, E.M. 等人, 治疗蛋白聚集体的免疫原性 (Immunogenicity of Therapeutic Protein Aggregates),《药学杂志 (J Pharm Sci)》105, 417-430, doi:10.1016/j.xphs.2015.11.002 (2016) .

[1041] 6 Buchanan, A. 等人, 对治疗IgG分子进行工程改造以解决半胱氨酸化、聚集且增强热稳定性和表达 (Engineering a therapeutic IgG molecule to address cysteinylated, aggregation and enhance thermal stability and expression). mAbs 5, 255-262, doi:10.4161/mabs.23392 (2013) .

[1042] 7 Dobson, C.L. 等人, 对人单克隆抗体的表面特性进行工程改造以阻止自缔合和体内快速清除 (Engineering the surface properties of a human monoclonal antibody prevents self-association and rapid clearance in vivo),《科学报道 (Scientific Reports)》6, 38644, doi:10.1038/srep38644 <https://www.nature.com/articles/srep38644-supplementary-information> (2016) .

[1043] 8 Schofield, D.J. 等人, 将噬菌体呈现应用于高通量抗体产生和表征 (Application of phage display to high throughput antibody generation and characterization),《基因组生物学 (Genome biology)》8, R254, doi:10.1186/gb-2007-8-

11-r254(2007) .

[1044] 9Jespers,L.,Schon,O.,Famm,K.和Winter,G.在噬菌体上通过热变性来选择抗聚集结构域抗体(Aggregation-resistant domain antibodies selected on phage),《自然·生物技术(Nat Biotechnol)》22,1161-1165,doi:10.1038/nbt1000(2004) .

[1045] 10Fennell,B.J.等人,人类原生scFv的局限于CDR的工程改造建立了高度稳定的可溶性双功能抗体用于皮下递送(CDR-restricted engineering of native human scFvs creates highly stable and soluble bifunctional antibodies for subcutaneous delivery),mAbs 5,882-895,doi:10.4161/mabs.26201(2013) .

[1046] 11Christ,D.,Famm,K.和Winter,G.抗聚集人类抗体结构域谱系(Repertoires of aggregation-resistant human antibody domains),《蛋白质工程、设计和选择(Protein engineering,design&selection:PEDS)》20,413-416,doi:10.1093/protein/gzm037(2007) .

[1047] 12Tiller,T.等人,基于固定VH/VL构架对、具有有利生物物理学特性的全合成人Fab抗体文库(A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties),mAbs5,445-470,doi:10.4161/mabs.24218(2013) .

[1048] 13Dudgeon,K.等人,用于产生抗聚集性增强的人类抗体可变域的通用策略(General strategy for the generation of human antibody variable domains with increased aggregation resistance),《美国国家科学院院刊(Proceedings of the National Academy of Sciences of the United States of America)》109,10879-10884,doi:10.1073/pnas.1202866109(2012) .

[1049] 14Bethea,D.等人,人类单克隆抗体CNT0607的自缔合机理(Mechanisms of self-association of a human monoclonal antibody CNT0607),《蛋白质工程设计和选择(Protein Engineering Design and Selection)》25,531-538(2012) .

[1050] 15Kelly,R.L.等人,人类IgG1抗体在小鼠中的高通量交叉相互作用度量与清除率相关(High throughput cross-interaction measures for human IgG1antibodies correlate with clearance rates in mice).mAbs 7,770-777,doi:10.1080/19420862.2015.1043503(2015) .

[1051] 16Hötzl,I.等人,快速清除的抗体的风险缓解策略(A strategy for risk mitigation of antibodies with fast clearance).mAbs 4,753-760,doi:10.4161/mabs.22189(2012) .

[1052] 17Pries,A.R.,Secomb,T.W.和Gaehtgens,P.内皮细胞表面层(The endothelial surface layer),《欧洲生理学杂志(Pflugers Arch)》440,653-666,doi:10.1007/s004240000307(2000) .

[1053] 18Nieuwdorp,M.等人,内皮细胞腊梅糖:健康与血管疾病之间的势垒(The endothelial glycocalyx:a potential barrier between health and vascular disease),《脂质学新见(Curr Opin Lipidol)》16,507-511(2005) .

[1054] 19Xu,Y.等人,解决选自试管内酵母呈现系统的抗体的多特异性:基于FACS的高通量选择和分析工具(Addressing polyspecificity of antibodies selected from an in

vitro yeast presentation system:a FACS-based,high-throughput selection and analytical tool),《蛋白质工程、设计和选择:PEDS(Protein engineering,design&selection:PEDS)》26,663-670,doi:10.1093/protein/gzt047(2013).

[1055] 20Martin,W.L.,West,A.P.,Jr.,Gan,L.和Bjorkman,P.J.FcRn/异二聚体Fc复合物的2.8A晶体结构:pH依赖性结合的机理(Crystal structure at 2.8A of an FcRn/heterodimeric Fc complex:mechanism of pH-dependent binding),《分子细胞(Mol Cell)》7,867-877(2001).

[1056] 21Suzuki,T.等人,新生儿FcR在调控含有人类IgG1的Fc域的治疗蛋白的血清半衰期方面的重要性:单克隆抗体和Fc融合蛋白对人类新生儿FcR的亲和力的比较研究(Importance of neonatal FcR in regulating the serum half-life of therapeutic proteins containing the Fc domain of human IgG1:a comparative study of the affinity of monoclonal antibodies and Fc-fusion proteins to human neonatal FcR),《免疫学杂志(Journal of immunology)》(Baltimore,Md.:1950)184,1968-1976,doi:10.4049/jimmunol.0903296(2010).

[1057] 22Schoch,A.等人,抗体可变域在电荷介导下对FcRn依赖性药物动力学的影响(Charge-mediated influence of the antibody variable domain on FcRn-dependent pharmacokinetics),《美国国家科学院院刊》112,5997-6002,doi:10.1073/pnas.1408766112(2015).

[1058] 23Kelly,R.L.等人,靶标非依赖性可变区介导的对抗体清除的影响可以不依赖于FcRn(Target-independent variable region mediated effects on antibody clearance can be FcRn independent),mAbs 8,1269-1275,doi:10.1080/19420862.2016.1208330(2016).

[1059] 24Boder,E.T.和Wittrup,K.D.用于筛选组合型多肽文库的酵母表面呈现(Yeast surface display for screening combinatorial polypeptide libraries),《自然·生物技术(Nat Biotechnol)》15,553-557,doi:10.1038/nbt0697-553(1997).

[1060] 25Chao,G.等人,利用酵母表面呈现来分离和工程改造人类抗体(Isolating and engineering human antibodies using yeast surface display.),《自然·方案(Nat.Protocols)》1,755-768,doi:<http://www.nature.com/nprot/journal/v1/n2/suppinfo/>

[1061] [nprot.2006.94\\_S1.html](http://www.ncbi.nlm.nih.gov/nucleotide/nprot.2006.94_S1.html)(2006).

[1062] 26Shusta,E.V.,Kieke,M.C.,Parke,E.,Kranz,D.M.和Wittrup,K.D.酵母多肽融合体表面呈现水平可预测热稳定性和可溶性分泌效率(Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency),《分子生物学杂志(Journal of molecular biology)》292,949-956,doi:<https://doi.org/10.1006/jmbi.1999.3130>(1999).

[1063] 27Kowalski,J.M.,Parekh,R.N.,Mao,J.和Wittrup,K.D.蛋白质折叠稳定性能够决定逃避内质网品质控制的效率(Protein Folding Stability Can Determine the Efficiency of Escape from Endoplasmic Reticulum Quality Control),《生物化学杂志(J.Biol.Chem.)》273,19453-19458(1998).

- [1064] 28Kowalski,J.M.,Parekh,R.N.和Wittrup,K.D.缺乏二硫键的牛胰脏胰蛋白酶抑制剂突变体在酿酒酵母中的分泌效率与热力学稳定性相关(Secretion Efficiency in *Saccharomyces cerevisiae* of Bovine Pancreatic Trypsin Inhibitor Mutants Lacking Disulfide Bonds Is Correlated with Thermodynamic Stability),《生物化学(Biochemistry)》37,1264-1273,doi:10.1021/bi9722397(1998).
- [1065] 29Hackel,B.J.,Ackerman,M.E.,Howland,S.W.和Wittrup,K.D.稳定性和CDR组成偏向富集结合剂功能概貌(Stability and CDR Composition Biases Enrich Binder Functionality Landscapes),《分子生物学杂志(Journal of molecular biology)》401,84-96,doi:<https://doi.org/10.1016/j.jmb.2010.06.004>(2010).
- [1066] 30Julian,M.C.等人,移植的淀粉样蛋白基序结构域抗体的亲和力和稳定性的协同进化(Co-evolution of affinity and stability of grafted amyloid-motif domain antibodies),《蛋白质工程设计和选择(Protein Eng.Des.Sel.)》28,339-350,doi:10.1093/protein/gzv050(2015).
- [1067] 31Rabia,L.A.,Desai,A.A.,Jhajj,H.S.和Tessier,P.M.了解和克服抗体亲和力、特异性、稳定性和溶解度之间的折衷(Understanding and overcoming trade-offs between antibody affinity,specificity,stability and solubility),《生物化学工程杂志(Biochemical Engineering Journal)》137,365-374,doi:<https://doi.org/10.1016/j.bej.2018.06.003>(2018).
- [1068] 32Igawa,T.,Haraya,K.和Hattori,K.将抗体作为新颖的治疗抗体模式扫描能够从循环中排除可溶性抗原(Sweeping antibody as a novel therapeutic antibody modality capable of eliminating soluble antigens from circulation),《免疫学评论(Immunol Rev)》270,132-151,doi:10.1111/imr.12392(2016).
- [1069] 33Qin,J.Y.等人,组成型启动子与多西环素诱导型启动子的系统比较(Systematic comparison of constitutive promoters and the doxycycline-inducible promoter),《公共科学图书馆·综合(PloS one)》5,e10611,doi:10.1371/journal.pone.0010611(2010).
- [1070] 34Walker,J.M.蛋白质方案手册(The protein protocols handbook)第3版(Humana出版社,2009).
- [1071] 35Janson,J.-C.《蛋白质提纯:原理、高分辨率方法和应用(Protein purification:principles,high resolution methods,and applications)》第3版(John Wiley&Sons,2011).
- [1072] 36Schiel,J.E.,Davis,D.L.和Borisov,O.《用于治疗单克隆抗体表征的工艺现状和新兴技术(State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization)》(美国化学学会,2014).
- [1073] 37Saro,D.等人,在《用于治疗单克隆抗体表征的工艺现状和新兴技术(State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization)》第2卷,《生物医药表征:NISTmAb案例分析(Biopharmaceutical Characterization:The NISTmAb Case Study)》第1201卷,ACS研讨会丛刊(ACS Symposium Series)第7章,329-355(美国化学学会,2015).

[1074] 38Sun,T.等人,通过生物层干涉量测术高通量检测抗体自身相互作用(High throughput detection of antibody self-interaction by bio-layer interferometry)mAbs 5,838-841,doi:10.4161/mabs.26186(2013).

[1075] 39Liu,Y.等人,在利用自身相互作用纳米颗粒光谱法的早期抗体发现期间根据可开发性进行的高通量筛选(High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy).mAbs 6,483-492,doi:10.4161/mabs.27431(2014).

[1076] 40Geng,S.B.,Wittekind,M.,Vigil,A.和Tessier,P.M.单克隆抗体自缔合的测量结果与复杂的生物物理学特性相关(Measurements of Monoclonal Antibody Self-Association Are Correlated with Complex Biophysical Properties),《分子制药学(Molecular Pharmaceutics)》13,1636-1645,doi:10.1021/acs.molpharmaceut.6b00071(2016).

[1077] 41Wu,J.等人,提纯之前利用亲和力捕捉自身相互作用纳米颗粒光谱法发现高度可溶性抗体(Discovery of highly soluble antibodies prior to purification using affinity-capture self-interaction nanoparticle spectroscopy),《蛋白质工程、设计和选择:PEDS》28,403-414,doi:10.1093/protein/gzv045(2015).

[1078] 42Nobmann,U.等人,动态光散射作为相对工具用于评估单克隆抗体的分子完整性和稳定性(Dynamic light scattering as a relative tool for assessing the molecular integrity and stability of monoclonal antibodies),《生物技术和基因工程评论(Biotechnology and Genetic Engineering Reviews)》24,117-128(2007).

[1079] 43Geng,S.B.,Cheung,J.K.,Narasimhan,C.,Shameem,M.和Tessier,P.M.利用胶体蛋白质相互作用的测量结果来改善单克隆抗体选择和工程化(Improving monoclonal antibody selection and engineering using measurements of colloidal protein interactions),《药学杂志(J Pharm Sci)》103,3356-3363,doi:10.1002/jps.24130(2014).

[1080] 44Howlett,G.J.,Minton,A.P.和Rivas,G.用于研究蛋白质缔合和组装的分析型超离心法(Analytical ultracentrifugation for the study of protein association and assembly),《化学生物学新见(Curr Opin Chem Biol)》10,430-436,doi:10.1016/j.cbpa.2006.08.017(2006).

[1081] 45Salinas,B.A.等人,了解和调节单克隆抗体配制物的乳白化和粘度(Understanding and modulating opalescence and viscosity in a monoclonal antibody formulation),《药学杂志》99,82-93,doi:10.1002/jps.21797(2010).

[1082] 46Gabel,F.等人,通过中子散射研究蛋白质动态(Protein dynamics studied by neutron scattering),《生物物理学季评(Q Rev Biophys)》35,327-367(2002).

[1083] 47Jacobs,S.A.,Wu,S.J.,Feng,Y.,Bethea,D.和O'Neil,K.T.交叉相互作用色谱:一种快速鉴定高度可溶性单克隆抗体候选物的方法(Cross - interaction chromatography:A rapid method to identify highly soluble monoclonal antibody candidates),《药物研究(Pharm.Res.)》27,65-71,doi:10.1007/s11095-009-0007-z(2010).

- [1084] 48 Josic,D., Bal,F. 和 Schwinn,H. 通过肝素亲和色谱法从凝血级联中分离血浆蛋白 (Isolation of plasma proteins from the clotting cascade by heparin affinity chromatography), 《色谱学杂志 (J Chromatogr)》632, 1-10 (1993).
- [1085] 49 Fekete,S., Veuthey,J.L., Beck,A. 和 Guillarme,D. 用于表征单克隆抗体和相关产物的疏水相互作用色谱 (Hydrophobic interaction chromatography for the characterization of monoclonal antibodies and related products), 《制药和生物医药分析杂志 (J Pharm Biomed Anal)》130, 3-18, doi:10.1016/j.jpba.2016.04.004 (2016).
- [1086] 50 Eriksson,K.O. 和 Belew,M. 疏水相互作用色谱 (Hydrophobic interaction chromatography), 《生物化学分析方法 (Methods Biochem Anal)》54, 165-181 (2011).
- [1087] 51 Bakalova,R. 和 Ohba,H. 大豆凝集素与白血病T细胞的相互作用以及其用于通过凝集素亲和色谱法与正常淋巴细胞试管内分离的用途 (Interaction of soybean agglutinin with leukemic T-cells and its use for their in vitro separation from normal lymphocytes by lectin-affinity chromatography), 《生物医学色谱 (Biomed Chromatogr)》17, 239-249, doi:10.1002/bmc.218 (2003).
- [1088] 52 Melidoni,A.N., Dyson,M.R., Wormald,S. 和 McCafferty,J. 选择拮抗性抗体以通过胚胎干细胞的诱导型表达控制分化 (Selecting antagonistic antibodies that control differentiation through inducible expression in embryonic stem cells), 《美国国家科学院院刊》110, 17802-17807, doi:10.1073/pnas.1312062110 (2013).
- [1089] 53 Xie,J., Zhang,H., Yea,K. 和 Lerner,R.A. 自分泌信号传导是基于使人类干细胞转分化的组合型抗体的选择 (Autocrine signaling based selection of combinatorial antibodies that transdifferentiate human stem cells), 《美国国家科学院院刊》110, 8099-8104, doi:10.1073/pnas.1306263110 (2013).
- [1090] 54 Zhang,H. 等人, 从组合型抗体文库感染的单一细胞中选择促效剂 (Selecting Agonists from Single Cells Infected with Combinatorial Antibody Libraries), 《化学与生物学 (Chemistry & Biology)》20, 734-741, doi:10.1016/j.chembiol.2013.04.012 (2013).
- [1091] 55 Igoucheva,O., Alexeev,V. 和 Yoon,K. 在哺乳动物细胞中利用小单链寡核苷酸进行靶基因校正 (Targeted gene correction by small single-stranded oligonucleotides in mammalian cells). 《基因疗法 (Gene Therapy)》8, 391, doi:10.1038/sj.gt.3301414 (2001).
- [1092] 56 Liang,X., Potter,J., Kumar,S., Ravinder,N. 和 Chesnut,J.D. 通过改善gRNA、Cas9核酸酶和供体DNA的设计和递送来增强CRISPR/Cas9介导的精确基因组编辑 (Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA). 《生物技术杂志 (J.Biotechnol.)》241, 136-146, doi:<https://doi.org/10.1016/j.jbiotec.2016.11.011> (2017).
- [1093] 57 Dyson,M.R. 等人, 通过将抗体亲和力成熟与质谱学组合来定位蛋白质相互作用 (Mapping protein interactions by combining antibody affinity maturation and mass spectrometry), 《分析生物化学 (Analytical biochemistry)》417, 25-35, doi:

10.1016/j.ab.2011.05.005(2011).

[1094] 58Gronwald,R.G.等人,编码人类血小板源生长因子受体的cDNA的克隆和表达:超过一种受体类别的证据(Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor:evidence for more than one receptor class),《美国国家科学院院刊》85,3435-3439(1988).

[1095] 59Dux,R.,Kindler-Rohrborn,A.,Lennartz,K.和Rajewsky,M.F.校准荧光强度以通过流式细胞术定量细胞亚群的抗体结合表面决定子(Calibration of fluorescence intensities to quantify antibody binding surface determinants of cell subpopulations by flow cytometry),《细胞测量术(Cytometry)》12,422-428,doi:10.1002/cyto.990120507(1991).

[1096] 60Brockhoff,G.,Hofstaedter,F.和Knuechel,R.流式细胞术对人类膀胱癌细胞系中的表皮生长因子受体的检测和量化与史卡查分析的比较(Flow cytometric detection and quantitation of the epidermal growth factor receptor in comparison to Scatchard analysis in human bladder carcinoma cell lines),《细胞测量术》17,75-83,doi:10.1002/cyto.990170110(1994).

[1097] 61Gordon,I.L.对荧光刀豆凝集素A结合到淋巴细胞的史卡查分析(Scatchard analysis of fluorescent concanavalin A binding to lymphocytes),《细胞测量术》20,238-244,doi:10.1002/cyto.990200307(1995).

[1098] 62Sheehan,J.和Marasco,W.A.噬菌体和酵母呈现(Phage and Yeast Display),《微生物学波谱(Microbiol Spectr)》3,AID-0028-2014,doi:10.1128/microbiolspec.AID-0028-2014(2015).

[1099] 63Puck,T.T.体细胞哺乳动物细胞的遗传学(The genetics of somatic mammalian cells),《实验医学与生物学进展(Adv Biol Med Phys)》5,75-101(1957).

[1100] 64Wurm,F. CHO准种一对制造过程的影响(Quasispecies-Implications for Manufacturing Processes),《过程(Processes)》1,296-311(2013).

[1101] 65Goh,J.B.和Ng,S.K.宿主细胞系选择对多糖特征的影响(Impact of host cell line choice on glycan profile),《生物技术重要评论(Critical Reviews in Biotechnology)》38,851-867,doi:10.1080/07388551.2017.1416577(2018).

[1102] 66Skerra,A.分子识别用的替代非抗体支架(Alternative non-antibody scaffolds for molecular recognition),《生物技术新见(Curr Opin Biotechnol)》18,295-304,doi:10.1016/j.copbio.2007.04.010(2007).

[1103] 67Gebauer,M.和Skerra,A.作为下一代抗体治疗剂的工程化蛋白质支架(Engineered protein scaffolds as next-generation antibody therapeutics),《化学生物学新见(Curr Opin Chem Biol)》13,245-255,doi:10.1016/j.cbpa.2009.04.627(2009).

[1104] 68Tiede,C.等人,得隆肽:一种用于分子识别应用的稳定和通用肽呈现支架(Adhiron:a stable and versatile peptide display scaffold for molecular recognition applications),《蛋白质工程、设计和选择:PEDS(Protein engineering, design&selection:PEDS)》27,145-155,doi:10.1093/protein/gzu007(2014).

- [1105] 69Koide,A.,Bailey,C.W.,Huang,X.和Koide,S.纤维结合蛋白III型结构域作为新颖结合蛋白的支架(The fibronectin type III domain as a scaffold for novel binding proteins),《分子生物学杂志(Journal of molecular biology)》284,1141-1151,doi:10.1006/jmbi.1998.2238(1998).
- [1106] 70Nygren,P.A.和Skerra,A.来自替代支架的结合蛋白(Binding proteins from alternative scaffolds),《免疫学方法杂志(J Immunol Methods)》290,3-28,doi:10.1016/j.jim.2004.04.006(2004).
- [1107] 71Chang,H.J.等人,胱氨酸稳定化小蛋白质作为稳定蛋白质结合剂的分子进化(Molecular evolution of cystine-stabilized miniproteins as stable proteinaceous binders),《结构(Structure)》17,620-631,doi:10.1016/j.str.2009.01.011(2009).
- [1108] 72Ward,E.S.,Güssow,D.,Griffiths,A.D.,Jones,P.T.和Winter,G.从大肠杆菌分泌的单一免疫球蛋白可变域谱系的结合活性(Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli),《自然(Nature)》341,544,doi:10.1038/341544a0(1989).
- [1109] 73McCafferty,J.,Griffiths,A.D.,Winter,G.和Chiswell,D.J.噬菌体抗体:呈现抗体可变域的丝状噬菌体(Phage antibodies: filamentous phage displaying antibody variable domains),《自然》348,552-554,doi:10.1038/348552a0(1990).
- [1110] 74Holt,L.J.,Herring,C.,Jespers,L.S.,Woolven,B.P.和Tomlinson,I.M.结构域抗体:用于治疗的蛋白质(Domain antibodies:proteins for therapy),《生物技术趋势(Trends in Biotechnology)》21,484-490,doi:10.1016/j.tibtech.2003.08.007(2003).
- [1111] 75Huston,J.S.等人,抗体结合位点的蛋白质工程:恢复大肠杆菌所产生的抗地高辛单链Fv类似物的特异活性(Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli),《美国国家科学院院刊》85,5879(1988).
- [1112] 76Bird,R.E.等人,单链抗原结合蛋白(Single-chain antigen-binding proteins),《科学(Science)》242,423-426(1988).
- [1113] 77Holliger,P.,Prospero,T.和Winter,G.“双链抗体”:二价双特异性抗体小片段(“Diabodies”:Small bivalent and bispecific antibody fragments),《美国国家科学院院刊》90,6444-6448(1993).
- [1114] 78Reiter,Y.,Brinkmann,U.,Lee,B.和Pastan,I.用于癌症检测和治疗的工程化抗体Fv片段:二硫键稳定化Fv片段(Engineering antibody Fv fragments for cancer detection and therapy:Bisulfide-stabilized Fv fragments),《自然·生物技术》14,1239,doi:10.1038/nbt1096-1239(1996).
- [1115] 79Holliger,P.和Hudson,P.J.工程化抗体片段和单一结构域的升高(Engineered antibody fragments and the rise of singledomains),《自然·生物技术》23,1126-1136,doi:10.1038/nbt1142(2005).
- [1116] 80Knappik,A.等人,基于模块化共同构架和三核苷酸随机化CDR的全合成人类组合型抗体文库(HuCAL)(Fully synthetic human combinatorial antibody libraries

(HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides),《分子生物学杂志 (Journal of molecular biology)》296, 57-86, doi: 10.1006/jmbi.1999.3444(2000).

[1117] 81Krebs, B. 等人, 重组人类抗体的高通量产生和工程化 (High-throughput generation and engineering of recombinant human antibodies),《免疫学方法杂志 (J Immunol Methods)》254, 67-84(2001).

[1118] 82Holliger, P. 和 Bohlen, H. 用于临床的工程化抗体 (Engineering antibodies for the clinic),《癌症和转移评述 (Cancer Metastasis Rev)》18, 411-419(1999).

[1119] 83Glennie, M.J., McBride, H.M., Worth, A.T. 和 Stevenson, G.T. 含有硫醚连接的 Fab'  $\gamma$  片段的双特异性 F(ab'  $\gamma$ )<sub>2</sub> 抗体的制备和性能 (Preparation and performance of bispecific F(ab' gamma)2 antibody containing thioether-linked Fab' gamma fragments),《免疫学杂志 (Journal of immunology)》(Baltimore, Md.: 1950) 139, 2367-2375(1987).

[1120] 84Repp, R. 等人, 使用针对 Fc  $\gamma$  RI 和 HER-2/neu 的双特异性抗体 (MDX-210) 免疫治疗乳癌的经 G-CSF 刺激的 PMN (G-CSF-stimulated PMN in immunotherapy of breast cancer with a bispecific antibody to Fc gamma RI and to HER-2/neu (MDX-210)),《血液治疗杂志 (J Hematother)》4, 415-421, doi: 10.1089/scd.1.1995.4.415(1995).

[1121] 85Suresh, M.R., Cuello, A.C. 和 Milstein, C. 得自杂合体融合瘤的双特异性单克隆抗体 (Bispecific monoclonal antibodies from hybrid hybridomas),《酶学方法 (Methods in enzymology)》121, 210-228(1986).

[1122] 86Staerz, U.D. 和 Bevan, M.J. 产生能够集中效应 T 细胞活性的双特异性单克隆抗体的杂合体融合瘤 (Hybrid hybridoma producing a bispecific monoclonal antibody that can focus effector T-cell activity),《美国国家科学院院刊》83, 1453(1986).

[1123] 87Merchant, A.M. 等人, 人类双特异性 IgG 的有效途径 (An efficient route to human bispecific IgG),《自然·生物技术》16, 677, doi: 10.1038/nbt0798-677(1998).

[1124] 88Von Kreudenstein, T.S. 等人, 改善双特异性抗体支架的生物物理学特性以促进可开发性: 分子设计的品质 (Improving biophysical properties of a bispecific antibody scaffold to aid developability: quality by molecular design), mAbs 5, 646-654, doi: 10.4161/mabs.25632(2013).

[1125] 89Ghetie, V. 等人, 通过随机诱变来增加 IgG 片段的血清持久性 (Increasing the serum persistence of an IgG fragment by random mutagenesis),《自然·生物技术》15, 637-640, doi: 10.1038/nbt0797-637(1997).

[1126] 90Borrok, M.J. 等人, 为了改善稳定性而设计的具有延长的半衰期的“Fc 静默” IgG1 形式 (An "Fc-Silenced" IgG1 Format With Extended Half-Life Designed for Improved Stability),《药学杂志 (J Pharm Sci)》106, 1008-1017, doi: 10.1016/j.xphs.2016.12.023(2017).

[1127] 91Schlothauer, T. 等人, 用于功能表征单克隆抗体的分析型 FcRn 亲和色谱法 (Analytical FcRn affinity chromatography for functional characterization of monoclonal antibodies). mAbs 5, 576-586, doi: 10.4161/mabs.24981(2013).

- [1128] 92Gossen,M.等人,四环素在哺乳动物细胞中引起的转录活化(Transcriptional activation by tetracyclines in mammalian cells),《科学(Science)》268,doi:10.1126/science.7792603(1995).
- [1129] 93Sadelain,M.,Papapetrou,E.P.和Bushman,F.D.用于将新DNA整合到人类基因组中的安全港(Safe harbours for the integration of new DNA in the human genome),《自然评论(Nature reviews)》,《癌症(Cancer)》12,51-58,doi:10.1038/nrc3179(2011).
- [1130] 94Foote,J.和Winter,G.影响高变环构象的抗体构架残基(Antibody framework residues affecting the conformation of the hypervariable loops),《分子生物学杂志(Journal of molecular biology)》224,487-499(1992).
- [1131] 95Haghparast,S.M.,Kihara,T.和Miyake,J.HEK293细胞在粘附和悬浮状态下的不同力学行为(Distinct mechanical behavior of HEK293 cells in adherent and suspended states),《PeerJ》3,e1131,doi:10.7717/peerj.1131(2015).
- [1132] 96Rivera,V.M.等人,通过控制内质网中的聚集来调控蛋白质分泌(Regulation of protein secretion through controlled aggregation in the endoplasmic reticulum),《科学(Science)》287,826-830(2000).
- [1133] 97Ferri,N.,Corsini,A.,Sirtori,C.R.和Ruscica,M.用于治疗高胆固醇血症的博可珠单抗(Bococizumab for the treatment of hypercholesterolaemia),《生物疗法专家意见(Expert Opin Biol Ther)》17,237-243,doi:10.1080/14712598.2017.1279602(2017).
- [1134] 98Liang,H.等人,前蛋白转化酶枯草杆菌蛋白酶/枯草溶菌素9型拮抗作用减少了经抑制素治疗的高胆固醇脂血症非人类灵长类动物的低密度脂蛋白胆固醇(Proprotein Convertase Subtilisin/Kexin Type 9Antagonism Reduces Low-Density Lipoprotein Cholesterol in Statin-Treated Hypercholesterolemic Nonhuman Primates),《药理学和实验治疗学杂志(Journal of Pharmacology and Experimental Therapeutics)》340,228(2012).
- [1135] 99Jones,T.D.,Crompton,L.J.,Carr,F.J.和Baker,M.P.在《治疗抗体:方法和方案(Therapeutic Antibodies:Methods and Protocols)》(Antony S.Dimitrov编)405-423(Humana出版社,2009).
- [1136] 100Bostrom,J.,Lee,C.V.,Haber,L.和Fuh,G.在《治疗抗体:方法和方案》(Antony S.Dimitrov编)353-376(Humana出版社,2009).
- [1137] 101Dyson,M.R.,Shadbolt,S.P.,Vincent,K.J.,Perera,R.L.和McCafferty,J.在大肠杆菌中产生可溶性哺乳动物蛋白质:与成功表达相关的蛋白质特征的鉴定(Production of soluble mammalian proteins in Escherichia coli:identification of protein features that correlate with successful expression),《BMC生物技术(BMC biotechnology)》4,32,doi:10.1186/1472-6750-4-32(2004).
- [1138] 102Neylon,C.用于使蛋白质编码DNA序列随机突变的化学和生物化学策略:用于定向进化的文库构建方法(Chemical and biochemical strategies for the randomization of protein encoding DNA sequences:library construction methods

for directed evolution),《核酸研究(Nucleic acids research)》32,1448-1459,doi:10.1093/nar/gkh315(2004).

[1139] 103Fujino,Y.等人,基于界面聚焦高通量突变扫描的稳健性试管内亲和力成熟策略(Robust in vitro affinity maturation strategy based on interface-focused high-throughput mutational scanning),《生物化学和生物物理学研究通讯(Biochemical and biophysical research communications)》428,395-400,doi:10.1016/j.bbrc.2012.10.066(2012).

[1140] 104Xin,Y.等人,抗神经纤毛蛋白-1(MNRP1685A):从临床前模型到人类跨越物种的出人意料的药物动力学差异(unexpected pharmacokinetic differences across species,from preclinical models to humans),《药物研究(Pharm Res)》29,2512-2521,doi:10.1007/s11095-012-0781-x(2012).

[1141] 105Patnaik,A.等人,评估MNRP1685A、完整人类抗NRP1单克隆抗体与贝伐单抗和太平洋紫杉醇的组合对晚期实体肿瘤患者的Ib期研究(A Phase Ib study evaluating MNRP1685A,a fully human anti-NRP1 monoclonal antibody,in combination with bevacizumab and paclitaxel in patients with advanced solid tumors),《癌症化疗与药理学(Cancer Chemother Pharmacol)》73,951-960,doi:10.1007/s00280-014-2426-8(2014).

[1142] 106Shin,D.S.和Ribas,A.检查点阻断作为癌症疗法的演变:现状、将来(The evolution of checkpoint blockade as a cancer therapy:What's here,what's next?),《免疫学新见(Curr.Opin.Immunol.)》33,23-35,doi:10.1016/j.coi.2015.01.006(2015).

[1143] 107Mouquet,H.等人,多反应性通过异型接合来增强抗HIV抗体的表观亲和力(Polyreactivity increases the apparent affinity of anti-HIV antibodies by heterooligation),《自然》467,591-595,doi:10.1038/nature09385(2010).

[1144] 108Fellouse,F.A.和Sidhu,S.S.在《制备和使用抗体:实用手册(Making and Using Antibodies:A Practical Handbook)》(G.C.Howard和M.R.Kaser编)第8章,157-180(CRC出版社,2007).

[1145] 109Peterson,M.L.在B细胞发育期间控制膜和分泌的免疫球蛋白的产生的机理(Mechanisms controlling production of membrane and secreted immunoglobulin during B cell development),《免疫学研究(Immunol Res)》37,33-46(2007).

[1146] 110Peterson,M.L.和Perry,R.P. $\mu$ m和 $\mu$ s mRNA的调控产生依赖于 $\mu$ s聚(腺苷酸)位点使用率和c $\mu$ 4至M1剪接的相对效率(The regulated production of mu m and mu s mRNA is dependent on the relative efficiencies of mu s poly(A) site usage and the c mu 4-to-M1 splice),《分子细胞生物学(Mol Cell Biol)》9,726-738(1989).

[1147] 111Yu,B.,Wages,J.M.和Larrick,J.W.用于促进细胞系发育的抗体-膜开关(AMS)技术(Antibody-membrane switch(AMS) technology for facile cell line development),《蛋白质工程设计和选择(Protein Engineering Design and Selection)》27,309-315,doi:10.1093/protein/gzu039(2014).

[1148] 112Wang,C.等人,抗PD-1抗体尼沃单抗、BMS-936558的试管内表征和在非人类灵

长类动物中的体内毒理学 (In vitro characterization of the anti-PD-1 antibody nivolumab, BMS-936558, and in vivo toxicology in non-human primates), 《癌症免疫学研究 (Cancer Immunol Res)》2, 846-856, doi:10.1158/2326-6066.CIR-14-0040 (2014).

[1149] 113D' Hautcourt, J.-L. 在《流式细胞术实验指南 (Current Protocols in Cytometry)》(John Wiley&Sons, Inc., 2001).

[1150] 114Gossen, M. 和 Bujard, H. 通过四环素响应性启动子紧密控制哺乳动物细胞中的基因表达 (Tight control of gene expression in mammalian cells by tetracycline-responsive promoters), 《美国国家科学院院刊》89, doi:10.1073/pnas.89.12.5547 (1992).

[1151] 115Ryan, M.D. 和 Drew, J. 口蹄疫病毒 2A 寡肽介导人工多聚蛋白的裂解 (Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein), 《欧洲分子生物学学会杂志 (The EMBO Journal)》13, 928-933 (1994).

[1152] 116Zhou, X., Vink, M., Klaver, B., Berkhout, B. 和 Das, A.T. 优化 Tet-On 系统以便通过病毒进化来调控基因表达 (Optimization of the Tet-On system for regulated gene expression through viral evolution), 《基因疗法 (Gene Ther)》13, 1382-1390, doi:10.1038/sj.gt.3302780 (2006).

[1153] 117Loew, R., Heinz, N., Hampf, M., Bujard, H. 和 Gossen, M. 背景表达最小化的经改善的 Tet 响应性启动子 (Improved Tet-responsive promoters with minimized background expression), 《BMC 生物技术 (BMC biotechnology)》10, 81, doi:10.1186/1472-6750-10-81 (2010).

[1154] 118To, T.L. 和 Maheshri, N. 杂讯能够诱导无双稳态的正向转录反馈回路产生双峰 (Noise can induce bimodality in positive transcriptional feedback loops without bistability), 《科学》327, 1142-1145, doi:10.1126/science.1178962 (2010).

[1155] 119Mullik, A. 等人, 二甲氨基二硫代甲酸铜基因开关: 一种用于调控哺乳动物细胞表达的系统 (The cumate gene-switch: a system for regulated expression in mammalian cells), 《BMC 生物技术》6, 1-18, doi:10.1186/1472-6750-6-43 (2006).

[1156] 120Yao, F. 等人, 四环素抑制因子 tetR 而非 tetR- 哺乳动物细胞转录因子融合体衍生物调控哺乳动物细胞中的诱导型基因表达 (Tetracycline repressor, tetR, rather than the tetR-mammalian cell transcription factor fusion derivatives, regulates inducible gene expression in mammalian cells), 《人类基因疗法 (Hum Gene Ther)》9, doi:10.1089/hum.1998.9.13-1939 (1998).

[1157] 121Wang, Y., O'Malley, B.W., Jr., Tsai, S.Y. 和 O'Malley, B.W. 用于基因转移的调控系统 (A regulatory system for use in gene transfer), 《美国国家科学院院刊》91, 8180-8184 (1994).

[1158] 122Datta-Mannan, A. 等人, 非特异性结合、靶标介导式清除和 FcRn 相互作用对人源化抗体的药物动力学的相互影响 (The interplay of non-specific binding, target-mediated clearance and FcRn interactions on the pharmacokinetics of humanized antibodies), mAbs 7, 1084-1093, doi:10.1080/19420862.2015.1075109 (2015).

[1159] 123Sigounas, G., Harindranath, N., Donadel, G. 和 Notkins, A.L. 多反应性抗体的

半衰期(Half-life of polyreactive antibodies),《临床免疫学杂志(J Clin Immunol)》14,134-140(1994).

[1160] 124Roopenian,D.C.和Akilesh,S.FcRn:新生儿Fc受体成熟(FcRn:the neonatal Fc receptor comes of age),《自然·免疫学评述(Nat Rev Immunol)》7,715-725,doi:10.1038/nri2155(2007).

[1161] 125Eid,J.等人,单一聚合酶分子的实时DNA测序(Real-time DNA sequencing from single polymerase molecules),《科学》323,133-138,doi:10.1126/science.1162986(2009).

[1162] 126Giordano,F.等人,来自MinION、PacBio和MiSeq平台的酵母基因组重新组装(De novo yeast genome assemblies from MinION, PacBio and MiSeq platforms),《科学报道(Scientific Reports)》7,3935,doi:10.1038/s41598-017-03996-z(2017).

[1163] 127Ehrenmann,F.,Duroux,P.,Giudicelli,V.和Lefranc,M.-P.《使用IMGT®对抗体进行的标准化序列和结构分析(Standardized Sequence and Structure Analysis of Antibody Using IMGT®)》11-31,doi:10.1007/978-3-642-01147-4\_2(2010).

[1164] 128Van Blarcom,T.等人,生产性共同轻链文库产生了一组多样性高亲和力双特异性抗体(Productive common light chain libraries yield diverse panels of high affinity bispecific antibodies).mAbs 10,256-268,doi:10.1080/19420862.2017.1406570(2018).

[1165] 129Spiess,C.,Zhai,Q.和Carter,P.J.双特异性抗体的替代分子形式和治疗应用(Alternative molecular formats and therapeutic applications for bispecific antibodies),《分子免疫学(Molecular Immunology)》67,95-106,doi:<http://dx.doi.org/10.1016/j.molimm.2015.01.003>(2015).

[1166] 130Husain,B.和Ellerman,D.通过双特异性抗体扩增生物治疗剂的边界(Expanding the Boundaries of Biotherapeutics with Bispecific Antibodies),《生物药物(BioDrugs)》32,441-464,doi:10.1007/s40259-018-0299-9(2018).

[1167] 131Fischer,N.等人,利用轻链对原生人类双特异性IgG进行可放大的产生和平台提纯(Exploiting light chains for the scalable generation and platform purification of native human bispecific IgG),《自然通讯(Nat Commun)》6,6113,doi:10.1038/ncomms7113(2015).

[1168] 132Ridgway,J.B.B.,Presta,L.G.和Carter,P.用于重链异二聚的抗体CH3域的'白包杵'工程化('Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization),《蛋白质工程、设计和选择(Protein Eng.Des.Sel.)》9,617-621,doi:10.1093/protein/9.7.617(1996).

[1169] 133Wozniak-Knopp,G.等人,将抗原结合位点引入免疫球蛋白恒定域的结构环中:具有工程化HER2/neu结合位点和抗体特性的Fc片段(Introducing antigen-binding sites in structural loops of immunoglobulin constant domains:Fc fragments with engineered HER2/neu-binding sites and antibody properties),《蛋白质工程、设计和选择:PEDS》23,289-297,doi:10.1093/protein/gzq005(2010).

[1170] 134Traxlmayr,M.W.等人,Her2/neu结合IgG1-Fc的定向进化以便利用酵母表面呈

现改善稳定性和聚集抗性(Directed evolution of Her2/neu-binding IgG1-Fc for improved stability and resistance to aggregation by using yeast surface display),《蛋白质工程、设计和选择:PEDS》26,255-265,doi:10.1093/protein/gzs102(2013) .

[1171] 135 Mahlangu,J.等人,艾密次单抗防治无抑制因子的A型血友病患者(Emicizumab Prophylaxis in Patients Who Have Hemophilia A without Inhibitors),《新英格兰医学杂志(N Engl J Med)》379,811-822,doi:10.1056/NEJMoa1803550(2018) .

[1172] 136 Sampei,Z.等人,鉴定和多维度优化不对称双特异性IgG抗体、从而模拟因子VIII辅因子活性功能(Identification and Multidimensional Optimization of an Asymmetric Bispecific IgG Antibody Mimicking the Function of Factor VIIIICofactor Activity),《公共科学图书馆·综合(PloS one)》8,e57479,doi:10.1371/journal.pone.0057479(2013) .

[1173] 137 Tabrizi,M.A.,Tseng,C.M.和Roskos,L.K.治疗性单克隆抗体的排除机理(Elimination mechanisms of therapeutic monoclonal antibodies),《今日药物发现(Drug Discov Today)》11,81-88,doi:10.1016/S1359-6446(05)03638-X(2006) .

[1174] 138 Dostalek,M.,Gardner,I.,Gurbaxani,B.M.,Rose,R.H.和Chetty,M.单克隆抗体基于药物动力学、药效学和生理学的药物动力学建模(Pharmacokinetics, pharmacodynamics and physiologically-based pharmacokinetic modelling of monoclonal antibodies),《临床药物动力学(Clin Pharmacokinet)》52,83-124,doi:10.1007/s40262-012-0027-4(2013) .

[1175] 139. Sondermann,P.和Szymkowski,D.E.利用Fc受体生物学设计治疗抗体(Harnessing Fc receptor biology in the design of therapeutic antibodies),《免疫学新见(Curr Opin Immunol)》40,78-87(2016) .

[1176] 140. Park,H.I.,Yoon,H.W.和Jung,S.T.高度可进化的抗体Fc域(The Highly Evolvable Antibody Fc Domain),《生物技术趋势(Trends in Biotechnology)》34,895-908(2016) .

[1177] 141. Hogarth,P.M.和Pietersz,G.A.用于治疗发炎、癌症和其它的Fc受体靶向疗法(Fc receptor-targeted therapies for the treatment of inflammation,cancer and beyond.Nature reviews),《自然评论(Nature reviews)》,《药物发现(Drug discovery)》11,311-331(2012) .

[1178] 142. Mayes,P.A.,Hance,K.W.和Hoos,A.癌症的免疫促效抗体开发的前景和挑战(The promise and challenges of immune agonist antibody development in cancer),《自然评论》,《药物发现》17,509-527(2018) .

[1179] 143. Jacobsen,F.W.等人,工程改造缺乏效应功能、但具有优化的可开发性的IgG支架(Engineering an IgG scaffold lacking effector function with optimized developability),《生物化学杂志》(2016) .

## 序列表

<110> 艾恩塔斯有限公司

<120> 根据可开发性选择真核细胞展示系统中的多肽药物

<130> NRS/CP7407091

<140> PCT/EP2018/083698

<141> 2018-12-05

<150> GB 1720351.4

<151> 2017-12-06

<160> 368

<170> PatentIn version 3.5

<210> 1

<211> 328

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 1

ttcgcttca cacgagttta ctccctatca gtgatagaga acgtatgtcg agtttactcc 60

[0001] ctatcagtga tagagaacga tgcgagttt actccctatc agtgatagag aacgtatgtc 120

gagtttactc cctatcagt atagagaacg tatgtcgagt ttactcccta tcagtgtatag 180

agaacgtatg tcgagtttat ccctatcagt gatagagaac gtatgtcgag ttactccct 240

atcagtgata gagaacgtat gtcgaggtag gcgtgtacgg tgggaggcct atataaggcag 300

agctcgttta gtgaaccgtc agatcgcc 328

<210> 2

<211> 123

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 2

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg

1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Thr Asn Tyr

20 25 30

Gly Met His Trp Gly Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val

35 40 45

Ala Val Ile Ser His Asp Gly Asn Asn Lys Tyr Tyr Val 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 Arg Glu Gly Ile Asp Phe Trp Ser Gly Leu Asn Trp Phe Asp Pro  
100 105 110  
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
115 120  
<210> 3  
<211> 108  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 3  
Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly  
1 5 10 15  
Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Ile Thr Gly Ser  
20 25 30  
[0002] Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu  
35 40 45  
Ile Cys Gly Ala Ser Ser Trp Ala Thr Gly Ile Pro Asp Arg Phe Ser  
50 55 60  
Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu  
65 70 75 80  
Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Ser Ser Ser Pro  
85 90 95  
Ile Thr Phe Gly Gln Gly Thr Arg Leu Glu Ile Lys  
100 105  
<210> 4  
<211> 108  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 4  
Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly  
1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Ile Thr Gly Ser  
20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu  
35 40 45

Ile Thr Gly Ala Ser Ser Trp Ala Thr Gly Ile Pro Asp Arg Phe Ser  
50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu  
65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Ser Ser Ser Pro  
85 90 95

Ile Thr Phe Gly Gln Gly Thr Arg Leu Glu Ile Lys  
100 105

<210> 5

<211> 124

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 5

[0003] Gln Val Gln Leu Val Glu Ser 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 Asn Ser Tyr  
20 25 30

Trp Ile Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val  
35 40 45

Ser Gly Ile Ala Tyr Asp Ser Ser Asn Thr Leu 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 Arg Gly Leu Gly Ala Phe His Trp Asp Met Gln Pro Asp Tyr Trp  
100 105 110

Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser  
115 120

<210> 6

<211> 122

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 6

Gln 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 Asn Ser Tyr

20 25 30

Trp Ile Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val

35 40 45

Ser Gly Ile Ala Tyr Asp Ser Ser Asn Thr Leu 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 Arg Gly Leu Gly Ala Phe His Ala Asp Met Gln Pro Asp Tyr Trp

[0004] 100 105 110

Gly Gln Gly Thr Leu Val Thr Val Ser Ser

115 120

<210> 7

<211> 107

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 7

Ser Tyr Glu Leu Thr Gln Pro Pro Ser Val Ser Val Ala Pro Gly Gln

1 5 10 15

Thr Ala Arg Ile Ser Cys Ser Gly Asp Asn Ile Gly Gly Thr Phe Val

20 25 30

Ser Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Ile Tyr

35 40 45

Asp Asp Asn Asp Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser

50 55 60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Gly Thr Gln Ala Glu

65 70 75 80

Asp Glu Ala Asp Tyr Tyr Cys Gly Thr Trp Asp Met Val Thr Asn Asn

85 90 95

Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu

100 105

<210> 8

<211> 128

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 8

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Trp Phe Gly

20 25 30

Ala Phe Thr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Gly Ile Ile Pro Ile Phe Gly Leu Thr Asn Leu Ala Gln Asn Phe

[0005]

50 55 60

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr Ser Thr Val Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys

85 90 95

Ala Arg Ser Ser Arg Ile Tyr Asp Leu Asn Pro Ser Leu Thr Ala Tyr

100 105 110

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

115 120 125

<210> 9

<211> 128

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 9

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Thr Gly

20 25 30

Ala Phe Thr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Gly Ile Ile Pro Ile Phe Gly Thr Thr Asn Leu Ala Gln Asn Phe

50 55 60

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr Ser Thr Val Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys

85 90 95

Ala Arg Ser Ser Arg Ile Tyr Asp Leu Asn Pro Ser Leu Thr Ala Tyr

100 105 110

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

115 120 125

<210> 10

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

[0006]

<400> 10

Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Ala Ala Pro Gly Gln

1 5 10 15

Lys Val Thr Ile Ser Cys Ser Gly Ser Ser Asp Ile Gly Asn Asn

20 25 30

Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu

35 40 45

Ile Tyr Asp Asn Asn Lys Arg Pro Ser Gly Ile Pro Asp Arg Phe Ser

50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Thr Leu Gly Ile Thr Gly Leu Gln

65 70 75 80

Thr Gly Asp Glu Ala Asp Tyr Tyr Cys Gly Thr Trp Asp Ser Ser Leu

85 90 95

Ser Ala Trp Val Phe Gly Gly Thr Lys Leu Thr Val Leu

100 105 110

<210> 11

<211> 123

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 11

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

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

35 40 45

Ser Gln Ile Ser Pro Ala Gly Gly Tyr Thr Asn Tyr Ala Asp Ser Val

50 55 60

Lys Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala 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 Gly Glu Leu Pro Tyr Tyr Arg Met Ser Lys Val Met Asp Val

100 105 110

[0007]

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser

115 120

<210> 12

<211> 107

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 12

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Tyr Phe Ser Ser Tyr

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

35 40 45

Tyr Gly Ala Ser Ser Arg Ala Ser Gly Val Pro Ser Arg Phe Ser Gly

50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro

65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Leu Gly Ser Pro Pro  
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys  
100 105

<210> 13

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 13

ccggaactct gccctctaac 20

<210> 14

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 14

tcctggata ccccgaagag 20

<210> 15

[0008] <211> 22

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 15

gagggcctgg atcttcttgc tc 22

<210> 16

<211> 22

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 16

gaagtagtcc ttgaccaggc ag 22

<210> 17

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 17

ccatggccca ggttcagctg 20

<210> 18  
<211> 54  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 18  
ctgtcgacc catgtaaagg cgccsnnsnn aaaggtgccg ccgcitgcct tgca 54  
<210> 19  
<211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 19  
ggcgccctta catgggtccg acag 24  
<210> 20  
<211> 51  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
[0009] <400> 20  
ctggaagttc tgggccagat tggtsnngcc gaagataggg atgatgccgc c 51  
<210> 21  
<211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 21  
accaaatctgg cccagaacctt ccag 24  
<210> 22  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 22  
actcgagacg gtgaccatttg tg 22  
<210> 23  
<211> 21  
<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 23

tttttgcga tggcccaagt g 21

<210> 24

<211> 22

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 24

aaaaaaaactc gagacgggtga cc 22

<210> 25

<211> 19

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 25

tttttgcca tggcccagg 19

<210> 26

[0010] <211> 23

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 26

aaaaaaaactc gagactgtca cgg 23

<210> 27

<211> 24

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 27

tttttgtca gcgacatcca gatg 24

<210> 28

<211> 22

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 28

tttttgcca tggcccaggt tc 22

<210> 29  
<211> 21  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 29  
tttttgcca tggcccaagt g 21  
<210> 30  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 30  
aaaaaaaactc gagacggtga cc 22  
<210> 31  
<211> 44  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
[0011] <400> 31  
ctgggcacgc cggtgtatct snngctggcg ctgtagatca gcag 44  
<210> 32  
<211> 48  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 32  
gtgccctggc caaatgtccg snnsnnagag taccgctgct ggcagtag 48  
<210> 33  
<211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 33  
tttttgcata gcgacatcca gatg 24  
<210> 34  
<211> 21  
<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 34

gctggcgctg tagatcagca g 21

<210> 35

<211> 20

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 35

agatacacccg gcgtgcccag 20

<210> 36

<211> 22

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 36

agagtaccgc tgctggcagt ag 22

<210> 37

[0012] <211> 56

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 37

aaaaaaagcgg ccgcggtaacg ctgttttcc agcttggtgc cctggccaaa tgtccg 56

<210> 38

<211> 24

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 38

tttttgcca tggcccagg ttag 24

<210> 39

<211> 26

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 39

aaaaaaaaactc gagactgtca cggtgg 26

<210> 40  
<211> 21  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 40

gtacccgggc cgcacccctcc g 21  
<210> 41  
<211> 27  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 41

cagccatggc gcctgtggag agaaagg 27  
<210> 42  
<211> 453  
<212> PRT  
<213> 人工序列(Artificial Sequence)

[0013]

<400> 42  
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  
Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val  
35 40 45  
Ser Gln Ile Ser Pro Ala Gly Gly Tyr Thr Asn Tyr Ala Asp Ser Val  
50 55 60  
Lys Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala 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 Gly Glu Leu Pro Tyr Tyr Arg Met Ser Lys Val Met Asp Val  
100 105 110  
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly  
115 120 125  
Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly

130 135 140  
Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val  
145 150 155 160  
Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe  
165 170 175  
Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val  
180 185 190  
Thr Val Pro Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val  
195 200 205  
Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys  
210 215 220  
Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu  
225 230 235 240  
Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr  
245 250 255  
Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val  
260 265 270  
[0014] Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val  
275 280 285  
Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser  
290 295 300  
Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu  
305 310 315 320  
Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala  
325 330 335  
Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro  
340 345 350  
Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln  
355 360 365  
Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala  
370 375 380  
Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr  
385 390 395 400  
Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu

405 410 415

Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser

420 425 430

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser

435 440 445

Leu Ser Pro Gly Lys

450

<210> 43

<211> 214

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 43

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Tyr Phe Ser Ser Tyr

20 25 30

[0015] Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

35 40 45

Tyr Gly Ala Ser Ser Arg Ala Ser Gly Val Pro Ser Arg Phe Ser Gly

50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro

65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Leu Gly Ser Pro Pro

85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala

100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly

115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala

130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln

145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser

165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr  
180 185 190  
Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser  
195 200 205  
Phe Asn Arg Gly Glu Cys  
210  
<210> 44  
<211> 440  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 44  
Gln Val Gln Leu Val Glu Ser Gly Gly Val Val Gln Pro Gly Arg  
1 5 10 15  
Ser Leu Arg Leu Asp Cys Lys Ala Ser Gly Ile Thr Phe Ser Asn Ser  
20 25 30  
Gly Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val  
35 40 45  
[0016] Ala Val Ile Trp Tyr Asp Gly Ser Lys Arg Tyr Tyr Ala Asp Ser Val  
50 55 60  
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Phe  
65 70 75 80  
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys  
85 90 95  
Ala Thr Asn Asp Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser  
100 105 110  
Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Ala Pro Cys Ser  
115 120 125  
Arg Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp  
130 135 140  
Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr  
145 150 155 160  
Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr  
165 170 175  
Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Lys

180 185 190  
Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp  
195 200 205  
Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala  
210 215 220  
Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro  
225 230 235 240  
Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val  
245 250 255  
Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val  
260 265 270  
Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln  
275 280 285  
Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln  
290 295 300  
Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly  
305 310 315 320  
[0017] Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro  
325 330 335  
Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr  
340 345 350  
Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser  
355 360 365  
Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr  
370 375 380  
Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr  
385 390 395 400  
Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe  
405 410 415  
Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys  
420 425 430  
Ser Leu Ser Leu Ser Leu Gly Lys  
435 440  
<210> 45

<211> 214

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 45

Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly

1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Tyr

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile

35 40 45

Tyr Asp Ala Ser Asn Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly

50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro

65 70 75 80

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Ser Ser Asn Trp Pro Arg

85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala

100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly

115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala

130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln

145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser

165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr

180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser

195 200 205

Phe Asn Arg Gly Glu Cys

210

<210> 46

<211> 114

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 46

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg

1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr

20 25 30

Gly Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val

35 40 45

Ala Phe Ile Arg Tyr Asp Gly Ser Asn Lys 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

Lys Thr His Gly Ser His Asp Asn Trp Gly Gln Gly Thr Met Val Thr

100 105 110

[0019] Val Ser

<210> 47

<211> 114

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 47

Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Gly Ala Pro Gly Gln

1 5 10 15

Arg Val Thr Ile Ser Cys Ser Gly Ser Arg Ser Asn Ile Gly Ser Asn

20 25 30

Thr Val Lys Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu

35 40 45

Ile Tyr Tyr Asn Asp Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser

50 55 60

Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Thr Gly Leu Gln

65 70 75 80

Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Ser Tyr Asp Arg Tyr Thr

85 90 95

His Pro Ala Leu Leu Phe Gly Thr Gly Thr Lys Val Thr Val Leu Gly

100 105 110

Gln Pro

<210> 48

<211> 118

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 48

Glu Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Glu

1 5 10 15

Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Thr Tyr

20 25 30

Trp Leu Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Asp Trp Ile

35 40 45

Gly Ile Met Ser Pro Val Asp Ser Asp Ile Arg Tyr Ser Pro Ser Phe

50 55 60

[0020] Gln Gly Gln Val Thr Met Ser Val Asp Lys Ser Ile Thr Thr Ala Tyr

65 70 75 80

Leu Gln Trp Asn Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys

85 90 95

Ala Arg Arg Arg Pro Gly Gln Gly Tyr Phe Asp Phe Trp Gly Gln Gly

100 105 110

Thr Leu Val Thr Val Ser

115

<210> 49

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 49

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Trp

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Glu Lys Ala Pro Lys Ser Leu Ile  
35 40 45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Tyr  
85 90 95

Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr Ala  
100 105 110

<210> 50

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 50

Asp Ser Arg Pro Pro Tyr Tyr Gly Met Asp Val  
1 5 10

[0021] <210> 51

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 51

Met Gln Ala Leu Gln Thr Pro Pro Thr  
1 5

<210> 52

<211> 8

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 52

Asp Gly Arg Gly Gly Phe Asp Tyr  
1 5

<210> 53

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 53

Gln Gln Phe Asn Ser Tyr His Leu Leu Thr

1 5 10

<210> 54

<211> 12

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 54

Gly Arg Ser Ser Val Ile Asp Tyr Gly Met Asp Val

1 5 10

<210> 55

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 55

Asn Ser Arg Asp Ser Ser Gly Asn His Val Val

[0022]

1 5 10

<210> 56

<211> 15

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 56

Asp Arg Val Ala Ala Thr His Tyr Tyr Tyr Gly Met Asp Val

1 5 10 15

<210> 57

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 57

Gln Gln Ser Tyr Gly Ser Pro Phe Thr

1 5

<210> 58

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 58

Gly Leu Leu Glu Lys Gly Ala Phe Asp Ile

1 5 10

<210> 59

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 59

Gln Gln Ser Tyr Ser Thr Pro Gln Thr

1 5

<210> 60

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

[0023] <400> 60

Asp Ser Arg Pro Pro Tyr Tyr Gly Met Asp Val

1 5 10

<210> 61

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 61

Gln Gln Ser Tyr Asn Ser Arg Pro Tyr Thr

1 5 10

<210> 62

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 62

Thr Ser Pro Tyr Ser Gly Ser Tyr Asn Asn

1 5 10

<210> 63

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 63

Ser Ser Tyr Gly Gly Asn Tyr Lys Tyr Leu

1 5 10

<210> 64

<211> 13

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 64

Gly Leu Ser Ser Thr Trp Ala Gly Gly Ala Phe Asp Ile

1 5 10

<210> 65

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

[0024]

<400> 65

Asn Ser Arg Asp Ser Ser Gly Asn His Val Val

1 5 10

<210> 66

<211> 12

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 66

Gly Ala His Ser Gly Tyr Asp Ser Asp Phe Asp Tyr

1 5 10

<210> 67

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 67

Gln Gln Tyr Asn Ser Tyr Pro Leu Thr

1 5

<210> 68

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 68

Val Ser Gly Ser Ser Asn His Ala Phe Asp Ile

1 5 10

<210> 69

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 69

His Gln Ser Ser Ser Phe Pro Leu Thr

1 5

<210> 70

<211> 13

<212> PRT

[0025]

<213> 人工序列(Artificial Sequence)

<400> 70

Asp Thr Ser Ser Arg Tyr Ala Gly Gly Ala Phe Asp Ile

1 5 10

<210> 71

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 71

Asn Ser Arg Asp Ser Ser Gly Asn His Val Val

1 5 10

<210> 72

<211> 13

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 72

Asp Thr Gly Ser Ser Ala Arg Gly Gly Asp Phe Asp Tyr

1 5 10  
<210> 73  
<211> 11  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 73

Asn Ser Arg Asp Ser Ser Gly Asn His Val Val

1 5 10  
<210> 74  
<211> 10  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 74

Asp Glu Arg Tyr Tyr Gly Asp Met Asp Val

1 5 10  
<210> 75  
<211> 9  
[0026] <212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 75

Met Gln Gly Leu Gln Thr Pro Arg Thr

1 5  
<210> 76  
<211> 13  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 76

Asp Tyr Ser Ser Gly Trp Ser Gly Asp Ala Phe Asp Ile

1 5 10  
<210> 77  
<211> 11  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 77

Asn Ser Arg Asp Ser Ser Gly Asn His Val Val

1 5 10

<210> 78

<211> 13

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 78

Arg Gly Thr Arg Gly Tyr Tyr Tyr Gly Met Asp Val

1 5 10

<210> 79

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 79

Gln Gln Ser Tyr Asn Ser Arg Pro Tyr Thr

1 5 10

<210> 80

[0027]

<211> 8

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 80

Thr Asn Asn Gly Phe Ile Asp Tyr

1 5

<210> 81

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 81

Ser Gln Ala Ser His Trp Pro Tyr Thr

1 5

<210> 82

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 82

Gly Arg Gly Arg Arg Tyr Phe Asp Leu

1 5

<210> 83

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 83

Gln Gln Ser Tyr Ser Thr Pro Arg Thr

1 5

<210> 84

<211> 12

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 84

Asp Gly Asp Glu Gly Glu Leu Gly Ala Phe Asp Ile

1 5 10

[0028]

<210> 85

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 85

Gln Gln Ser Tyr Gly Ser Pro Phe Thr

1 5

<210> 86

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 86

Gly Asp Asn Asn Tyr Tyr Phe Asp Tyr

1 5

<210> 87

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 87

Gln Gln Tyr Tyr Ser Arg Pro Ile Thr

1 5

<210> 88

<211> 12

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 88

Glu Thr Gly Glu Gly Arg Trp Glu Leu Leu Gly Tyr

1 5 10

<210> 89

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 89

[0029] Met Gln Gly Thr His Trp Pro Arg Thr

1 5

<210> 90

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 90

Leu Ser His Thr Ala Pro Leu Val Asp Tyr

1 5 10

<210> 91

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 91

Gln Gln Tyr Ser Lys Trp Pro Leu Thr

1 5

<210> 92

<211> 13

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 92

Ala Ile Ala Pro Arg Arg Tyr Tyr Gly Met Asp Val

1 5 10

<210> 93

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 93

Gln Gln Ser Tyr Ser Thr Pro Arg Thr

1 5

<210> 94

<211> 13

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 94

[0030]

Ala Ile Ala Pro Arg Arg Tyr Tyr Gly Met Asp Val

1 5 10

<210> 95

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 95

Gln Gln Tyr Asn Ser Tyr Pro Leu Thr

1 5

<210> 96

<211> 7

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 96

Asp Gly Tyr Asn Ser Asp Tyr

1 5

<210> 97

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 97

Gln Gln Tyr Tyr Ser Lys Pro Leu Thr

1 5

<210> 98

<211> 12

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 98

Arg Arg Tyr Asn Trp Asp Tyr Asp Tyr Val Asp Val

1 5 10

<210> 99

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

[0031]

<400> 99

Gln Gln Ser Tyr Ser Thr Pro Arg Thr

1 5

<210> 100

<211> 14

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 100

Asp Lys Pro Val Gly Ser Ser Gly Trp Tyr Pro Phe Asp Tyr

1 5 10

<210> 101

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 101

Gln Gln Ser Tyr Ser Thr Pro Tyr Thr

1 5

<210> 102  
<211> 10  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 102  
Gln Ile Asn Trp Gly Tyr Tyr Phe Asp Tyr  
1 5 10  
<210> 103  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 103  
Met Gln Ala Leu Gln Thr Pro Pro Thr  
1 5  
<210> 104  
<211> 18  
<212> DNA  
[0032] <213> 人工序列(Artificial Sequence)  
<400> 104  
ggtgctcgac tccaccaa 18  
<210> 105  
<211> 21  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 105  
gatggaaaggtt gccatgaaag a 21  
<210> 106  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 106  
tcttgtattg ccgggatcct tc 22  
<210> 107  
<211> 22

<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 107  
taactccca g ccctac ctac tc 22  
<210> 108  
<211> 30  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 108  
ctccacccat ac caccatgg actataattt 30  
<210> 109  
<211> 36  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 109  
tttttatgc atcttatgcc agctttgga tgacgg 36  
[0033] <210> 110  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 110  
ctcctctgag tctagccagg cc 22  
<210> 111  
<211> 30  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 111  
caaataatgt ccatgagggtg gtaggtggag 30  
<210> 112  
<211> 38  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 112

tttttgcgta tcgcgatggc ttacatcccg tgccttc 38  
<210> 113  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 113  
ggcctggcta gactcagagg ag 22  
<210> 114  
<211> 44  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 114  
tatattgtat acggcgcgcc tgtcagggac aagattagtc acag 44  
<210> 115  
<211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
[0034] <400> 115  
gactttggtg ataatgtgag cagc 24  
<210> 116  
<211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 116  
gctgctcaca ttatcaccaa agtc 24  
<210> 117  
<211> 35  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 117  
tatattccctg caggctccctg caaaggccctg aagag 35  
<210> 118  
<211> 38  
<212> DNA

<213> 人工序列(Artificial Sequence)  
<400> 118  
tttttatgc atcttcatga cggggagata aaagcata 38  
<210> 119  
<211> 44  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 119  
tatattgtat acggcgccgccc tgtcaggac aagattagtc acag 44  
<210> 120  
<211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 120  
gactttggtg ataatgttag cagc 24  
<210> 121  
[0035] <211> 24  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 121  
gctgctcaca ttatcaccaa agtc 24  
<210> 122  
<211> 35  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 122  
tatattcctg caggctcctg caaaggccctg aagag 35  
<210> 123  
<211> 38  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 123  
tttttatgc atcttcatga cggggagata aaagcata 38

<210> 124  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 124  
ggaatcatgg gaaataggcc ct 22  
<210> 125  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 125  
cgctcacaat tccacacaac at 22  
<210> 126  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
[0036] <400> 126  
agggcctatt tcccatgatt cc 22  
<210> 127  
<211> 22  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 127  
atgttgttg gaattgtgag cg 22  
<210> 128  
<211> 18  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 128  
tccccgtcat ccaaaagc 18  
<210> 129  
<211> 18  
<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 129

tctgctgtga ctaatctt 18

<210> 130

<211> 448

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 130

agggcctatt tcccatgatt cttcatatt tgcatatacg atacaaggct gtttagagaga 60

taattagaat taatttgact gtaaacacaa agatattatg acaaaaatcg tgacgtagaa 120

agtaataatt tcttggtag ttgcagttt taaaattatg tttaaaatg gactatcata 180

tgcttaccgt aacttgaag tatttcgatt tcttggctt atatatctt gggaaaggac 240

gaaacaccga tccaaaagct ggcattgtcg ttttagagct agaaatagca agttaaaata 300

aggcttagtcc gttatcaact taaaaaagtgc gcaccgagtc ggtgccttt tctatgtatac 360

cgtcgacctc tagctagac ttggcgtaat catggcata gctgttcct gtgtgaaatt 420

gttatecgct cacaattcca cacaacat 448

<210> 131

[0037] <211> 448

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 131

agggcctatt tcccatgatt cttcatatt tgcatatacg atacaaggct gtttagagaga 60

taattagaat taatttgact gtaaacacaa agatattatg acaaaaatcg tgacgtagaa 120

agtaataatt tcttggtag ttgcagttt taaaattatg tttaaaatg gactatcata 180

tgcttaccgt aacttgaag tatttcgatt tcttggctt atatatctt gggaaaggac 240

gaaacaccgt ctccccgtca tccaaaageg ttttagagct agaaatagca agttaaaata 300

aggcttagtcc gttatcaact taaaaaagtgc gcaccgagtc ggtgccttt tctatgtatac 360

cgtcgacctc tagctagac ttggcgtaat catggcata gctgttcct gtgtgaaatt 420

gttatecgct cacaattcca cacaacat 448

<210> 132

<211> 448

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 132

agggcctatt tcccatgatt cttcatatt tgcatatacg atacaaggct gtagagaga 60  
taattagaat taatttgact gtaaacacaa agatattgt acaaatacg tgacgtagaa 120  
agtaataatt tcttgggtag ttgcagttt taaaattatg ttttaaaatg gactatcata 180  
tgcttaccgt aacctgaaag tatttcgatt tcttggctt atatatcttg tggaaaggac 240  
gaaacaccga tgccagctt tggatgacgg ttttagagct agaaatagca agttaaaata 300  
aggcttagcc gtatcaact taaaaaagtgc acaccgagtc ggtgccttt tctagtatac 360  
cgtcgaccc tc tagctagac ttggcgtaat catggcata gctgttcct gtgtgaaatt 420  
gttatccgct cacaattcca cacaacat 448  
<210> 133  
<211> 37  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 133  
ctttctctcc acaggcgccc atggccgaag tgcagcc 37  
<210> 134  
<211> 28  
<212> DNA  
[0038] <213> 人工序列(Artificial Sequence)  
<400> 134  
tttttctcg agacggtgac cagggttc 28  
<210> 135  
<211> 28  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 135  
tttttgcta gtcctatga gtcactc 28  
<210> 136  
<211> 32  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 136  
gtcacgctt gtcggccgc gggctgacct ag 32  
<210> 137  
<211> 18

<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 137  
ggccgcacca agcgtgac 18  
<210> 138  
<211> 19  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 138  
ggcgccctgtg gagagaaaag 19  
<210> 139  
<211> 28  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 139  
tttttgcta gtccttatga gtcgtactc 28  
<210> 140  
[0039] <211> 28  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 140  
tttttctcg agacggtgac cagggttc 28  
<210> 141  
<211> 899  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 141  
taataagcta gcagaggaga catccagatg acacagagcc ctgcgcgt gtctgccage 60  
gtgggagaca gagtgaccat cacatgcaag gccagccgga acatcgagag acagctggcc 120  
tgttatcgc agaaggctgg acaggctct gagctgctga tctatcaggc cagcagaaaa 180  
gaaagcggcg tgcccgatag attcagccggc agcagatacg gcacggactt caccctgaca 240  
atatccagcc tccagccctga ggatatcgcc acctactact gccagcagta cagcgaccct 300  
ccactgacat ttggcggagg caccaagggtg gaaatcaagc ggacagccggc cgccccctagc 360  
gtttcatct ttccacccatcg cgacgagcagc ctgaagtctg gcacagccctc tgtcggtgc 420

ctgctgaaca acttctaccc cagagaagcc aagggtcagt ggaagggtgga caacgccctc 480  
 cagagccgca atagccaaga gagcgtgacc gagcaggaca gcaaggactc tacctacagc 540  
 ctgagcagca cactgaccct gagcaaggcc gactacgaga agcacaaagt gtacgcctgc 600  
 gaagtgaccc accagggcct ttctagccct gtgaccaaga gcttcaaccg gggcgaatgt 660  
 taataatcta gagectcgac tgtgccttc agttgccagc catctgtgt ttgccttc 720  
 cccgtgcctt ccttgaccct ggaagggtgcc actcccactg tccttccta ataaaatgag 780  
 gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag 840  
 gacagcaagg gggaggattg ggaagacaat agcaggcatg ctgggagaaga tcttaataa 899  
 <210> 142  
 <211> 1557  
 <212> DNA  
 <213> 人工序列(Artificial Sequence)  
 <400> 142

ttattagcca tggcccaggt gcagctgggt gaatctggcg gaggactggc ttagcctggc 60  
 ggatctctga gactgtctg tgccgccagc ggcttcacct tcagctacta cgatatccag 120  
 tgggtccgac aggcccctgg caaaggactt gaatgggtgt ccagcatcg cccctctggc 180  
 cagttcacct actaccggcg agaagtgaag ggcagattca ccatcagccg ggacaacagc 240  
 [0040] aagaacaccc tgtacctgca gatgaacacgc ctgagagccg aggacaccgc cgtgtactac 300  
 tggccagaa gaaccggcag agagtacggc ggaggctgg actttgatta ctggggccag 360  
 ggcacccctgg tcacagtctc gagcgcctct acaaaggccc ccagcgttt cccactggct 420  
 ccctgttagca gaagcaccag cgaatctaca gccgcctctgg gctgcctggt caaggactac 480  
 ttccctgagc ctgtgaccgt gtcctggaaac tctggcgctc tgacatctgg cgtgcacacc 540  
 ttccagccg tgctgcaaag cagccgcctg tacagtcgtc gcagcgtcgt gacagtgcct 600  
 agcagctctc tgggcaccca gacctacacc tgtaatgtgg accacaagcc tagcaacacc 660  
 aaggtggaca agcgcgtgaa atctaagtac ggcctccctt gtcctccatg tccctgcacct 720  
 gagtttctcg gcggaccctc cgtgtccctg ttccctccaa agcctaagga caccctgatg 780  
 atctccagaa cacccgaagt gacctgcgtg gtggggacg tttcacaaga ggaccccgag 840  
 gtgcagtta attggtagt ggacggcggt gaagtgcaca acgccaagac caagcctaga 900  
 gaggaacagt acaacagcac ctacagagt gtgcctgc tgacagtgc gcaccaggat 960  
 tggctgaacg gcaaagatgta caagtgcacg gtgcacaaca agggctgc aagcagcatc 1020  
 gagaaaacca tcagcaaggc caagggccag cctaggaaac cccaggatca cacactgcct 1080  
 ccaagccaga aagagatgac caagaaccag gtgcctgc cctgcctcgta gaagggttc 1140  
 tacccttcgg atatgcctgtt ggaatgggag agcaatggcc agccagagaa caactacaag 1200  
 accacaccc tcgtgctgga cagcgacggc tcattctcc tgtacagcaa gctgaccgtg 1260

gacaagagca gatggcaaga gggcaacgtg tttagctgca gcgtgatgca cgaggccctg 1320  
cacaacagat acacccagaa gtccctgtct ctgagccccg aacaaaact catctcgaa 1380  
gaggatctga atgcgttggg ccaggacacg caggaggta tcgtggtgcc acactcttg 1440  
cccttaagg tggtggtgat ctccggatc ctggccctgg tggtgctcac catcatctcc 1500  
cttatcatcc tcatcatgtt ttggcagaag aagccacgtt agtaaaagct tttatta 1557  
<210> 143  
<211> 2152  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 143  
ttattagaat tcaacatgga ctggaccctgg agggtcttc gcttgcgtgc ttttagctcca 60  
ggtaaaggc caactggttc caggcgtgag gaaggattt ttccagttt agaggactgt 120  
cattctctac tgggtccctc ccgcagggtgc tcactccag gttcagctgg tgcaaggctt 180  
cagcgagctg aaaaaacctg ggcgcctccgt gaagggtgtcc tgcaaggctt ctggctacac 240  
ctttaccgac aacaacatgg actgggtccg acaggcccct ggacaaggac ttgagtggat 300  
ggcgacatc aacaccagaa gggcgccag catctacaac gaaggttcc aggacagagt 360  
catcatgacc gtggacaaga gcaccgacac cgcctacatg gaactgagca gcctgagaag 420  
cgaggacacc gccaccatc actggccag aagaaagagc tacggctact acctggacga 480  
gtggggcgag ggaacactgg tcacagtgtc taggcgcagc acaaaggccc ctagegttt 540  
cccaactggct ccctgttagca gaagcaccag cgaatctaca ggcgcctgg gtcgcctgt 600  
gaaggactac ttccctgagc ctgtgaccgt tagctggaaac agcggaggcac tgacaagcgg 660  
cgtgcacaca ttccctggcc tgctgcaag cagcggccctg tactctctga gcaagcgtcgt 720  
gacagtgcct agcagcttc tgggcaccca gacctacacc tgtaatgtgg accacaagcc 780  
tagcaacacc aaggtggaca agcgcgtgaa atctaagtac ggcgcctt gtcctccatg 840  
tcctgttccaa gagtttctcg gggaccctc cgtgttccctg ttccctccaa agcctaagga 900  
caccctgtatc atctccagaa cacccgaagt gacccgtg ggggtggacg ttccataaga 960  
ggaccccgag gtgcagtca attggtaatcgt ggacggccgtg gaagtgcaca acgccaagac 1020  
caagccctaga gaggaacagt acaacagcac ctacagagtg gtgtccgtgc tgacagtgt 1080  
gcaccaggat tggctgaacg gcaagagta caagtgcacg gtgtccaaaca agggccctg 1140  
aagcagcatc gagaaaacca tcagcaaggc caagggccag cctaggaaac cccagggtta 1200  
cacactgctt ccaagccaa agggaaatgac caagaaccag gtgtccctga cctgectgtt 1260  
caaggcccttcc tacccttccg atatcgccgt ggaatgggag agcaatggcc agccagagaa 1320  
caactacaag accacacccctc ctgtgctgaa cagcgcacggc tcattctcc tgtacagcaa 1380  
gtgtactgtg gataagagcc ggtggcaaga gggcaacgtg tttagctgta gcgtgatgca 1440

cgaggccctg cacaaccact acacccaaga gaggcgtct ctgagccctg aacaaaaact 1500  
catctcagaa gaggatctga atgctgtggg ccaggacacg caggaggtca tcgtgggcc 1560  
acactcccttgc ccccttaagg tgggtgtat ctcagccatc ctggccctgg tggctcac 1620  
catcatctcc cttatcatcc tcatcatgtct ttggcagaag aagccacgtt agtaactaag 1680  
tcgacatcca gacatgataa gatacatgttg tgagtttggaa caaaccacaa ctagaatgca 1740  
gtgaaaaaaaaa tgctttattt gtgaaatttg tgatgctatt gctttatttga taaccattat 1800  
aagctgcaat aaacaagttt acaacaacaa ttgcattcat ttatgtttc aggttcaggg 1860  
ggaggtgtgg gagggtttttaa aagcaagta aaacctctac aatgtggta tggctgatta 1920  
tgatccctgca agcctcgatcg tcctggccgg accacgttat ctgtcaagg tccccggccc 1980  
cggaacgcgcg ctccatgagc agagcgcccg ccggcgaggc gaagactcgg gcggcgccct 2040  
ggccgtccca ccaggtcaac aggccgttaac cgcccttc atcgggaatg cgcgacacct 2100  
tcagcatcgc cggcatgtcc cccctggcgga cggaaatgtat gtatacttat ta 2152  
<210> 144  
<211> 389  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 144

[0042] <400> 144  
ttattagcca tggcccagggt gcagctggtt gaatctggcg gaggactggt tcagcctggc 60  
ggatctctga gactgtcttg tgccgccagc ggcttcacct tcagctacta cgatatccag 120  
tgggtccgac aggcccctgg caaaggactt gaatgggtgt ccagcatcag cccctctggc 180  
cagtccacct actaccggcg agaagtgaag ggcagattca ccatcagccg ggacaacacgc 240  
aagaacaccc tgtacctgca gatgaacagc ctgagagccg aggacaccgc cgtgtactac 300  
tgcgccagaa gaaccggcag agagtacggc ggaggctggt actttgatta ctggggccag 360  
ggcacccctgg tcacagtctc gagttatta 389  
<210> 145  
<211> 353  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 145  
ttattagcta gcgacatcca gatgacacag agcccttagca gcctgtctgc cagcgtggga 60  
gacagagtga ccatcacatg caaggccagc cggaacatcg agagacagct ggcctggat 120  
cagcagaagc ctggacaggc tcctgagctg ctgtatctac aggccagcag aaaagaaaagc 180  
ggcgtgcccc atagattcag cggcagcaga tacggcaccc acttcacct gacaataatcc 240  
agecctccagc ctgaggatata cgccacccatc tactgccagc agtacagcga ccctccactg 300

acatttggcg gaggcaccaa ggtggaaatc aagcggacag cggccgctta tta 353

<210> 146

<211> 353

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 146

tttagcttgcgacatcca gatgacacag agccctagca gcctgtctgc cagcgtggga 60

gacagagtga ccatcacatg caaggccagc cgaaacatct atagacagct ggcctggat 120

cagcagaagc ctggacaggc tcctgagctg ctgatctatc aggcagcag aaaatata 180

ggcgtgccc atagattcag cggcagcaga tacggcacccg acttcaccct gacaatatcc 240

agcctccagc ctgaggatat cgccacccat tactgccagc agtacagcga ccctccactg 300

acatttggcg gaggcaccaa ggtggaaatc aagcggacag cggccgctta tta 353

<210> 147

<211> 353

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 147

[0043]

tttagcttgcgacatcca gatgacacag agccctagca gcctgtctgc cagcgtggga 60

gacagagtga ccatcacatg caaggccagc cgaaacatct atagacagct ggcctggat 120

cagcagaagc ctggacaggc tcctgagctg ctgatctatc aggcagcag aaaagaaagc 180

ggcgtgccc atagattcag cggcagcaga tacggcacccg acttcaccct gacaatatcc 240

agcctccagc ctgaggatat cgccacccat tactgccagc agtacagcga ccctccactg 300

acatttggcg gaggcaccaa ggtggaaatc aagcggacag cggccgctta tta 353

<210> 148

<211> 353

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 148

tttagcttgcgacatcca gatgacacag agccctagca gcctgtctgc cagcgtggga 60

gacagagtga ccatcacatg caaggccagc cgaaacatct atagacagct ggcctggat 120

cagcagaagc ctggacaggc tcctgagctg ctgatctatc aggcagcag aaaatata 180

ggcgtgccc atagattcag cggcagcaga tacggcacccg acttcaccct gacaatatcc 240

agcctccagc ctgaggatat cgccacccat tactgccagc agtacagcag ccctccactg 300

acatttggcg gaggcaccaa ggtggaaatc aagcggacag cggccgctta tta 353

<210> 149

<211> 353

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 149

tttagtgcata gggacatcca gatgacacag agccctagca gcctgtctgc cagcgtggga 60  
gacagagtga ccatcacatg caaggccagc cgaaacatct atagacagct ggctggtat 120  
cagcagaagc ctggacaggc tcctgagctg ctgatctatc agggcagcag aagatatacg 180  
ggcgtgcccgg atagattcag cggcagcaga tacggcaccg acttcaccct gacaatatcc 240  
agcctccagc ctgaggatata cgccacccat tactgccagc agtacagcag ccctccactg 300  
acatttggcg gaggcaccaa ggtggaaatc aaggccacag cggccgccta tta 353

<210> 150

<211> 138

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 150

[0044] Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Glu Ala Pro Arg Gln

1 5 10 15

Arg Val Thr Ile Thr Cys Ser Gly Ser Ser Asn Ile Gly Asn Asn

20 25 30

Ala Val Asn Trp Tyr Gln Gln Leu Pro Gly Lys Ala Pro Lys Leu

35 40 45

Ile Tyr Ala Ala Gly Arg Cys Pro Arg Ile Leu Met Arg Cys Lys Gln

50 55 60

Asp Ser Asp Cys Leu Ala Gly Cys Val Cys Gly Pro Asn Gly Phe Cys

65 70 75 80

Gly Ala Asn Ser Gly Val Ser Asp Arg Phe Ser Ala Ala Lys Ser Gly

85 90 95

Thr Ser Ala Ser Leu Ala Ile Asn Gly Leu Arg Ser Glu Asp Glu Ala

100 105 110

Asp Tyr Tyr Cys Ala Ala Trp Asp Asp Ser Leu Asn Gly Tyr Val Phe

115 120 125

Gly Thr Gly Thr Lys Leu Thr Val Leu Gly

130 135

<210> 151

<211> 145

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 151

Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Glu Ala Pro Arg Gln

1 5 10 15

Arg Val Thr Ile Thr Cys Ser Gly Ser Ser Asn Ile Gly Asn Asn

20 25 30

Ala Val Asn Trp Tyr Gln Gln Leu Pro Gly Lys Ala Pro Lys Leu Leu

35 40 45

Ile Tyr Ala Ala Gly Arg Gly Cys Leu Lys Phe Gly Trp Lys Cys Asn

50 55 60

Pro Arg Asn Asp Lys Cys Cys Ser Gly Leu Lys Cys Gly Ser Asn His

65 70 75 80

Asn Trp Cys Lys Trp His Ile Gly Ala Asn Ser Gly Val Ser Asp Arg

85 90 95

[0045]

Phe Ser Ala Ala Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Asn Gly

100 105 110

Leu Arg Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala Trp Asp Asp

115 120 125

Ser Leu Asn Gly Tyr Val Phe Gly Thr Gly Thr Lys Leu Thr Val Leu

130 135 140

Gly

145

<210> 152

<211> 145

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 152

Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Glu Ala Pro Arg Gln

1 5 10 15

Arg Val Thr Ile Thr Cys Ser Gly Ser Ser Asn Ile Gly Asn Asn

20 25 30

Ala Val Asn Trp Tyr Gln Gln Leu Pro Gly Lys Ala Pro Lys Leu Leu  
35 40 45

Ile Tyr Ala Ala Gly Arg Ala Ser Cys Arg Thr Pro Lys Asp Cys Ala  
50 55 60

Asp Pro Cys Arg Lys Glu Thr Gly Cys Pro Tyr Gly Lys Cys Met Asn  
65 70 75 80

Arg Lys Cys Lys Cys Asn Arg Cys Ala Asn Ser Gly Val Ser Asp Arg  
85 90 95

Phe Ser Ala Ala Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Asn Gly  
100 105 110

Leu Arg Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala Trp Asp Asp  
115 120 125

Ser Leu Asn Gly Tyr Val Phe Gly Thr Gly Thr Lys Leu Thr Val Leu  
130 135 140

Gly  
145

<210> 153

[0046] <211> 33

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 153

Gly Cys Leu Lys Phe Gly Trp Lys Cys Asn Pro Arg Asn Asp Lys Cys  
1 5 10 15

Cys Ser Gly Leu Lys Cys Gly Ser Asn His Asn Trp Cys Lys Trp His  
20 25 30

Ile

<210> 154

<211> 33

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 154

Gly Cys Xaa Xaa Xaa Xaa Xaa Lys Cys Asn Pro Arg Asn Asp Lys Cys  
1 5 10 15

Cys Ser Gly Leu Lys Cys Gly Ser Asn His Asn Trp Cys Lys Trp His

20 25 30

Ile

<210> 155

<211> 33

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 155

Gly Cys Xaa Xaa Xaa Xaa Xaa Lys Cys Asn Pro Arg Asn Asp Lys Cys

1 5 10 15

Cys Ser Gly Leu Xaa Cys Gly Ser Asn His Asn Trp Cys Lys Xaa Xaa

20 25 30

Xaa

<210> 156

<211> 34

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 156

[0047]

Ala Ser Cys Arg Thr Pro Lys Asp Cys Ala Asp Pro Cys Arg Lys Glu

1 5 10 15

Thr Gly Cys Pro Tyr Gly Lys Cys Met Asn Arg Lys Cys Lys Cys Asn

20 25 30

Arg Cys

<210> 157

<211> 34

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 157

Ala Ser Cys Arg Thr Pro Lys Asp Cys Ala Asp Pro Cys Xaa Xaa Glu

1 5 10 15

Thr Gly Cys Pro Tyr Gly Lys Cys Xaa Asn Arg Xaa Cys Lys Cys Asn

20 25 30

Arg Cys

<210> 158

<211> 30

<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 158  
acaggggtcc gcgcagtctt taggagtcg 30  
<210> 159  
<211> 96  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 159  
cctaaagact gcgcggaccc ctgtvnsvns gagactggat gtccatacgg taagtgcvns 60  
aatagavnst gcaa atgtaa ccgatgcgca aacagt 96  
<210> 160  
<211> 29  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 160  
catcccccttc ccgcgtcgta aatgagaag 29  
[0048] <210> 161  
<211> 66  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 161  
ctcatttacg cagcgggaag gggatgcvns vns vns vns sgaatgcaa cccaagaaac 60  
gataaa 66  
<210> 162  
<211> 81  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 162  
gagtcttgag cagcatttat cgttcttgg gttgcatttc sns nbs nbsn bsn bgcatcc 60  
ccttcccgt gcgtaaatga g 81  
<210> 163  
<211> 93  
<212> DNA



tttatctgt cccctccacc ccacagtggg gcaagatgca tcttctgacc tcttccttc 840  
ctcccacagg gcatggcaaa acctctgagc caggaagaaa gcacactgt tgaaagagca 900  
accgctacta tcaacagcat ccccatctcc gaagactatt ctgtggctag tgccgcctcg 960  
tccagcgacg ggagaatctt caccgggttg aacgtctacc actttacagg cggaccatgc 1020  
gcagagctgg tggcctggg gactgcagcc gctgcagecg ctggtaatct gacctgtate 1080  
gtggccattg gcaacgaaaa tagggcatac ctgtccccat gcggcagggt tcggcagggt 1140  
ctgctggatc tgcatctgg catcaaggca attgtcaaaag actctgtatgg acagcctacc 1200  
gccgtcggtt tccgtgaact gctgccttagc ggctatgtct gggagggata atgagcttgg 1260  
cttcgaaatg accgaccaag cgacgccccaa cctgccatca cgagatttcg attccaccgc 1320  
cgccctctat gaaagggtgg gcttcggat cggttccgg gacgccccgtt ggtatgtatcc 1380  
ccagcgcggg gatctcatgc tggagttctt cggccaccccc aacttgtta ttgcagctta 1440  
taatggttac aaataaaagca atagcatcac aaatttcaca aataaagcat tttttcact 1500  
gcattctagt tgtggagatc tcgtgaggct ccggtgcccc tcagtgggca gagcgcacat 1560  
cccccacagt ccccgagaag ttggggggag gggtcggcaa ttgaaccggc gcctagagaa 1620  
ggtggcgcgg ggtaaaactgg gaaagtgtatc tcgtgtactg gctccgcctt ttcccggagg 1680  
gtgggggaga accgtatata agtgcactag tcgcccgtaa cgttctttt cgcaacgggt 1740  
ttgccgcccag aacacaggta agtgcgtgt gtggtcccc cgggcctggc ctcttacgg 1800  
gttatggccc ttgcgtgcct tgaattactt ccacctggct gcagtagctg attctgtatc 1860  
ccgagcttcg ggttggaaagt gggtgggaga gttcgtggcc ttgcgtttaa ggagccccctt 1920  
cgccctgtgc ttgagttgtg gcctggccctg ggcgcgtggg cggccgcgtg cgaatctgg 1980  
ggcaccttcg cgcctgtctc gtcgtttcg ataagtctct agccattaa aattttgtat 2040  
gacctgcgtgc gacgttttt ttcgttgcag atagcttgtt aatgcgggc caagatcagc 2100  
acactggat ttcggttttt gggccgcgg gggcgcacgg gggccgtgcg tcccaagcga 2160  
catgttcggc gaggcggggc ctgcgagcgc ggcacccgg aatcggacgg gggtagtctc 2220  
aagctggccg gctgtctcg gtgcctggcc tcgcgcggcc gtgtatgcgc ccgcctggg 2280  
cgcaaggctt gggccggcgtc gcaccagggt cgtgagcgg aagatggccg tttccggcc 2340  
ctgctgcagg gaccaaaaa tggaggacgc ggcgcgtggg agagcggggc ggttagtctc 2400  
ccacacaaag gaaaaggccc ttccgttct cagccgtcgc ttcatgtgac tccacggagt 2460  
accggggcgc gtcaggcac ctgcattgt ttcgcgtt tttggatgtc tcttccttt 2520  
gttgggggggaa ggggtttat gcgatggagt ttccccacac tgagtgggtg gagactgaag 2580  
ttaggccagc ttggcacttg atgttaattct ccttggaaatt tggccctttt gagttggat 2640  
tttgggttcat tctcaaggctt cagacagtgg ttcaaagttt ttttctcca ttccagggtgt 2700  
cgtaaaaact accccctaaaaa gccaaaagat cggcgtggc caccatgagg gcctggatct 2760  
ttttctctt ttcgtggcc gggagggttc tggcagctag cgacatccag atgacccaga 2820



[0052] ggctgcctgg tgaaggacta cttccctgag cctgtgaccg tgtcctggaa ctctggcgcc 4920  
 ctgacccctcg gcggtgcatac ctccctgcc gtcctccagt cctccggct gtactccctg 4980  
 tcctccgtgg tgaccgtgcc ttccctctt ctgggcaccc agacctacat ctgcaacgtg 5040  
 aaccacaaggc ctccaacac caagggtggac aagaagggtgg agccctaagtc ctgcgacaag 5100  
 acccacaccc gcccctcatg tcctgcctt gagctgtgg gcggaccctc cggttcctg 5160  
 ttccctctta agccataagga caccctgtat atctcccgga cccctgaagt gacctgcgtg 5220  
 gtggggacg tggccacgaa agatcctgaa gtgaagttca attggtaacgt ggacggcggtg 5280  
 gaggtgcaca acgccaagac caagccctgg gaggaaacagt acaactccac ctaccgggtg 5340  
 gtgtctgtgc tgaccgtgt gcaccaggac tggctgaacg gcaaagaata caagtgcag 5400  
 gtgtccaaaca aggccctgccc tgccctatc gaaaagacca tctccaaggc taaggccag 5460  
 ccacgggaac ctcaggtcta cacactgcct cctagccggg acgagctgac caagaaccag 5520  
 gtgtccctga cctgtcttgtt gaagggttcc tacccctccg atatgcgtt ggagtggag 5580  
 tctaacggcc agcctgagaa caactacaag accacccttc ctgtgtggc ctccgacggc 5640  
 tccttcctcc tggactccaa gctgaccgtg gacaagtccc ggtggcagca gggcaacgtg 5700  
 ttctctgttcc cctgtatgtc cgaggccctg cacaaccact acaccaggaa gtcctgtcc 5760  
 ctgtctccgt gcaaggaaca aaaactcatc tcagaagagg atctgaatgc tggggccag 5820  
 gacacgcagg aggtcatgtt ggtgccacac tccttgccct ttaaggttgtt ggtgtatca 5880  
 gccatctgg ccctgggtt gctcaccate atctccctta tcattctcat catgtttgg 5940  
 cagaagaaggc cacgttagta aaagcttgc acttggaaag taatagttt tcctgcacgg 6000  
 gttagaatca gcctcgactg tgccttctag ttgcctggca tctgttgtt gcccctcccc 6060  
 cgtgccttcc ttgaccctgg aagggtccac tcccactgtc cttectaataa aaaaatgagga 6120  
 aattgcattcg cattgtctga gtaggtgtca ttctattctg ggggggtgggg tggggcagga 6180  
 cagcaaggggg gaggattggg aagacaatag caggcatgtc ggggatggcc cgggcatgtat 6240  
 aacttcgtat aatgtatgtt atacgaagtt atgtatacgg cgcccccact agggacagga 6300  
 ttggtgacag aaaagccccca tccttaggec tccttcctcc tagtctccgt atattgggtc 6360  
 taaccccccac ctcctgttag gcagattccct tatctggta cacacccca ttctggag 6420  
 ccatctctct cttggccaga acctctaagg ttgtcttacg atggagccag agaggatct 6480  
 gggagggaga gcttggcagg ggggtgggggg gaagggggggg atgcgtgacc tgcccggttc 6540  
 tcagtggcca ccctgcgtca cccttcctca gaacctgagc tgctctgacg cggctgtctg 6600  
 tgctgttca ctgtcctgg tgctgcagtc tccttacact tcccaagagg agaaggcgtt 6660  
 tggaaaaaca aatcagaat aagtggtcc tgagttctaa ctgttgttct tcacccttct 6720  
 agtccccat ttatattgtt ctcctggcg tcagttttac ctgtgagata aggccagtag 6780  
 ccagccccgt ctcggcaggc ctgtggtag gagggggggtg tccgtgtgga aaactccctt 6840  
 tgtgagaatg gtgcgtcata ggtgttcacc aggtcggtggc cgcctctact cccttctct 6900

ttctccatcc ttcttcctt aaagagtccc cagtgtatc tggacatat tcctccgccc 6960  
agagcagggt cccgcctccc taaggccctg ctctgggctt ctgggttga gtcctggca 7020  
agcccaggag aggcgcgtcg gcttccctgt ccccccttctt cgtccaccat ctcatcccc 7080  
tggctctctt gccccctccc tacaggggtt cctggctctg ctctcttgca ggcatctct 7140  
cgatctctcg atttcgatca agacattctt ttaatggctt ttctggaca ccactagggg 7200  
tcagaagtag ttcataaaac ttcttcctt ccataatctc attgggttacc ttgggtatc 7260  
gaaacttaat taagccacct gacgcgcct gtacggcgcg attaagcgcg gcgggtgtgg 7320  
tggttacgcg cagcgtgacc gctacacttg ccagcgcctt agcgcgcctt ccttcgtt 7380  
tctcccttc ctttctcgcc acgttcgcgg gcttccccc tcaagctcta aatcgggggc 7440  
tccctttagg gttccgattt agtgccttac ggcacctcga ccccaaaaaa ctgatttagg 7500  
gtgatgggtc acgttagtggg ccategcctt gatagacggt ttctgcctt tgacgttgg 7560  
agtccacgtt cttaatagt ggactcttgc tccaaactgg aacaacactc aaccctatct 7620  
cggtctattt tttgattta taagggattt tgccgatttc ggcctattgg taaaaaaatg 7680  
agctgattt aaaaaattt aacgcgaattt ttaacaaaat attaacgctt acaatttacg 7740  
cgtaagata cattgatgag ttggacaaa ccacaactag ttaattaacc agtcaagtca 7800  
gctacttggc gagatcgact tgtctgggtt tcgactacgc tcagaattgc gtcaagtcaag 7860  
ttcgtatctgg tcctgttat tgcacccgtt ctccgattac gagtttcatt taaatcatgt 7920  
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaaggc cgctgtgtc gcgttttcc 7980  
ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcaag aggtggcga 8040  
acccgacagg actataaaga taccaggctt ttccccctgg aagctccctc gtgcgtctc 8100  
ctgtccgac cctgcgcgtt accggatacc tgcgcctt ctcccttcg ggaagcgtgg 8160  
cgctttctca tagtcacgc tgttaggtatc tcagttcggtt gttaggtcggtt cgctccaagc 8220  
tggctgtgt gcacgaaccc cccgttcage ccgaccgcgtg cgccttatcc gtaactatc 8280  
gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca 8340  
ggatttagcag agcgaggatgtt gtagggcgtt ctacagagttt ctgaaagtgg tggcttaact 8400  
acggctacac tagaagaaca gtatggta tctgcgtctt gctgaagcca gttaccctcg 8460  
gaaaaagagt tggtagtctt tgatccggca aacaaaccac cgctggtagc ggtgggtttt 8520  
ttgtttgcaa gcagcagattt acgacgcagaa aaaaaggatc tcaagaagat ctgttgcatt 8580  
tttctacggg gtctgcgtc cagtgaaacg aaaaactcactt ttaagggattt tggctatga 8640  
gattatcaaa aaggatctt accttagatcc tttaaattt aaaaatgaagt tttaaatcaa 8700  
tctaaagtat atatgagttt acgttgtctg acagttacca atgcttaatc agtgaggcac 8760  
ctatctcagc gatctgtctt ttctgtcat ccatagttgc atttaaattt ccgaactctc 8820  
caaggccctc gtccggaaaat ttccaaaccc ttctgtcccgat ccatacttgc ggcgttccct 8880  
cgaacgaact atcgcaagtc tcttggccgg ccttgcgcctt tggctattgc ttggcagggc 8940

[0054]

ctatgccag gtattactcc aatccgaat atccgagatc gggatcaccc gagagaagtt 9000  
 caacccat cctcaatccc gatctatccg agatccgagg aatatcgaa tcggggcg 9060  
 cctgggtac cgagaacat cctctcagtg cgagtctcga cgatccatat cggtgcgtgg 9120  
 cagtcagcca gtcggaatcc agcttggac ccaggaagtc caatcgicag atattgtact 9180  
 caagcctggt cacggcagecg taccgatctg tttaaaccta gatattgata gtctgatcgg 9240  
 tcaacgtata atcgagtcct agctttgca aacatctatc aagagacagg atcagcagga 9300  
 ggcttcgca tgattgaaca agatggattg cacgcaggct ctccggcggc ttgggtggag 9360  
 aggctattcg gctatgactg ggcacaacag acaatcgct getctgatgc cgccgtgttc 9420  
 cggtgtcag cgcaggggcg tccggttctt ttgtcaaga ccgacctgtc cggtgccctg 9480  
 aatgaactgc aagacgagge agcgcggcta tcgtggctgg cgacgcacggg cggtccttgc 9540  
 gccccgtgtc tcgacgttgt cactgaagcg ggaaggact ggctgttatt gggcgaagtg 9600  
 ccggggcagg atctccgtc atctcacctt gctccgtccg agaaagtatc catcatggct 9660  
 gatgcaatgc ggcggctgca tacgcttgcgatccgcttccgatccgatccgatccgatccg 9720  
 aaacatcgca tcgagcggc acgtactcgg atggaagccg gtctgtcga tcaggatgtat 9780  
 ctggacgaag agcatcagg gctcgccca gccgaactgt tcggcaggct caaggcgtct 9840  
 atgcccgcacg gcgaggatct cgtcgtgacc cacggcgatg cctgctgcc gaatatcatg 9900  
 gtggaaaatg gccgttttc tggattcatc gactgtggcc gtctgggtgt ggccggaccgc 9960  
 tatcaggaca tagcgttggc taccgtgtat attgtcaag agcttggcgg cgaatggct 10020  
 gaccgcgtcc ttgtgcattt cggatcgcc gcccggatt cgacgcgtat cgccttctat 10080  
 cgccttcttgc acgaggatctt ctgaccgatt ctggatgtatc tggcgcagaa aaaaatgcct 10140  
 gatgcgacgc tgcgcgtctt atactccac atatgcaga ttcagcaacg gatacggctt 10200  
 ccccaacttgc cccacttcca tacgtgtctt ctttaccaga aatttatctt taaggcgat 10260  
 taaaactcgac tctggctata tcgaatctcc gtcgttgcgatccgatccgatccgatccgatccg 10320  
 cgctcatttgc ctgcgtggc atcgaatctc gtcagctatc gtcagcttac ctgttggca 10380

&lt;210&gt; 167

&lt;211&gt; 132

&lt;212&gt; PRT

&lt;213&gt; 人工序列(Artificial Sequence)

&lt;400&gt; 167

Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala

1 5 10 15

Thr Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala

20 25 30

Ser Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val

35 40 45

Tyr His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr

50 55 60

Ala Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly

65 70 75 80

Asn Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val

85 90 95

Leu Leu Asp Leu His Pro Gly Ile Lys Ala Ile Val Lys Asp Ser Asp

100 105 110

Gly Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr

115 120 125

Val Trp Glu Gly

130

<210> 168

<211> 18

<212> PRT

<213> 人工序列(Artificial Sequence)

[0055]

<400> 168

Met Arg Ala Trp Ile Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu

1 5 10 15

Ala Ala

<210> 169

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 169

Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val

1 5 10 15

Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gly Asn Ile His Asn

20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu

35 40 45

Ile Tyr Tyr Thr Thr Leu Ala Asp Gly Val Pro Ser Arg Phe Ser

50 55 60

Gly Ser Gly Ser Gly Thr Asp Tyr Thr Phe Thr Ile Ser Ser Leu Gln  
65 70 75 80  
Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln His Phe Trp Ser Thr Pro  
85 90 95  
Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr  
100 105 110  
<210> 170  
<211> 102  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 170  
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly  
1 5 10 15  
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala  
20 25 30  
Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln  
35 40 45  
[0056] Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser  
50 55 60  
Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Leu Tyr  
65 70 75 80  
Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser  
85 90 95  
Phe Asn Arg Gly Glu Cys  
100  
<210> 171  
<211> 15  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 171  
Met Ser Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr  
1 5 10 15  
<210> 172  
<211> 116

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 172

Met Ala Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg Pro

1 5 10 15

Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Thr Phe Ser

20 25 30

Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Arg Gly Leu Glu

35 40 45

Trp Ile Gly Met Ile Trp Gly Asp Gly Asn Thr Asp Tyr Asn Ser Ala

50 55 60

Leu Lys Ser Arg Val Thr Met Leu Val Asp Thr Ser Lys Asn Gln Phe

65 70 75 80

Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr

85 90 95

Cys Ala Arg Glu Arg Asp Tyr Arg Leu Asp Tyr Trp Gly Gln Gly Ser

100 105 110

[0057]

Leu Val Thr Val

115

<210> 173

<211> 332

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 173

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

1 5 10 15

Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys

20 25 30

Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu

35 40 45

Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu

50 55 60

Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr

65 70 75 80

Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val  
85 90 95  
Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro  
100 105 110  
Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe  
115 120 125  
Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val  
130 135 140  
Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe  
145 150 155 160  
Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro  
165 170 175  
Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr  
180 185 190  
Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val  
195 200 205  
Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala  
[0058] 210 215 220  
Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg  
225 230 235 240  
Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly  
245 250 255  
Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro  
260 265 270  
Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser  
275 280 285  
Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln  
290 295 300  
Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His  
305 310 315 320  
Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
325 330  
<210> 174  
<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 174

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1 5 10

<210> 175

<211> 50

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 175

Asn Ala Val Gly Gln Asp Thr Gln Glu Val Ile Val Val Pro His Ser

1 5 10 15

Leu Pro Phe Lys Val Val Ile Ser Ala Ile Leu Ala Leu Val Val

20 25 30

Leu Thr Ile Ile Ser Leu Ile Ile Leu Ile Met Leu Trp Gln Lys Lys

35 40 45

Pro Arg

[0059]

50

<210> 176

<211> 264

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 176

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val

1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser

20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe

35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala

50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val

65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu

85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys

100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro

115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala

130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu

145 150 155 160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala

165 170 175

Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys

180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp

195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala

210 215 220

[0060] Thr Arg Asp Ile Ala Glu Glu Leu Gly Glu Trp Ala Asp Arg Phe

225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe

245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe

260

<210> 177

<211> 392

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 177

ccatggcca gggtcagctg gttcagtctg ggcggaaagt gaagaaaacct ggcagcagcg 60

tgaagggtgc ctgcaaagca agcggggca cctttgggt cggcgcctt acatgggtcc 120

gacaggctcc aggacagggc ctgtaatgga tggcgccat catccctatc ttccggctga 180

ccaatctggc ccagaacttc cagggcagag tgaccatcac agccgacgag agcaccagca 240

ccgtgtcat ggaactgagc agcctgagaa gcgaggacac cgccgtgtac tactgtgcca 300

gaagcagccg gatctacgat ctgaacccta gcctgaccgc ctactacgac atggatgtgt 360

ggggccaggg cacaatggtc accgtctcga gt 392  
<210> 178  
<211> 129  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 178  
Met Ala Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro  
1 5 10 15  
Gly Ser Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Trp  
20 25 30  
Phe Gly Ala Phe Thr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu  
35 40 45  
Trp Met Gly Gly Ile Ile Pro Ile Phe Gly Leu Thr Asn Leu Ala Gln  
50 55 60  
Asn Phe Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr Ser Thr  
65 70 75 80  
Val Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr  
[0061] 85 90 95  
Tyr Cys Ala Arg Ser Ser Arg Ile Tyr Asp Leu Asn Pro Ser Leu Thr  
100 105 110  
Ala Tyr Tyr Asp Met Asp Val Gly Gln Gly Thr Met Val Thr Val Ser  
115 120 125  
Ser  
<210> 179  
<211> 8  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 179  
Gly Gly Thr Phe Trp Phe Gly Ala  
1 5  
<210> 180  
<211> 8  
<212> PRT  
<213> 人工序列(Artificial Sequence)

<400> 180

Ile Ile Pro Ile Phe Gly Leu Thr

1 5

<210> 181

<211> 21

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 181

Ala Arg Ser Ser Arg Ile Tyr Asp Leu Asn Pro Ser Leu Thr Ala Tyr

1 5 10 15

Tyr Asp Met Asp Val

20

<210> 182

<211> 225

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 182

[0062]

Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys Pro Gly Ala

1 5 10 15

Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr

20 25 30

Trp Met His Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile

35 40 45

Gly Glu Ile Asn Pro Ser Asn Gly Arg Thr Asn Tyr Asn Glu Lys Phe

50 55 60

Lys Ser Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr

65 70 75 80

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys

85 90 95

Ala Arg Glu Arg Pro Leu Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr

100 105 110

Ser Val Thr Val Ser Ser Asp Ile Val Met Thr Gln Ser His Lys Phe

115 120 125

Met Ser Thr Ser Val Gly Asp Arg Val Ser Ile Thr Cys Lys Ala Ser

130 135 140  
Gln Asp Val Ser Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Gln  
145 150 155 160  
Ser Pro Lys Leu Leu Ile Tyr Ser Ala Ser Tyr Arg Tyr Thr Gly Val  
165 170 175  
Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr  
180 185 190  
Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Gln  
195 200 205  
Arg Tyr Ser Thr Pro Arg Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile  
210 215 220  
Lys  
225  
<210> 183  
<211> 225  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
[0063]  
<400> 183  
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 Glu Ile Asn Pro Ser Gly Gly Arg Thr Asn Tyr Asn Glu Lys Phe  
50 55 60  
Lys Ser Arg Val Thr Met Thr Arg Asp Thr Ser Thr Ser Thr Val Tyr  
65 70 75 80  
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys  
85 90 95  
Ala Arg Glu Arg Pro Leu Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr  
100 105 110  
Thr Val Thr Val Ser Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser  
115 120 125

Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser  
 130 135 140  
 Gln Asp Val Ser Thr Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys  
 145 150 155 160  
 Ala Pro Lys Leu Leu Ile Tyr Ser Ala Ser Tyr Arg Tyr Thr Gly Val  
 165 170 175  
 Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr  
 180 185 190  
 Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln  
 195 200 205  
 Arg Tyr Ser Thr Pro Arg Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile  
 210 215 220  
 Lys  
 225  
 <210> 184  
 <211> 225  
 <212> PRT  
 [0064] <213> 人工序列(Artificial Sequence)  
 <400> 184  
 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 Glu Ile Ser Pro Phe Gly Gly Arg Thr Asn Tyr Asn Glu Lys Phe  
 50 55 60  
 Lys Ser Arg Val Thr Met Thr Arg Asp Thr Ser Thr Ser Thr Val Tyr  
 65 70 75 80  
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys  
 85 90 95  
 Ala Arg Glu Arg Pro Leu Tyr Ala Ser Asp Leu Trp Gly Gln Gly Thr  
 100 105 110  
 Thr Val Thr Val Ser Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser

115 120 125

Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser

130 135 140

Gln Gly Ile Ser Ser Ala Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys

145 150 155 160

Ala Pro Lys Leu Leu Ile Tyr Ser Ala Ser Tyr Arg Tyr Thr Gly Val

165 170 175

Pro Ser Arg Phe Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr

180 185 190

Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln

195 200 205

Arg Tyr Ser Leu Trp Arg Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile

210 215 220

Lys

225

<210> 185

<211> 90

[0065]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 185

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 Glu Ile Ser Pro Phe Gly Gly Arg Thr Asn Tyr Asn Glu Lys Phe

50 55 60

Lys Ser Arg Val Thr Met Thr Arg Asp Thr Ser Thr Ser Thr Val Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 186

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 186

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

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 187

<211> 90

[0066]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 187

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

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 188

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 188

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 Asn 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

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 189

<211> 90

[0067]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 189

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 Gly Tyr

20 25 30

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

35 40 45

Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe

50 55 60

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

65 70 75 80

Met Glu Leu Ser Arg Leu Arg Ser Asp Asp

85 90

<210> 190

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 190

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 Gly Tyr

20 25 30

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

35 40 45

Gly Arg Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe

50 55 60

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

65 70 75 80

Met Glu Leu Ser Arg Leu Arg Ser Asp Asp

85 90

<210> 191

<211> 90

[0068]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 191

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 Gly Tyr

20 25 30

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

35 40 45

Gly Arg Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe

50 55 60

Gln Gly Arg Val Thr Ser Thr Arg Asp Thr Ser Ile Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Arg Leu Arg Ser Asp Asp

85 90

<210> 192

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 192

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 Gly Tyr

20 25 30

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

35 40 45

Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe

50 55 60

Gln Gly Trp Val Thr Met Thr Arg Asp Thr Ser Ile Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Arg Leu Arg Ser Asp Asp

85 90

<210> 193

<211> 90

[0069]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 193

Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Leu Gly Ala

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Gly Tyr

20 25 30

Tyr Met His Trp Val Xaa Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe

50 55 60

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

65 70 75 80

Met Glu Leu Ser Arg Leu Arg Ser Asp Asp

85 90

<210> 194

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 194

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

Ala Met His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Met

35 40 45

Gly Trp Ile Asn Ala Gly Asn Gly Asn Thr Lys Tyr Ser Gln Lys Phe

50 55 60

Gln Gly Arg Val Thr Ile Thr Arg Asp Thr Ser Ala Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 195

<211> 90

[0070]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 195

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

Gly Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Trp Ile Ser Ala Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Lys Leu

50 55 60

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

65 70 75 80

Met Glu Leu Arg Ser Leu Arg Ser Asp Asp

85 90

<210> 196

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 196

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

Gly Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Trp Ile Ser Ala Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Lys Leu

50 55 60

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

65 70 75 80

Met Glu Leu Arg Ser Leu Arg Ser Asp Asp

85 90

<210> 197

<211> 90

[0071]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 197

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

Gly Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Trp Ile Ser Ala Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Lys Leu

50 55 60

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

65 70 75 80

Met Glu Leu Arg Ser Leu Arg Ser Asp Asp

85 90

<210> 198

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 198

Gln Met Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Thr Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Tyr Arg

20 25 30

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

35 40 45

Gly Trp Ile Thr Pro Phe Asn Gly Asn Thr Asn Tyr Ala Gln Lys Phe

50 55 60

Gln Asp Arg Val Thr Ile Thr Arg Asp Arg Ser Met Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 199

<211> 90

[0072]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 199

Gln Met Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Thr Gly Ser

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Tyr Arg

20 25 30

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

35 40 45

Gly Trp Ile Thr Pro Phe Asn Gly Asn Thr Asn Tyr Ala Gln Lys Phe

50 55 60

Gln Asp Arg Val Thr Ile Thr Arg Asp Arg Ser Met Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 200

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 200

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

Asp Ile Asn Trp Val Arg Gln Ala Thr Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Trp Met Asn Pro Asn Ser Gly Asn Thr Gly Tyr Ala Gln Lys Phe

50 55 60

Gln Gly Arg Val Thr Met Thr Arg Asn Thr Ser Ile Ser Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp

85 90

<210> 201

<211> 375

[0073]

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 201

tttttgcca tggcccagg tcagctggtt cagtctggcg ccgaagtcaa gaaacctggc 60

gcctctgtga aggtgtcctg caagggccage ggctacacct ttaccagcta cgctatgcac 120

tgggtccgac aggccccctgg acaaggactt gagtggtatgg gcgagatcag cccattcggc 180

ggcaggacca actacaacga gaagtcaag agccgcgtga ccatgaccag agacaccagc 240

acctccaccc tgtacatggc actgagcagc ctgagaagcg aggacaccgc cgtgtactac 300

tgtgccagag agaggccact gtacgcctct gatctttggg gccagggcac caccgtgaca 360

gtctcgagtt tttt 375

<210> 202

<211> 375

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 202

tttttgcca tggcccagg tcagctggtt cagtctggcg ccgaagtcaa gaaacctggc 60

gcctctgtga aggtgtcctg caagggccage ggctacacct ttaccagcta cgatatgcac 120

tgggtccgac aggcccctgg acaaggactt gagtgatgg gcgagatcg cccattcggc 180  
ggcaggacca actacaacga gaagtcaag agccgcgtga ccatgaccag agacaccagc 240  
acctccaccc tgcataatggaa actgagcagc ctgagaagcg aggacaccgc cggtactac 300  
tgtgccagag agaggccact gtacgcctct gatctttggg gccagggcac caccgtgaca 360  
gtctcgagtt tttt 375  
<210> 203  
<211> 375  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 203  
tttttgcca tggcccaggt tcagctggtt cagtcgtggc ccgaagtgaa gaaacctggc 60  
gcctctgtga aggtgtcctg caaggccagc ggctacacct ttaccagcta ctacatgcac 120  
tgggtccgac aggcccctgg acaaggactt gagtgatgg gcgagatcaa cccatctggc 180  
ggcagcacca actacaacga gaagtcaag agccgcgtga ccatgaccag agacaccagc 240  
acctccaccc tgcataatggaa actgagcagc ctgagaagcg aggacaccgc cggtactac 300  
tgtgccagag agaggccact gtacgcctct gatctttggg gccagggcac caccgtgaca 360  
gtctcgagtt tttt 375  
[0074]  
<210> 204  
<211> 375  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 204  
tttttgcca tggcccaggt tcagctggtt cagtcgtggc ccgaagtgaa gaaacctggc 60  
gcctctgtga aggtgtcctg caaggccagc ggctacacct ttaccagcta cgctatgcac 120  
tgggtccgac aggcccctgg acaaggactt gagtgatgg gcgagatcaa cccatctggc 180  
ggcagcacca actacaacga gaagtcaag agccgcgtga ccatgaccag agacaccagc 240  
acctccaccc tgcataatggaa actgagcagc ctgagaagcg aggacaccgc cggtactac 300  
tgtgccagag agaggccact gtacgcctct gatctttggg gccagggcac caccgtgaca 360  
gtctcgagtt tttt 375  
<210> 205  
<211> 375  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 205

tttttgcca tggcccaggt tcagctggtt cagtcgtgcg ccgaagtgaa gaaacctggc 60  
 gcctctgtga aggtgtcctg caaggccagc ggctacacct ttaccagcta cgtatgcac 120  
 tgggtccgac aggccccctgg acaaggactt gagtggatgg gcgagatcaa cccatctggc 180  
 ggcagcacca actacaacga gaagtcaag agccgcgtga ccatgaccag agacaccagc 240  
 acctccaccc tgtacatgga actgagcagc ctgagaagcg aggacaccgc cgtgtactac 300  
 tgtgccagag agaggccact gtacgcctct gatctttggg gccagggcac caccgtgaca 360  
 gtctcgagtt tttt 375  
 <210> 206  
 <211> 375  
 <212> DNA  
 <213> 人工序列(Artificial Sequence)  
 <400> 206

tttttgcca tggcccaggt tcagctggtt cagtcgtgcg ccgaagtgaa gaaacctggc 60  
 gcctctgtga aggtgtcctg caaggccagc ggctacacct ttaccagcta ctacatgcac 120  
 tgggtccgac aggccccctgg acaaggactt gagtggatgg gcgagatcg cccattggc 180  
 ggcaggacca actacaacga gaagtcaag agccgcgtga ccatgaccag agacaccagc 240  
 acctccaccc tgtacatgga actgagcagc ctgagaagcg aggacaccgc cgtgtactac 300  
 [0075] tgtgccagag agaggccact gtacgcctct gatctttggg gccagggcac caccgtgaca 360  
 gtctcgagtt tttt 375  
 <210> 207  
 <211> 353  
 <212> DNA  
 <213> 人工序列(Artificial Sequence)  
 <400> 207

tttttgtcta gegacatcca gatgacacag agccctagca gcctgtctgc cagcgtggga 60  
 gacagagtga ccatcacctg tagagccagc cagggcatct ctctgcctt ggcattgttat 120  
 cagcagaagc ctggcaaggc ccctaagctg ctgatctaca gcgccagcta aagatacacc 180  
 ggcgtgcca gcagatttc tggcagcggc tctggcaccc acttcaccc taccataagc 240  
 agccgtcagc ctgaggatata cgccaccc tacgtccagc agcggtaatc ttagtaacgg 300  
 acatttggcc agggcacca aactggaaatc aagcgttaccg cggccgcattt ttt 353  
 <210> 208  
 <211> 8  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)

<400> 208

Gly Tyr Thr Phe Thr Ser Tyr Tyr

1 5

<210> 209

<211> 8

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 209

Gly Tyr Thr Phe Thr Ser Tyr Ser

1 5

<210> 210

<211> 8

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 210

Gly Tyr Thr Phe Thr Ser Tyr Asp

[0076]

1 5

<210> 211

<211> 8

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 211

Gly Tyr Thr Phe Thr Ser Tyr Ala

1 5

<210> 212

<211> 8

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 212

Gly Tyr Thr Phe Thr Ser Tyr Trp

1 5

<210> 213

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 213

Ile Asn Pro Ser Gly Gly Ser Thr Asn Tyr Asn

1 5 10

<210> 214

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 214

Ile Asn Pro Ser Gly Gly Arg Thr Asn Tyr Asn

1 5 10

<210> 215

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 215

Ile Ser Pro Phe Gly Gly Arg Thr Asn Tyr Asn

[0077]

1 5 10

<210> 216

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 216

Ile Asn Pro Ser Asn Gly Arg Thr Asn Tyr Asn

1 5 10

<210> 217

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 217

Ser Leu Arg Tyr Thr

1 5

<210> 218

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 218

Ser Glu Arg Tyr Thr

1 5

<210> 219

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 219

Ser Gln Arg Tyr Thr

1 5

<210> 220

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

[0078] <400> 220

Ser Ser Arg Tyr Thr

1 5

<210> 221

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 221

Ser Ala Arg Tyr Thr

1 5

<210> 222

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 222

Ser Pro Arg Tyr Thr

1 5

<210> 223

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 223

Ser Asp Arg Tyr Thr

1 5

<210> 224

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 224

Ser Thr Arg Tyr Thr

1 5

<210> 225

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

[0079] <400> 225

Ser Met Arg Tyr Thr

1 5

<210> 226

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 226

Ser Val Arg Tyr Thr

1 5

<210> 227

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 227

Ser Asn Arg Tyr Thr

1 5

<210> 228  
<211> 5  
<212> PRT  
<213> 人工序列(Artificial Sequence)

<400> 228  
Ser Tyr Leu Glu Ser

1 5

<210> 229  
<211> 5  
<212> PRT  
<213> 人工序列(Artificial Sequence)

<400> 229

Ser Tyr Arg Tyr Thr

1 5

<210> 230  
<211> 5  
<212> PRT  
[0080] <213> 人工序列(Artificial Sequence)

<400> 230

Ser His Arg Tyr Thr

1 5

<210> 231  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)

<400> 231

Gln Gln Arg Tyr Ser Leu Gln Arg Thr

1 5

<210> 232  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)

<400> 232

Gln Gln Arg Tyr Ser Pro Trp Arg Thr

1 5

<210> 233

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 233

Gln Gln Arg Tyr Ser Asn Ser Arg Thr

1 5

<210> 234

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 234

Gln Gln Arg Tyr Ser Leu Lys Arg Thr

1 5

<210> 235

<211> 9

[0081] <212> PRT

<213> 人工序列(Artificial Sequence)

<400> 235

Gln Gln Arg Tyr Ser Ser Pro Arg Thr

1 5

<210> 236

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 236

Gln Gln Arg Tyr Ser Asn Asn Arg Thr

1 5

<210> 237

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 237

Gln Gln Arg Tyr Ser Ser Gly Arg Thr

1 5

<210> 238

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 238

Gln Gln Arg Tyr Ser Leu Ala Arg Thr

1 5

<210> 239

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 239

Gln Gln Arg Tyr Ser Thr Leu Arg Thr

1 5

<210> 240

[0082]

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 240

Gln Gln Arg Tyr Ser Leu Ala Arg Thr

1 5

<210> 241

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 241

Gln Gln Arg Tyr Ser Ser Leu Arg Thr

1 5

<210> 242

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 242

Gln Gln Arg Tyr Ser Thr His Arg Thr

1 5

<210> 243

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 243

Gln Gln Arg Tyr Ser Phe Ala Arg Thr

1 5

<210> 244

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 244

Gln Gln Arg Tyr Ser Ile Arg Arg Thr

1 5

[0083]

<210> 245

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 245

Gln Gln Arg Tyr Ser Val Cys Arg Thr

1 5

<210> 246

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 246

Gln Gln Arg Tyr Ser Trp Ala Arg Thr

1 5

<210> 247

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 247

Gln Gln Arg Tyr Ser Thr Thr Arg Thr

1 5

<210> 248

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 248

Gln Gln Arg Tyr Ser Gln Met Arg Thr

1 5

<210> 249

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 249

Gln Gln Arg Tyr Ser Ala Asp Arg Thr

[0084]

1 5

<210> 250

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 250

Gln Gln Arg Tyr Ser Phe Asn Arg Thr

1 5

<210> 251

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 251

Gln Gln Arg Tyr Ser Tyr Asp Arg Thr

1 5

<210> 252

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 252

Gln Gln Arg Tyr Ser Ser Gln Arg Thr

1 5

<210> 253

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 253

Gln Gln Arg Tyr Ser Tyr Val Arg Thr

1 5

<210> 254

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 254

[0085]

Gln Gln Arg Tyr Ser Leu Tyr Arg Thr

1 5

<210> 255

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 255

Gln Gln Arg Tyr Ser Ala Asp Arg Thr

1 5

<210> 256

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 256

Gln Gln Arg Tyr Ser Ser Met Arg Thr

1 5

<210> 257

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 257

Gln Gln Arg Tyr Ser Tyr Tyr Arg Thr

1 5

<210> 258

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 258

Gln Gln Arg Tyr Ser Leu Ile Arg Thr

1 5

<210> 259

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

[0086] <400> 259

Gln Gln Arg Tyr Ser Glu Arg Arg Thr

1 5

<210> 260

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 260

Gln Gln Arg Tyr Ser

1 5

<210> 261

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 261

Gln Gln Arg Tyr Ser Gln Met Arg Thr

1 5

<210> 262  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 262  
Gln Gln Arg Tyr Ser Val Asn Arg Thr  
1 5  
<210> 263  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 263  
Gln Gln Arg Tyr Ser Phe Phe Arg Thr  
1 5  
<210> 264  
<211> 9  
<212> PRT  
[0087] <213> 人工序列(Artificial Sequence)  
<400> 264  
Gln Gln Arg Tyr Ser Asn Val Arg Thr  
1 5  
<210> 265  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 265  
Gln Gln Arg Tyr Ser Ala Val Arg Thr  
1 5  
<210> 266  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 266  
Gln Gln Arg Tyr Ser Asp Pro Arg Thr

1 5

<210> 267

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 267

Gln Gln Arg Tyr Ser Met Thr Arg Thr

1 5

<210> 268

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 268

Gln Gln Arg Tyr Ser Ile Ser Arg Thr

1 5

<210> 269

<211> 9

[0088]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 269

Gln Gln Arg Tyr Ser Leu Tyr Arg Thr

1 5

<210> 270

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 270

Gln Gln Arg Tyr Ser Asn Met Arg Thr

1 5

<210> 271

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 271

Gln Gln Arg Tyr Ser Gln Met Arg Thr

1 5

<210> 272

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 272

Gln Gln Arg Tyr Ser Ala Asp Arg Thr

1 5

<210> 273

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 273

Gln Gln Arg Tyr Ser Thr Ile Arg Thr

1 5

<210> 274

[0089]

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 274

Gln Gln Arg Tyr Ser Cys Ala Arg Thr

1 5

<210> 275

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 275

Gln Gln Arg Tyr Ser Asn His Arg Thr

1 5

<210> 276

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 276

Gln Gln Arg Tyr Ser Ala Lys Arg Thr

1 5

<210> 277

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 277

Gln Gln Arg Tyr Ser His Arg Arg Thr

1 5

<210> 278

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 278

Gln Gln Arg Tyr Ser Cys Thr Arg Thr

1 5

[0090]

<210> 279

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 279

Gln Gln Arg Tyr Ser Tyr Met Arg Thr

1 5

<210> 280

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 280

Gln Gln Arg Tyr Ser His Phe Arg Thr

1 5

<210> 281

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 281

Gln Gln Arg Tyr Ser Ser Ala Arg Thr

1 5

<210> 282

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 282

Gln Gln Arg Tyr Ser Asn Arg Arg Thr

1 5

<210> 283

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 283

[0091] Gln His Ser Arg Asp Leu Pro Leu Thr

1 5

<210> 284

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 284

Gln Gln Arg Tyr Ser Thr Gly Arg Thr

1 5

<210> 285

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 285

Gln Gln Arg Tyr Ser Asp Thr Arg Thr

1 5

<210> 286

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 286

Gln Gln Arg Tyr Ser Glu Thr Arg Thr

1 5

<210> 287

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 287

Gln Gln Arg Tyr Ser Asn Ser Arg Thr

1 5

<210> 288

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 288

[0092]

Gln Gln Arg Tyr Ser Pro Pro Arg Thr

1 5

<210> 289

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 289

Gln Gln Arg Tyr Ser Ser Ala Arg Thr

1 5

<210> 290

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 290

Gln Gln Arg Tyr Ser Leu Glu Arg Thr

1 5

<210> 291

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 291

Gln Gln Arg Tyr Ser His Ser Arg Thr

1 5

<210> 292

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 292

Gln Gln Arg Tyr Ser Val Ala Arg Thr

1 5

<210> 293

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

[0093] <400> 293

Gln Gln Arg Tyr Ser Leu Thr Arg Thr

1 5

<210> 294

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 294

Gln Gln Arg Tyr Ser His Pro Arg Thr

1 5

<210> 295

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 295

Gln Gln Arg Tyr Ser Asp Thr Arg Thr

1 5

<210> 296  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 296  
Gln Gln Arg Tyr Ser Asn Met Arg Thr  
1 5  
<210> 297  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 297  
Gln Gln Arg Tyr Ser Ser Lys Arg Thr  
1 5  
<210> 298  
<211> 9  
<212> PRT  
[0094] <213> 人工序列(Artificial Sequence)  
<400> 298  
Gln Gln Arg Tyr Ser His Phe Arg Thr  
1 5  
<210> 299  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 299  
Gln Gln Arg Tyr Ser Pro Pro Arg Thr  
1 5  
<210> 300  
<211> 9  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 300  
Gln Gln Arg Tyr Ser Glu Ser Arg Thr

1 5

<210> 301

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 301

Gln Gln Arg Tyr Ser Val Tyr Arg Thr

1 5

<210> 302

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 302

Gln Gln Arg Tyr Ser Cys Ile Arg Thr

1 5

<210> 303

<211> 9

[0095]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 303

Gln Gln Arg Tyr Ser Ala Thr Arg Thr

1 5

<210> 304

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 304

Gln Gln Arg Tyr Ser Ala Thr Arg Thr

1 5

<210> 305

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 305

Gln Gln Arg Tyr Ser Asp Cys Arg Thr

1 5

<210> 306

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 306

Gln Gln Arg Tyr Ser Met Ala Arg Thr

1 5

<210> 307

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 307

Gln Gln Arg Tyr Ser Gly Glu Arg Thr

1 5

<210> 308

[0096]

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 308

Gln Gln Arg Tyr Ser Ile Arg Arg Thr

1 5

<210> 309

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 309

Gln Gln Arg Tyr Ser Leu Met Arg Thr

1 5

<210> 310

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 310

Gln Gln Arg Tyr Ser Leu Asp Arg Thr

1 5

<210> 311

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 311

Gln Gln Arg Tyr Ser Tyr Val Arg Thr

1 5

<210> 312

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 312

Gln Gln Arg Tyr Ser Ser Ala Arg Thr

[0097]

1 5

<210> 313

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 313

Gln Gln Arg Tyr Ser Leu Asn Arg Thr

1 5

<210> 314

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 314

Gln Gln Arg Tyr Ser Leu Trp Arg Thr

1 5

<210> 315

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 315

Gln Gln Arg Tyr Ser Thr Pro Arg Thr

1 5

<210> 316

<211> 224

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 316

aagtccctaa gcttgtctcc gggtaaatga gtgccacggc cggcaagccc ccgctcaagt 60

ccctaagctt gtctccaggt aaatgagtgc cacggccggc aagccccgc tcaagtccct 120

aagcttgtct ccggtaagt gagtgccacg gccggcaagc ccccgctcaa gtccctaagc 180

ttgtctccag gtaagtgagt gccacggcccg gcaagccccc gtc 224

<210> 317

<211> 3640

<212> DNA

<213> 人工序列(Artificial Sequence)

[0098]

<400> 317

agtctcgagt gcctccacca agggccctag cgtctttcct ctggccctt cttccaagtc 60

tacctctggc ggcaccgtcg ctctgggtcg cctggtaag gactacttcc ctgagcgtgt 120

gaccgtgtcc tggaaacttgc gcccctgac ctccggcggtg cataaccttcc ctggcgctct 180

ccagtccctcc ggcctgtact ccctgtcctc cgtggtgacc gtgccttccctctctggg 240

cacccagacc tacatctgca acgtgaacca caagccttcc aacaccaagg tggacaagaa 300

ggtggagccct aagtccctgcg acaagaccca cacctgccttccatgtcctg cccctgagct 360

gctggggcggaa ccctccgtgt tcctgttccc tcctaagccct aaggacaccc tgatgtatc 420

ccggaccctt gaagtgacccctt gctgggtgggt ggacgtgtcc cacgaagatc ctgaagtgaa 480

gttcaattgg tacgtggacg gctggggagggt gcacaacgcc aagaccaagg ctggggaggaa 540

acagtacaac tccacccatcc ggggtgggtc tgtgtgtacc gtgtgtacc accactggct 600

gaacggcaaa gaataacaagg gcaagggtgtc caacaaggcc ctgcctgccc ctatgaaaa 660

gaccatctcc aaggctaagg gcccggccacg ggaaccctcgt gctcacacac tgccctccatgg 720

ccggggacggag ctgaccaaga accagggtgtc cctgacccgtgt ctggtaagg gcttctaccc 780

ttccgatatac gccgtggagggt gggaggctaa cggccggccctt gagaacaact acaagaccc 840

ccctccgtgt ctggactcccg acggccctt ctccctgtac tccaagctga ccgtggacaa 900

gtcccggtgg cagcaggcga acgtgttctc ctgtccgtg atgcacggagg ccctgcacaa 960

[0099]

ccactacacc cagaagtccc tgtccctaag ctgtctcca gtaagttag tgccacggcc 1020  
 ggcaagcccc cgctcccccag gctctcgaaaa tcgcgcgagg atgttggca cgtaccccg 1080  
 ctacataactt cccgggcacc cagcatggaa ataaagcacc cagcgctgcc ctggggccct 1140  
 gcgagactgt gatggtttcc tccgtgggtc aggccgagtc tgaggccctga gtggcatgag 1200  
 ggaggcagag cgggttccac tgtccccaca ctggggcagg ctgtgcaggt gtgcctggc 1260  
 cgcctagggt ggggctcagc caggggctgc cctcgccagg ggggggatt tgccagcgtg 1320  
 gccctccctc cagcagcage tgccctggc tggggcacgg gaagccctag gagcccctgg 1380  
 ggacagacac acagccccctg cctctgttagg agactgtctt gtctgtgag cgcctgtcc 1440  
 tccgacctcc atgcccactc gggggcatgc ctatgtccatg tgcttaggga caggccctcc 1500  
 ctcacccatc tacccttacg gcactaaccct ctggctgccc tgccctagc cgcacccgca 1560  
 tggggacaca accgactccg gggacatgca ctctcgccct ctgtggaggg actggtccag 1620  
 atgeccacac acacactcag cccagacccg ttcaacaaac cccgcgtga ggttggccgg 1680  
 ccacacggcc accacacaca cacgtgcacg ctcacacac ggagcctcac cggggcgaac 1740  
 cgcacagcac ccagaccaga gcaaggctt cgcacacgtg aacactcctc agacacagggc 1800  
 cccacagggc cccacgcggc acctcaaggc ccacgagccg ctggcagct tctccacatg 1860  
 ctgacctgtc cagacaaacc cagccctctt ctacaaggt gcccgtcag cggccacaca 1920  
 cacacaggcc cccacacaca ggggaaacaca cgcacacgtc cgtccctggc actggccac 1980  
 ttcccaatgc egcccttccc tgctgttagg gtcacatgag gtgtggctt caccatcctc 2040  
 ctggccctgt ggccctcagg agggacacag gagatggggc gggggctcg ctgagggcca 2100  
 ggtcgctatc tagggctggg tgcgtggctg agtccccggg ccaaagctgg tgcccaggc 2160  
 aggcagctgt ggggagctga ctcaggaca ctgctggccc atccccggg gcccctacat 2220  
 cctgggtctt gccacagagg gaatcacccc cagaggcccg agcccagcag gacacagtat 2280  
 tgaccaccca ctccctgtcc agagctgcaaa ctggaggaga gctgtgcggg ggcgcaggac 2340  
 ggggagctgg acgggctgtg gaccacatc accatcttca tcacactttt cctgtcaagc 2400  
 gtgtgctaca gtgccaccat cacccttcc aaggtaaat aataatagct tacgacgtga 2460  
 tcagcctcga ctgtgccttc tagttccag ccatctgtt tttgccttc cccctgcct 2520  
 tccttgaccc tggaaagggtgc cactccact gtccttcctt aataaaatga gggaaattgca 2580  
 tcgtttgtc tgtagtaggtg tcatttttattt ctgggggtt ggggtggggca ggacagcaag 2640  
 ggggaggatt gggaaagacaa tagcaggcat gctggggaaat tctggccgg gcatgataac 2700  
 gtgacagaaa agccccatcc ttaggcctcc tccttccttgc ttcctgtata ttgggtctaa 2760  
 ccccccaccc tcgtttaggca gattttttt ctgggtacac accccctttt cctggagcc 2820  
 tctcttcctt tgccagaacc tctaaagggtt gttttttttt gggccagaga ggatctggg 2880  
 agggagagct tggcagggggg tggggggaa gggggggatg cgtgacccctgc cgggttctca 2940  
 gtggccaccc tgcgttaccc tctcccttgc cctgtttttt cttttttttt cttttttttt 3000

cgtttcactg atcctggc tgcagcttcc ttacacttcc caagaggaga agcagttgg 3060  
 aaaaacaaaa tcagaataag ttggcctga gttctaatt tggcttca ccttctagt 3120  
 ccccaattta tattgttccct ccgtgcgtca gtttacctg tgagataagg ccagtagcca 3180  
 gccccgtctt ggcagggctg tggtgaggag ggggggtcc gtgtggaaaa ctcccttgt 3240  
 gagaatggtg cgtcttaggt gttcaccagg tcgtggccgc ctctactccc ttctctttc 3300  
 tccatccttc ttcccttaaa gagtccccag tgctatctgg gacatattcc tccgcccaga 3360  
 gcagggtccc gcttccctaa ggccctgctc tgggcctctg ggtttgagtc ctggcaagc 3420  
 ccaggagagg cgctcaggct tccctgtccc cttcctctgt ccaccaatctc atgeccctgg 3480  
 ctctcctgcc cttccctac aggggttccct ggctctgctc tcctgcaggc gatctctcga 3540  
 tctctcgatt tcgatcaaga cattccctta atggctttt ctggacacca ctagggtca 3600  
 gaagtagttc atcaaacttt cttccctccc taatctcatt 3640  
 <210> 318  
 <211> 334  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <400> 318  
 Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser  
 [0100] 1 5 10 15  
 Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys  
 20 25 30  
 Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu  
 35 40 45  
 Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu  
 50 55 60  
 Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr  
 65 70 75 80  
 Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val  
 85 90 95  
 Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro  
 100 105 110  
 Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe  
 115 120 125  
 Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val  
 130 135 140

Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe  
 145 150 155 160  
 Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro  
 165 170 175  
 Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr  
 180 185 190  
 Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val  
 195 200 205  
 Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala  
 210 215 220  
 Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg  
 225 230 235 240  
 Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly  
 245 250 255  
 Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro  
 260 265 270  
 Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser  
 [0101] 275 280 285  
 Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln  
 290 295 300  
 Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His  
 305 310 315 320  
 Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
 325 330  
 <210> 319  
 <211> 45  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <400> 319  
 Leu Gln Leu Glu Glu Ser Cys Ala Glu Ala Gln Asp Gly Glu Leu Asp  
 1 5 10 15  
 Gly Leu Trp Thr Thr Ile Thr Ile Phe Ile Thr Leu Phe Leu Leu Ser  
 20 25 30  
 Val Cys Tyr Ser Ala Thr Ile Thr Phe Phe Lys Val Lys

35 40 45  
<210> 320  
<211> 7680  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 320

gcgatcgctg cttctctga cctgcattct ctccccctggg cctgtgccgc ttctgtctg 60  
cagcttgcgg cctgggtcac ctctacggct ggcccaagatc ctccctgccc gcctccctca 120  
ggttccgtct tctccactc ccttcccc ttgtctctgt ctgtgtgtc gcccaaggat 180  
gcctttccg gaggacttcc ttctcggcgc tgcaccacgt gatgtctct gagcggatcc 240  
tccccgtgtc tgggtctctt ccgggcatct ctccctccctt acccaacccc atggcgttt 300  
caactcgctgg gttccctttt ctttcctt ctggggcttg tgccatctct cgtttcttag 360  
gatggccttc tccgacggat gtctcccttg cgtcccgctt ccccttcttg taggcctgca 420  
tcatcaccgt ttttctggac aaccccaaag taccccgctt ccctggcttt agccacctct 480  
ccatccttctt gtttcttctt ctggacacc ccttctctt gtggattcgg gtcacccctc 540  
actcccttca ttgggcage tcccttaccc cccttaccc tctagtctgt gcaagcttt 600  
ccagccccctt gtcacggat ctccagggg tccgagagct cagctgtct tcttcctcca 660  
accggggccc ctatgtccac ttcaaggacag catgtttgtc gcctccagggtt atcctgtgtc 720  
cccgagctgg gaccaccta tattccagg gccggtaat gtggctctgg ttctgggtac 780  
tttatctgt cccctccacc ccacagtggg gcaagatgca tcttctgacc tcttctctc 840  
ctccccacagg gcatggcaaa acctctgagc caggaagaaa gcacactgtat tgaaagagca 900  
accgctacta tcaacagcat ccccatctcc gaagactatt ctgtggcttag tgccgtctg 960  
tccagcgacg ggagaatctt caccgggttg aacgtctacc actttacagg cggaccatgc 1020  
gcagagctgg tggctctggg gactgcagcc gctgcagccg ctggtaatctt gacctgtatc 1080  
gtggccattt gcaacgaaaa taggggcatc ctgtccccat gggcagggtg tcggcagggtg 1140  
ctgctggatc tgcattcttgc catcaaggca attgtcaaag actctgtatgg acagcctacc 1200  
gccgtcggtt tccgtgaact gctgccttgc ggctatgtt gggagggtata atgagcttgg 1260  
cttcgaaaca cacagatcta cacggatcca gcttacgacg tggatcgttgc ctactgtgcc 1320  
ttcttagttgc cagccatctt ttttttgcctt ctccccctgg ctttccttgc ccctggaaagg 1380  
tgccactccc actgtctttt cctataaaaa tgaggaaatt gcatgcattt gtgtgatgtt 1440  
gtgtcattctt attctggggg gtgggggtggg gcaggacacg aagggggagg attgggaaga 1500  
caatagcagg catgtgggg acgtatgtca gctggatcta gtaatcaattt acggggatcat 1560  
tagttcatag cccatatatg gagttccgcgtt tacataact tacggtaat ggccggctgt 1620  
gctgaccggcc caacgacccc cgcccatgtt gctcaataat gacgtatgtt cccatagtt 1680

[0103]

cgccaatagg gactttccat tgacgtcaat gggtggagta tttacggtaa actgccact 1740  
 tggcagtaca tcaagtgtat catatgcca gtacgcccc tatgcacgt aatgacggta 1800  
 aatggcccgc ctggcattat gcccagtaca tgacctatg ggacttct acitggcagt 1860  
 acatctacgt attagtcatc gctattacca tgctgatgcg gtttggcag tacatcaatg 1920  
 ggcgtggata gcggtttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg 1980  
 ggagtttgtt ttggcaccaa aatcaacggg actttccaaa atgtcgtaac aactccgccc 2040  
 cattgacgca aatggcggtt aggcgtgtac ggtgggaggt ctatataagc agagctggtt 2100  
 tagtgaaccg tcagatcaga tccatcgatc taggaattca ccatgccaaa gagacccaga 2160  
 ccctctagat tagataaaag taaagtgatt aacagcgcat tagagctgct taatgaggc 2220  
 ggaatcgaag gtttaacaac ccgtaaactc gcccagaagc tgggtgtaga gcagcctaca 2280  
 ctgtattggc atgtaaaaaa taagcgggct ttgcgcacg ccttagccat tgagatgta 2340  
 gataggcacc atactcactt ttgccttta aaaggggaaa gctggcaaga tttttacgc 2400  
 aataacgcta aaagtttag atgtgcctta ctaagtcatc gcaatggagc aaaagtacat 2460  
 tcagatacac ggcctacaga aaaacagtat gaaactctcg aaaatcaatt accctttta 2520  
 tgccaacaag gttttcact agagaacgcg ttatgcac tcagcgtgt gggcattt 2580  
 acttttaggtt gcgtattgga agatcaagag catcaagtcg ctaaagaaga aagggaaaca 2640  
 cctactactg atagtatgcc gccattatta cgacaagctc tcgaattatt tgatcaccaa 2700  
 ggtgcagagc cagccttctt attcggcctt gaattgatca tatgeggatt agaaaaacaa 2760  
 cttaaatgtt aaagtgggc cgcgtacagc cgcgcgcgt aaaaaacaa ttacgggtct 2820  
 accatcgagg gcctgcctga tctccggac gacgacgccc cgaagaggc gggcgtggcg 2880  
 gctccgcgecc tgctttcttccccgcggca cacacgcgcgac gactgtcgac gccccccccc 2940  
 accgatgtca gcctggggca cgagctccac ttagacggcg aggacgtggc gatggcgcatt 3000  
 gccgacgcgc tagacgattt cgatctggac atgtgggg acggggattt cccgggtccg 3060  
 ggatttaccc cccacgactc egccccctac ggccgtctgg atatggccga cttcgagttt 3120  
 gagcagatgtt acccgatgc ctttggaaatt gacgagtgac gtgggttaggg ggcgcgagga 3180  
 tccagacatg ataagataca ttgtgagggt tggacaaacc acaactagaa tgcagtgaaa 3240  
 aaaatgcattt atttgtgaaa ttgtgatgc tattgcctta ttgttaacca ttataagctg 3300  
 caataaaacaa gtaacaaca acaattgcatttcatatgtt ttcagggttc agggggaggt 3360  
 gtgggaggtt tttaaagca agtaaaacct ctacaaatgt ggtatggctg attatgtacc 3420  
 tgcaaggctc gtcgtccctgg cggaccacg ctatctgtc aaggccccggccgcg 3480  
 cgccgtccat gagcagagcg cccgcgcgcg aggcaagac tcgggcggcg ccctgcccgt 3540  
 cccaccaggt caacaggcgg taaccggctt ctcatcggtt aatgcgcgcg accttcagca 3600  
 tcgcggcat gtccccctgg cggacggaa gatccagct cgaccatgct tggcgagatt 3660  
 ttcaggagct aaggttagctt cgtttcaca cgagttact ccctatcagt gatagagaac 3720

[0104]

gtatgtcgag ttactccct atcagtgata gagaacgatg tcgagttac tccctatcg 3780  
 tgtagagaa cgtatgtcga gttactccc tatcagtgtat agagaacgtt tgtcgagtt 3840  
 actccctatc agtgcatac aacgtatgtc gagtttatcc ctatcagtga tagagaacgt 3900  
 atgtcgagtt tactccctat cagtgcatac gaacgtatgt cgaggtaggc gtgtacggtg 3960  
 ggaggcctat ataaggcatac ctcgttagt gaaccgtcag atcgcctgga ttcaaatgag 4020  
 ggcctggatc ttcttcctcc ttgcctggc cgggagggt ctggcagcta gcgagattgt 4080  
 cctgacccag agccccgcaa cactgtccct gtccccggaa gaaagagcaa ccctgtcctg 4140  
 tagagcatca aaagggtgtt ctaccgtgg gtacagctat ctgcacttgtt accaggcaga 4200  
 gcccgacag gcccctaggc tgctgatcta cctggctct tatctggaga gtggagtgcc 4260  
 agcacggttc tcaggttccg gcagcggaaac agactttacc ctgacaattt ccagcctgga 4320  
 gccagaagac ttgcgggtt actattgcca gcattccaga gatctgcctt tgactttgg 4380  
 cggagggacc aagggtcgaaa tcaaaaggac tgccggcga accgtggctg cccctccgt 4440  
 ttcatcttc cccctcccg acgagcagct gaagtccggc accgcctctg tgggtgcct 4500  
 gctgaacaac ttctaccctc gggaggccaa ggtgcagtgg aagggtggaca acgcctgca 4560  
 gtccggcaac tcccaggaaat ccgtcaccga gcaggacttcc aaggactcta ctactccct 4620  
 gtccctccacc ctgaccctgtt ccaaggccga ctacgagaag cacaagctgt acgcctgca 4680  
 agtgcacccac cagggccctgtt ccttcctccgtt gaccaagtcc ttcaaccggg gcgagtgtc 4740  
 tagacgagca aagagaggca gcccggcgcac caacttttgc ctgctgaaac aggccggcga 4800  
 tggtaagaa aaccctaggac ccgagctcat gagttggagc tgatcatcc ttctttgg 4860  
 agcaacagctt acaggtaagg ggttaacagt agcaggctt aggtctggac atatatatgg 4920  
 gtgacaatga catccactt gccttcctctt ccacaggcgc catggccctt gtgcagctgg 4980  
 tgcagagcgg cgtggaagtg aaaaaggctt gggcatccgtt gaaggctcc ttgaaagcaa 5040  
 gccccgttacac attcaccaac tactatatgtt actgggtgag acaggccacca ggacaggac 5100  
 tggagttggat gggcggaatc aacccttcta atgggggtac aaactcaac gaaaaggta 5160  
 aaaacagggtt cactctgacc acagattcca gcactaccac agcctatatg gagctgaagt 5220  
 ccctccctt cgcgcatacc ggcgttactt attgcgttagt gcccggactac aggttcgata 5280  
 tggctttaatgaa ctattttttttt cagggaacta ccgtcacagt ctgcgttgcctt tccaccaagg 5340  
 gccccttgcgtt ctttccctctt gccccttccctt ccaaggcttac ctgcgttgcctt accgtgtc 5400  
 tggcttgcctt ggtgaaggac tacttccctt ggcgttgcac cgtgttgcctt aactctggcg 5460  
 cccttgcgttcc cggcggttgcac accttccctt ccgttccctt gtcctccggc ctgtacttcc 5520  
 tgcctccgtt ggtgaccgtt ccgttccctt ctctgggcac ccagacccatc atctgcaacg 5580  
 tgaaccacaa cccttccaaac accaagggtgg acaagaagggtt ggaggcttcaag tccgtgcaca 5640  
 agacccacac ctgccttccca tgccttgcctt ctgagctgtt gggcggttccccc tccgttgcctt 5700  
 tgcctccctcc taagecttaag gacacccttgcata tgcctcccttccca gaccccttgcata gtcacccgttgcgtt 5760

[0105]

tgggtggga cgtgtccac gaagatcctg aagtgaagtt caattggta cttggacggcg 5820  
 tggagggtgca caacgcga accaaggcctc ggggagaaca gtacaactcc acctaccggg 5880  
 tgggtctgt gctgaccgtg ctgcaccagg actggctgaa cgccaaagaa tacaagtgc 5940  
 aggtgtccaa caaggccctg cctgccccta tcgaaaagac catctccaag gctaaggcc 6000  
 agccacggga acctcaggta tacacactgc ctccatgccg ggacgagctg accaagaacc 6060  
 aggtgtccct gacctgtctg gtgaagggtt tctacccttc cgatatgcgc gtggagtggg 6120  
 agtctaacgg ccagccgtag aacaactaca agaccacccc tcctgtgtg gactccgacg 6180  
 gctccttctt cctgtactcc aagctgaccg tggacaagtc ccgggtggcag cagggcaacg 6240  
 tgggtctctg ctccgtgtatg cacgaggccc tgcacaacca ctacacccag aagtccctgt 6300  
 cccgtctcc tggcaaggaa caaaaactca ttcagaaga ggatctaat gctgtggcc 6360  
 aggacacgca ggaggcatc gtgggccac actccttgcc cttaagggtg gtggatct 6420  
 cagccatctt ggcctggtg gtgctcacca tcatctccct tatcatctc atcatgttt 6480  
 ggcagaagaa gccacgttag taaaagcttg tcacttgaa agtaatagtt ttccctgcac 6540  
 gggtagata cagcctcgac tgtgccttc agttgccagc catctgtgt ttgcctcc 6600  
 cccgtgcctt ctttgcaccc ggaagggtgcc actcccactg tccttcata ataaaatgag 6660  
 gaaattgtcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag 6720  
 gacagcaagg gggaggatg ggaagacaat agcaggcatg ctggggatgg cccgggcatg 6780  
 ataacttcgt ataatgtatg ctatacgaag ttatgtatac ggcgcgccta ctgggacag 6840  
 gattggtgac agaaaagccc catccttagg ctccttcctt cctagtcctc tgatattggg 6900  
 tctaaccccc accttcgttt aggcagattc ctatcttgt gacacacccc catttcctgg 6960  
 agccatctct ctccttgcca gaaccttaa ggttgctta cgatggagcc agagaggatc 7020  
 ctgggaggga gagctggca gggggggg gggaaaggggg ggtatgcgtg cctgcccgt 7080  
 ttccttcgc caccctgcgc tacccttc cagaacctga gctgcttga cggcgtgtc 7140  
 tggcgcttt cactgatcct ggtgctgcag ctccctaca ctccccaaaga ggagaagcag 7200  
 ttggaaaaaa caaaatcaga ataagtttgtt cctgagttt aactttggctt ctccacctt 7260  
 ctatccccca atttatattt tccctccgtg cgtcagtttt acctgtgaga taaggccagt 7320  
 agccagcccc gtcctggcag ggctgtggtg aggagggggg tgcctgtg gaaaactccc 7380  
 ttgtgagaa tggcgctcc taggtgtca ccaggctgtg gcccctcta ctcccttc 7440  
 ctcttcctt ccttcattcc taaaagatc cccagtgta tctgggacat attcctccgc 7500  
 ccagagcagg gtcccgcttc cctaaggccc tgctcgggc ttctgggttt gaggcttgg 7560  
 caagcccagg agaggcgctc aggctccctt gtccttc ctcgtccacc atctcatgcc 7620  
 ctggcttc ctgcccctt cctacagggtt ttccctggctc tgctctctg cagggatct 7680  
 <210> 321  
 <211> 132

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 321

Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala

1 5 10 15

Thr Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala

20 25 30

Ser Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val

35 40 45

Tyr His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr

50 55 60

Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly

65 70 75 80

Asn Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val

85 90 95

Leu Leu Asp Leu His Pro Gly Ile Lys Ala Ile Val Lys Asp Ser Asp

100 105 110

[0106] Gly Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr

115 120 125

Val Trp Glu Gly

130

<210> 322

<211> 335

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 322

Pro Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu

1 5 10 15

Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln

20 25 30

Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys

35 40 45

Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His

50 55 60

Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg  
65 70 75 80  
Asn Asn Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly  
85 90 95  
Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr  
100 105 110  
Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu  
115 120 125  
Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys  
130 135 140  
Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr  
145 150 155 160  
Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu  
165 170 175  
Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu  
180 185 190  
Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala  
[0107] 195 200 205  
Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly  
210 215 220  
Leu Leu Asp Leu Pro Asp Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala  
225 230 235 240  
Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser  
245 250 255  
Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp  
260 265 270  
Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp  
275 280 285  
Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro  
290 295 300  
His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe  
305 310 315 320  
Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp Glu Tyr Gly Gly  
325 330 335

<210> 323

<211> 17

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 323

Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu

1 5 10 15

Ala

<210> 324

<211> 111

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 324

Leu Ala Arg Leu Ser Pro Arg Ala Pro Gln His Cys Pro Cys Pro Pro

1 5 10 15

Glu Lys Glu Gln Pro Cys Pro Val Glu His Gln Lys Val Cys Leu Pro

20 25 30

[0108] Val Gly Thr Ala Ile Cys Thr Gly Thr Ser Arg Ser Pro Asp Arg Pro

35 40 45

Leu Gly Cys Ser Thr Trp Leu Leu Ile Trp Arg Val Glu Cys Gln His

50 55 60

Gly Ser Gln Val Pro Ala Ala Glu Gln Thr Leu Pro Gln Phe Pro Ala

65 70 75 80

Trp Ser Gln Lys Thr Ser Pro Cys Thr Ile Ala Ser Ile Pro Glu Ile

85 90 95

Cys Pro Leu Leu Ala Glu Gly Pro Arg Ser Lys Ser Lys Gly Leu

100 105 110

<210> 325

<211> 106

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 325

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln

1 5 10 15

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr

20 25 30

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser

35 40 45

Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr

50 55 60

Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys

65 70 75 80

His Lys Leu Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro

85 90 95

Val Thr Lys Ser Phe Asn Arg Gly Glu Cys

100 105

<210> 326

<211> 20

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 326

[0109]

Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val

1 5 10 15

Glu Glu Asn Pro

20

<210> 327

<211> 15

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 327

Met Ser Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr

1 5 10 15

<210> 328

<211> 120

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 328

Met Ala Gln Val Gln Leu Val Gln Ser Gly Val Glu Val Lys Lys Pro

1 5 10 15

Gly Ala Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr

20 25 30

Asn Tyr Tyr Met Tyr Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu

35 40 45

Trp Met Gly Gly Ile Asn Pro Ser Asn Gly Gly Thr Asn Phe Asn Glu

50 55 60

Lys Phe Lys Asn Arg Val Thr Leu Thr Thr Asp Ser Ser Thr Thr Thr

65 70 75 80

Ala Tyr Met Glu Leu Lys Ser Leu Gln Phe Asp Asp Thr Ala Val Tyr

85 90 95

Tyr Cys Ala Arg Arg Asp Tyr Arg Phe Asp Met Gly Phe Asp Tyr Trp

100 105 110

Gly Gln Gly Thr Thr Val Thr Val

115 120

<210> 329

<211> 332

[0110]

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 329

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

1 5 10 15

Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys

20 25 30

Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu

35 40 45

Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu

50 55 60

Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr

65 70 75 80

Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val

85 90 95

Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro

100 105 110

Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe  
115 120 125

Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val  
130 135 140

Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe  
145 150 155 160

Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro  
165 170 175

Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr  
180 185 190

Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val  
195 200 205

Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala  
210 215 220

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg  
225 230 235 240

Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly  
[0111] 245 250 255

Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro  
260 265 270

Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser  
275 280 285

Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln  
290 295 300

Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His  
305 310 315 320

Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
325 330

<210> 330

<211> 23

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 330

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala

1 5 10 15

Pro Glu Leu Leu Gly Gly Pro

20

&lt;210&gt; 331

&lt;211&gt; 10

&lt;212&gt; PRT

&lt;213&gt; 人工序列(Artificial Sequence)

&lt;400&gt; 331

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1 5 10

&lt;210&gt; 332

&lt;211&gt; 50

&lt;212&gt; PRT

&lt;213&gt; 人工序列(Artificial Sequence)

&lt;400&gt; 332

Asn Ala Val Gly Gln Asp Thr Gln Glu Val Ile Val Val Pro His Ser

1 5 10 15

[0112]

Leu Pro Phe Lys Val Val Val Ile Ser Ala Ile Leu Ala Leu Val Val

20 25 30

Leu Thr Ile Ile Ser Leu Ile Ile Leu Ile Met Leu Trp Gln Lys Lys

35 40 45

Pro Arg

50

&lt;210&gt; 333

&lt;211&gt; 6280

&lt;212&gt; DNA

&lt;213&gt; 人工序列(Artificial Sequence)

&lt;400&gt; 333

ggagatctc cccagcatgc ctgttattgt ctccccatc ctcccccttg ctgtcctgcc 60

ccacccacc ccccagaata gaatgacacc tactcagaca atgcgtatgc atttcctcat 120

tttatttagga aaggacagtg ggagttggcac ctcccgaggta caaggaaggc acgggggagg 180

ggcaaacaac agatggctgg caactagaag gcacagtgcg ggctcttagat tattagcact 240

cgcccccgtt gaaggacttg gtcacggag aggacaggcc ctggtgggtc acttcgcagg 300

cgtacagctt gtgcctctcg tagtcggccct tggacagggt cagggtggag gacagggagt 360

[0113]

aggttagatc cttggagtcc tgctcggtga cggattcctg ggagttgccg gactgcaggg 420  
 cgttgtccac ctccactgc accttggct cccgagggtt gaagttgttc agcaggcaca 480  
 ccacagaggc ggtgccggac ttcagctgtc cgtcgagggg aggaaagatg aacacggaag 540  
 gggcgccgc ggtacgttig atttccacct tggtcccttg gccgaacgtc ctgggggtgc 600  
 tccagaagtg ctggcagtag taggtggcga tgcctctgg ctggaggctg ctgtatggta 660  
 aggttagtc ggtaccgcta ccgctaccgc tgaatctgtc tggcacaccc tcagccaggg 720  
 tggtgtgtta gtagatcagc agcttggag ctttacctgg ctctgtctgg taccaagcca 780  
 ggtatgtgt gatgttaccg ctggctctac aggtgtatgt cactctgtca cccacgttgg 840  
 cgctcaggct gcttgggctc tgggtcatct ggtatgttcc tctgtctgt gaaaatatac 900  
 agcaaacatc agtacaacat aaatatctgt gtatgaaaat caccttaat ctgtcttagac 960  
 atgaagaaag aatatgcaat acattttaa aattaggatt taaaatcaa gccaatatac 1020  
 acctatttca gagtccccaa tgaaagaaat tacagattga aagaatatac ccgccttagt 1080  
 ttgtggaaat atttcaccc ggaacccaga gcagcagaaa cccaatgagt tttgtatggca 1140  
 acatggggc cactccggat ctgatctgcg atctgacggt tcactaaacg agctctgtt 1200  
 atataggccct cccacccgtac acggccaccc gacatactcg tttttactcc etatcagtga 1260  
 tagagaacgt atgaagagtt tactccctat cagtgtataga gaacgtatgc agactttact 1320  
 ccctatcagt gatagagaac gtataaggag tttactccct atcagtgtata gagaacgtat 1380  
 gaccagtttta ctccctatca gtgtataga acgtatctac agtttactcc etatcagtga 1440  
 tagagaacgt atatccagtt tactccctat cagtgtataga gaacgttataa gctttagggcg 1500  
 tgtacggtgg ggcctataa aagcagagct cgtttagtga accgtcagat cgcctggagc 1560  
 aattccacaa cacttttgc ttataccaac ttccgttacc acttcttacc ctgttaaaag 1620  
 atcagatcca tcgattggcc accatgagtt ggagctgtat catcccttc ttggtagcaa 1680  
 cagctacagg taaggggtta acagtagcag gcttggggtc tggacatata tatgggtgac 1740  
 aatgacatcc actttgcctt tctctccaca ggcccattgg cccaggtcca actgcaggag 1800  
 agcggtccag gtcgtgtgag acctagccag accctgagcc tgacctgcac cgtgtctggc 1860  
 agcaccctca gcccgtatgg tgtaaactgg gtgagacage caccctggacg aggtcttgag 1920  
 tggattggaa tgatttttttttggg tggatggaaac acagactata attcagctct caaatccaga 1980  
 gtgacaatgc tggtagacac cagcaagaac cagttcagcc tgagacttag cagcgtgaca 2040  
 gcccggaca cccgggtcta ttattgtgca agagagagag attataggct tgactactgg 2100  
 ggtcaaggca gcctcgac agtctcgatg gcctccacca agggccctag cgtctttcc 2160  
 ctggccctt cctccaagtc tacctctggc ggcaccgtg ctctgggtcg cctgggtgaag 2220  
 gactacttcc ctgagccigt gaccgtgtcc tggaaactctg ggcctgtac ctccggcggt 2280  
 cataccctcc ctgcggctcc ccagtcctcc ggcctgtact ccctgtctc cgtgggtgacc 2340  
 gtgccttccct cctctctggg cacccagacc tacatctgca acgtgaacca caagecttcc 2400

[0114]

aacaccaagg tggacaagaa ggtggagcct aagtccctgcg acaagaccca cacctgcct 2460  
 ccatgtccctg cccctgagct gctggggcgga ccctccgtgt tcctgtcccc tcctaagcct 2520  
 aaggacaccc tcatgtatctc ccggaccctt gaagtgacct gcgtgggtt ggacgtgtcc 2580  
 cacgaagatc ctgaagtgaa gtcaattgg tacgtggacg gcgtggagggt gcacaacgcc 2640  
 aagaccaagc ctcgggagga acagtacaac tccacctacc ggggtgggttc tgtgtgacc 2700  
 gtgtgcacc aggactggct gaacggcaaa gaatacaagt gcaagggttc caacaaggcc 2760  
 ctgcctgccc ctatcgaaaa gaccatctcc aaggctaagg gccagccacg ggaacctcag 2820  
 gtctacacac tgcctcttag ccgggacgag ctgaccaaga accaggtgtc cctgacctgt 2880  
 ctggtaagg gcttctaccc ttccgatatac gccgtggagt gggagtcata cggccagcct 2940  
 gagaacaact acaagaccac ccctccgttg ctggactccg acggctccctt ctccctgtac 3000  
 tccaagctga ccgtggacaa gtcccggtgg cagcagggca acgtttctc ctgctccgtg 3060  
 atgcacgagg ccctgcacaa ccactacacc cagaagtcctc tgccctgtc tcctggcaag 3120  
 gaacaaaaac tcaatctcaga agaggatctg aatgtgtgg gccaggacac gcaggaggc 3180  
 atcgtggtgc cacactcctt gccccttaag gtgggtggta ttcagccat cttggccctg 3240  
 gtgggtgtca ccatcatctc ctttatcatc cttcatcatc ttggcagaa gaagccacgt 3300  
 tagtaaaagc ttgtcacitg gaaagtaata gttttccctg cacgggttagt aatcagccctc 3360  
 gactgtgcct tctagttgcc agccatctgt tgtttgcctt tcccccgtgc ctcccttgac 3420  
 ccttggagggt gcccactccca ctgtcccttc ctaataaaat gagggaaattt catgcattt 3480  
 tctgagtagg tgcattcta ttctgggggg tgggggtgggg caggacagca agggggaggaa 3540  
 ttgggaagac aatagcaggc atgctggggta tggcccgggc atgataactt cgtataatgt 3600  
 atgcataacg aagtatgtt gatgttca tacaaaagga ctgccttcgtt cttttggggaa 3660  
 tcccaaggac cgtcgtaaaa ctcccactaa cgtagaaccc agagatcgct gctttccgc 3720  
 cccctcaccc gcccgcctc gtcatactg agtggaggaa gagcatgcgtt gaggctccgg 3780  
 tgcccggtcag tggcagagc gcacatcgcc cacatcccc gagaagttgg ggggggggtt 3840  
 cgccaatttga accgggtgcctt agagaagggtg ggcgggggtt aactggggaa gtatgtcg 3900  
 gtactggctc cgccttttc ccgggggtgg gggagaaccg tatataatgtt cagtagtcgc 3960  
 cgtgaacgtt cttttcgca acgggtttgc cggccagaaca cagttttttt ccgtgtgtgg 4020  
 ttcccgccggg cttggccctt ttacgggtta tggcccttgc gtccttggaa ttactccac 4080  
 gcccctggctt gcaatgtttt attttttttt ccggatccgtt ggttggaaatgggggggg 4140  
 gttcgaggcc ttgcgttaa ggagccctt cgcctgtgc ttggatgttgc gcttggcttgc 4200  
 ggcgtgggg ccggccgtt cgaatgtttt ggcacccgtt cgcctgttc gtcgtttttt 4260  
 ataagtctct agccatttaa aatttttttt gacccgttgc gacgtttttt ttctggcaag 4320  
 atatgtttttt aaatgcgggc catgtatctc acactggat ttccgggtttt gggccgggg 4380  
 gggccgttgcg gggccgttgc tccctggca catgttgcgc gaggccggggc ctgcgttgc 4440

[0115]

ggccaccgag aatcgacgg gggtagtctc aagctggccg gcctgctcg gtgcctggcc 4500  
tcgcgccccc gtgtatcgcc cgcgcctggg cgcaaggct gcccggctc gcaccagt 4560  
cgtagcggaa aagatggcg ctccccggcc ctgctgcagg gagctcaaaa tggaggacgc 4620  
ggcgctcgaa agagcggcg ggtgagtcac ccacacaaag gaaaaggcc ttccgtct 4680  
cagccgtcgc ttcatgtac tccacggagt accgggcgcc gcaggcac ctgcattag 4740  
tcacgagctt ttggagtacg tcgtcttag gtggggggg ggggtttat gcgatggagt 4800  
ttccccacac tgagtgggtg gagactgaag ttaggccagc ttggcacttg atgtaattct 4860  
ccttggaaat tgccctttt gagttggat ctggttcat tctcaagcct cagacagtgg 4920  
ttcaaagttt ttcttcca ttccagggtt ctgagagct cgtttagtga accgtcagat 4980  
cgccctggaga cgccatccac gctgtttga cctccataga agacacccggg accgatccag 5040  
cccccgcgc cccgaattca ccatgtctag actggacaag agcaaagtca taaactctgc 5100  
tcttggaaat tcaatggag tcggatcgaa aggctgtacg acaaggaaac tcgtcaaaa 5160  
gctgggagtt gagcagccctt ccctgtactg gcaactgtaa aacaagcggg ccctgtcg 5220  
tgccctgcca atcgagatgc tggacaggca tcataccac tcctgcccc tggaaaggcg 5280  
gtcatggcaa gactttctgc ggaacaacgc caagtctac cgctgtgctc tcctctcaca 5340  
tcgcgacggg gctaaagtgc atcggcac ccggccaaaca gagaacacgt acgaaacct 5400  
ggaaaaatcag ctgcgttcc tgtgtcagca aggttctcc ctggagaacg cactgtacg 5460  
tctgtccgcctt gggccact ttacactggg ctgcgtattt gggaaacagg agcatcaagt 5520  
agcaaaagag gaaagagaga cacctaccac cgattctatg cccctactt taaaacaacg 5580  
aattgagctt ttcgaccggc agggagccga acctgccttc ctggcccg tggaaactaat 5640  
catatgtggc ctggagaaac agctaaagtgc cggaaacgc gggccgaccg acgcccctga 5700  
cgattttgac ttagacatgc tcccgccga tggccctgac gactttgacc ttgatatgt 5760  
gcctgtgac gctctgtacg attttgaccc tgacatgctc cccggtaac taagtaagga 5820  
tccagacatg ataagataca ttgatgagtt tggacaaacc acaactagaa tgcagtgaaa 5880  
aaaatgcattt atttgtgaaa ttgtgtatgc tattgttttta ttgttaacca ttataagctg 5940  
caataaacaat gttacaacaaca acaattgcattt tttcagggttcc agggggaggt 6000  
gtgggaggtt tttaaagca agttaaaacctt ctacaaatgtt ggtatggctt attatgtatcc 6060  
tgcaaggccctc gtcgtcttgg ccggaccacg ctatctgtgc aaggccccgg gccccggacg 6120  
cgcgccat gggcggcgcc cccggccggc agggcgaagac tggggccggcc ccctggccgt 6180  
cccaccaggtaacaggccctt cttcatcgaa aatgcgcggc accttcagca 6240  
tcggccggcat gtccccctgg cggacggaa gtatgtatac 6280  
<210> 334  
<211> 103  
<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 334

Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser

1 5 10 15

Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu

20 25 30

Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser

35 40 45

Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu

50 55 60

Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Leu

65 70 75 80

Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys

85 90 95

Ser Phe Asn Arg Gly Glu Cys

100

<210> 335

[0116]

<211> 109

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 335

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gly Asn Ile His Asn Tyr

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile

35 40 45

Tyr Tyr Thr Thr Leu Ala Asp Gly Val Pro Ser Arg Phe Ser Gly

50 55 60

Ser Gly Ser Gly Thr Asp Tyr Thr Phe Thr Ile Ser Ser Leu Gln Pro

65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln His Phe Trp Ser Thr Pro Arg

85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr

100 105  
<210> 336  
<211> 4  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 336  
Ala Ser Arg Gly  
1  
<210> 337  
<211> 15  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 337  
Met Leu Pro Ser Gln Leu Ile Gly Phe Leu Leu Leu Trp Val Pro  
1 5 10 15  
<210> 338  
<211> 15  
[0117] <212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 338  
Met Ser Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr  
1 5 10 15  
<210> 339  
<211> 4  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 339  
Gly Ala Met Ala  
1  
<210> 340  
<211> 115  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 340

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg Pro Ser Gln  
1 5 10 15  
Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Thr Phe Ser Gly Tyr  
20 25 30  
Gly Val Asn Trp Val Arg Gln Pro Pro Gly Arg Gly Leu Glu Trp Ile  
35 40 45  
Gly Met Ile Trp Gly Asp Gly Asn Thr Asp Tyr Asn Ser Ala Leu Lys  
50 55 60  
Ser Arg Val Thr Met Leu Val Asp Thr Ser Lys Asn Gln Phe Ser Leu  
65 70 75 80  
Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala  
85 90 95  
Arg Glu Arg Asp Tyr Arg Leu Asp Tyr Trp Gly Gln Gly Ser Leu Val  
100 105 110  
Thr Val Ser  
115  
<210> 341  
[0118] <211> 331  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 341  
Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser  
1 5 10 15  
Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp  
20 25 30  
Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr  
35 40 45  
Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr  
50 55 60  
Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln  
65 70 75 80  
Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp  
85 90 95  
Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro

100 105 110  
Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro  
115 120 125  
Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr  
130 135 140  
Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn  
145 150 155 160  
Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg  
165 170 175  
Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val  
180 185 190  
Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser  
195 200 205  
Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys  
210 215 220  
Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp  
225 230 235 240  
[0119] Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe  
245 250 255  
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu  
260 265 270  
Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe  
275 280 285  
Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly  
290 295 300  
Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr  
305 310 315 320  
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
325 330  
<210> 342  
<211> 23  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 342

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala  
1 5 10 15  
Pro Glu Leu Leu Gly Gly Pro  
20  
<210> 343  
<211> 10  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 343

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  
1 5 10  
<210> 344  
<211> 50  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 344

Asn Ala Val Gly Gln Asp Thr Gln Glu Val Ile Val Val Pro His Ser  
[0120] 1 5 10 15  
Leu Pro Phe Lys Val Val Val Ile Ser Ala Ile Leu Ala Leu Val Val  
20 25 30  
Leu Thr Ile Ile Ser Leu Ile Ile Leu Ile Met Leu Trp Gln Lys Lys  
35 40 45  
Pro Arg  
50  
<210> 345  
<211> 248  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 345

Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu  
1 5 10 15  
Leu Asn Gly Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln  
20 25 30  
Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys

35 40 45

Arg Ala Leu Leu Asp Ala Leu Pro Ile Glu Met Leu Asp Arg His His

50 55 60

Thr His Ser Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg

65 70 75 80

Asn Asn Ala Lys Ser Tyr Arg Cys Ala Leu Leu Ser His Arg Asp Gly

85 90 95

Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr

100 105 110

Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu

115 120 125

Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys

130 135 140

Val Leu Glu Glu Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr

145 150 155 160

Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Lys Gln Ala Ile Glu Leu

165 170 175

[0121]

Phe Asp Arg Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu

180 185 190

Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Gly Pro

195 200 205

Thr Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Pro Ala Asp Ala

210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Pro Ala Asp Ala Leu Asp Asp

225 230 235 240

Phe Asp Leu Asp Met Leu Pro Gly

245

<210> 346

<211> 99

<212> DNA

<213> 智人(Homo sapiens)

<400> 346

cagggccggt taatgtggct ctgggtctgg gtacttttat ctgtccctc cacccacag 60

tggggccact agggacagga ttggtgacag aaaagcccc 99



1 5  
<210> 350  
<211> 1040  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 350

gggcatgata acttcgtata atgtatgcta tacgaagttt tgatacggc gcgcctgtca 60  
gggacaagat tagtcacagc agagccccca tccccgttct tcttcctctt ggccacagtg 120  
ttggttcac ttcagtcattt catggcaact tccatctcctt gggttgtcg cctcttgctta 180  
ggatgctcgt ggagggtgggg aataaagggtt cagtgttca ctcatccagg cagagggaaaa 240  
acagaatccg tctgtctgg gctgggttgtt cttttagaca gagtgctgc ctagagtgc 300  
ggaaggcctca ggttccaacc ccaataccac agaaatcagg tgcttgaaaa gtagaggcag 360  
gagggttgaa attcagtgtc acettcatctt acacagtgtac tttagggcca tcttgatagc 420  
ccacatgaga cccagttgca aaatcttgac caaacaaaaa aagtcgtgc tgacttttg 480  
tgtcagtctg ctcttgtgtt ctgttatctt gagtgcttt tctcggtgcg ccttagttt 540  
tttacctgttta gaatgggacc agtggtcacc cctgtccccca gaggactct atggtgagag 600  
gctgcctgtg gaaaacttcc ttatggggc cggcggtgac tgggtggac atcgccctct 660  
atatatacgca gggtcgttgc ttcccgaggc cacactcagg ggtcagegtt ctgtgaccag 720  
tccaggagat cagtcaggc tgctcacattt atcaccaaaag tccctggccc tcccaagagt 780  
tccctcccaa ctgcattttt ttcctctgca tctgtggag gtcccttagag catcctctct 840  
gacagggacc tggacactca ggcctgacgt gtcagccctt cagggtggac ttcctcccg 900  
ccatgcagac accttggtga cctgactctt caggcccttg caggagccgt caggcgatct 960  
ctcgatctctt cgatttcgtt caagacattt cttaatggt ctgttggat caccactagg 1020  
ggtcagaagt agttcatcaa 1040

<210> 351  
<211> 408  
<212> DNA  
<213> 人工序列(Artificial Sequence)  
<400> 351

gccttctctt ccacaggcgc catggccgaa gtgcagctgg tgcagttcttgg agctgagggtg 60  
aaggaggctgg gggcctcagt gacagttcc tgcaggctt ccgggttacac ctttaccaac 120  
cacggatatac gctgggtgcg acaggccctt ggacaaggcc ttgagttggat gggatggaac 180  
agcccttaca atggaaacac aaactatgca cagaggttcc agggcagagt caccatgacc 240  
acagacacat ccacgaacac agcctacatg gagtgagga ccctgacatc tgacgacacg 300

gccatgtatt actgtgccag agatagggat tactatgatg cggggagcta ctggggccag 360  
 ggaaccctgg tcacccgtc gagtgccagc accaaggccc ccagcgta 408  
 <210> 352  
 <211> 127  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <400> 352  
 Glu Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Arg Pro Gly Ala  
 1 5 10 15  
 Ser Val Thr Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asn His  
 20 25 30  
 Gly Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met  
 35 40 45  
 Gly Trp Asn Ser Pro Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Arg Phe  
 50 55 60  
 Gln Gly Arg Val Thr Met Thr Asp Thr Ser Thr Asn Thr Ala Tyr  
 65 70 75 80  
[0124] Met Glu Leu Arg Thr Leu Thr Ser Asp Asp Thr Ala Met Tyr Tyr Cys  
 85 90 95  
 Ala Arg Asp Arg Asp Tyr Tyr Asp Ala Gly Ser Tyr Trp Gly Gln Gly  
 100 105 110  
 Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val  
 115 120 125  
 <210> 353  
 <211> 345  
 <212> DNA  
 <213> 人工序列(Artificial Sequence)  
 <400> 353  
 gctagctct atgagctgac tcagccaccc tcgggttcag tatccccagg acagacggcc 60  
 aggatcacct gctctggaga ttcattgcca aagcaatatg cttatggta ccagcaaaag 120  
 ccaggccagg cccctgtatt agtgcataat aaagacagcg agaggccctc agggatccct 180  
 gagcgattct ctggctccgg ctcaggaca acagtcatcg tgaccatcg tggagtccag 240  
 gcagaagacg aggctgacta ttattgtcaa tcagcagaca gtgataatgc ttttgtctt 300  
 ggaagagggaa ccaagctgac cgtccttaggt cagcccgccgg ccgtt 345

<210> 354

<211> 7390

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 354

gagatctcc ccagcatgcc tgctattgtc ttcccaatcc tccccttgc tgcctgc 60  
caccccaccc cccagaatag aatgacacct actcagacaa tgcgatgcaa ttcttcatt 120  
ttattaggaa aggacagtgg gagtggcacc ttccagggtc aggaaggca cgggggaggg 180  
gcaaacaaca gatggtggc aactagaagg cacagtcgag gctctagatt attaacattc 240  
gccccgggtg aagctcttgg tcacaggct agaaaggccc tggtggtca ctgcagggc 300  
gtacactttg tgcttcgt agtcggcctt gtcagggtc agtgtgtgc tcaggctgt 360  
ggtagagtc ttgtgtcct gtcggtcac gtcgtttgg ctattgccc tctggaggc 420  
gttgtccacc ttccactgca ccttggcttc tctgggttag aagttgtca gcaggcacac 480  
gacagagget gtccagact tcagtcgtc gtcgttaggt ggaaagatga acacgttagg 540  
ggcggccgct gtccgcttga ttccacattt ggtgcctcg ccaaagtca gtggagggtc 600  
gttgtacigc tggcagtagt aggtggcgat atcctcaggc tggaggctgg atattgtcag 660  
ggtaagtcg gtccgtatc tgctggcgt gaatctatcg ggcacgccc tttctttct 720  
gtggccttga tagatcagca gtcaggagc ctgtccaggc ttctgtgtat accaggccag 780  
ctgtctctcg atgttccggc tggccttgc tggtatggc actctgtc tcacgctggc 840  
agacaggctg ctaggctct gtgtcatctg gatgtcttc ctgttagctg aaaatacata 900  
gcaaacatca gtacaacata aatatctgt tatgaaaatc accttaatc ttgttagaca 960  
tgaagaaaga atatgcaata cattttaaa attaggatt taaaatcaag caaaaaatca 1020  
cttattgcag agtccccat gaaagaaatt acagattgaa agaataatc cgccttaggtt 1080  
tgtggaaata ttctcacctg gaacccagag cagcagaaac ccaatgagtt gtgtatggcaa 1140  
catggtggcc actccggatc tgatctgacg gttcactaaa ccagctctgc ttatataagac 1200  
ctccccccgtt acacgcctac cggccatttgg ggtacttagt atcaattacg gggcatttag 1260  
ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct 1320  
gaccgcctaa cgaccccccgc ccattgacgt caataatgac gtatgtccc atagtaacgc 1380  
caatagggac ttccatttga cgtcaatggg tggagtattt acggtaaact gcccacttgg 1440  
cagtagatca agtgttatcat atgccaagta cggccctat tgacgtcaat gacggtaat 1500  
ggccgcctg gcattatgcc cagtacatga cttatggga ctttcctact tggcgttaca 1560  
tctacgtatt agtcatcgctt attaccatgc tgatgcgggtt ttggcgttaca atcaatggc 1620  
gtggatagcg gtttgactca cggggatttc caagtcttca cccatttgc gtcaatggga 1680  
gttgttttg geaccaaata caacgggact ttccaaaatc tcgtaaacaac tccgccttcat 1740

tgacgcaaat gggcggttagg cgtgtacggt gggaggtcta tataaggcaga gctggtttag 1800  
tgaaccgtca gatcagatcc atcgattggc caccatgagt tggagctgtat tcatttcctt 1860  
cttggtagca acagctacag gtaaggggtt aacagtagca ggcttgaggctt ctggacat 1920  
atatgggtga caatgacatc cacttgcctt ttcttcac aggcgccatg gcccagggtgc 1980  
agctgggtga atctggcgga ggactgggtc agcctggcggtt atctctgaga ctgtcttgt 2040  
ccgcccaggcggtt cttcacccatc agctactacg atatccagtg ggtccgacag gcccctggca 2100  
aaggacttga atgggtgtcc agcatcagcc cctctggcca gtccacccatc taccggcgag 2160  
aagtgaaggg cagattcacc atcagccggg acaacagcaa gaacaccctt tacctgcaga 2220  
tgaacagccctt gagagccgag gacaccggcccg tgtactactg cgccagaaga accggcagag 2280  
agtacggcggtt aggctggtac ttgttattt gggccaggg caccctggtc acagtctcgaa 2340  
gcgcctctac aaaggcccccc agcggtttcc cactggctcc ctgttagcaga agcaccagcg 2400  
aatctacacgc cgctctgggc tgcctggtca aggactactt tcctgagccctt gtgaccgtgt 2460  
cctggaaatc tggcgctctg acatctggcg tgcacaccctt tccagccgtg ctgcaaagca 2520  
gccccctgtt cagtctgagc agcgtcggtt cagtccttagt cagctctgtt ggcaccaggaa 2580  
cctacacctttaatgtggac cacaaggctt gcaacaccaa ggtggacaag cgctgtt 2640  
ctaagtacgg ccctccctgtt cctccatgtc ctgcacccatc gtttctggc ggaccctccg 2700  
tgttcctgtt tcctccaaag cctaaggaca ccctgatgtt ctccagaaca cccgaagtgtt 2760  
cctgcgttgtt ggtggacgtt tcacaagagg acccccgaggtt gcagtttaat tggtagtgg 2820  
acggcggttga agtgcacaac gccaagacca agccttagaga ggaacagtttac aacagcacctt 2880  
acagagtggt gtccctgtt acagtgttgc accaggattt gctgaacccggc aaagagtaca 2940  
agtgcacagggtt gtccctgtt ggcctggccaa gcagcatcga gaaaaccatc agcaaggccaa 3000  
agggccagcc taggaaaccc cagggttaca cactgcctcc aagccagaaa gagatgacca 3060  
agaaccagggtt gtccctgttacc tgcctgttca agggcttctt cccttccatc atgcctgtt 3120  
aatgggagag caatggccag ccagagaaca actacaagac cacacccctt gtgtggaca 3180  
gcgcacggctt attcttcctt tacagcaagc tgaccgtggaa caagagcaga tggcaagagg 3240  
gcaacgtgtt cagctgcagc gtgtgttgc accggccctgtca caacagatac acccagaagt 3300  
ccctgtcttctt gagccccggaa caaaaactca ttcagaaga ggttcttcaat gctgtggcc 3360  
aggacacgcacca ggagggttgc acgttgcaccatc cttttaagggtt gttgggtatctt 3420  
cagccacccatc ggccctgggtt gtgttgcacca tcaatccctt tatcatccctt atcatgtttt 3480  
ggcagaagaa gccaacgttagt taaaagcttgc tcaatggaa agtaatagttt ttccctgcac 3540  
gggttagataat cagccctgcac tggcccttctt agttgtccagc catctgtgtt tggccctcc 3600  
cccggttgcctt ctttgcaccctt ggaagggttgc acccccacttgc tcccttcata ataaaatgag 3660  
gaaattgtcat cgcattgtctt ggttgggttgc cttttaagggtt gttgggtatctt 3720  
gacagcaagg gggaggattt ggaagacaat agcaggcatc ctggggatgg cccggccatc 3780

[0127]

ataacttcgt ataatgtatg ctatacgaag ttatgttagag taattcatac aaaaggactc 3840  
gcccctgcct tggggaatcc cagggaccgt cgtaaactc ccactaacgt agaaccgcaga 3900  
gatcgctgcg tcccccccccc ctcacccgcc cgctctcgic atcaactgagg tggagaagag 3960  
calgcgtgag gctccgggtgc ccgtcagtgg gcagagcgca catgcccac agtccccgag 4020  
aagtggggg gaggggtcgg caattgaacc ggtgcctaga gaaggtggcg cggggtaaac 4080  
tgggaaagtg atgtcggtta ctggctccgc cttttcccg agggggggg agaaccgtat 4140  
ataagtgcag tagtcgcgt gaacgttctt ttgcacaacg ggtttgcgc cagaacacag 4200  
gtaagtgcgc tgtgtgggtc ccgcgggcct ggccttta cgggttatgg ccctgcgtg 4260  
cctgaatta ctccacgccc cctggctgca gtacgtgatt ctgtatccc agcttcgggt 4320  
tggaaatggg tggagaggtt cgaggccitg cgcttaagga gcccctcgc ctgcgtgtt 4380  
agtggaggcc tggctgggc gctggggccg ccgcgtgcga atctgggtc accttcgcgc 4440  
ctgtctcgct gcttcgata agtctctagc catttaaaat ttgtatgac ctgctgcgac 4500  
gccttttttc tgccaagata gtcttgaaa tgccggccat gatctgcaca ctggatttc 4560  
gttttttggg gccgcggccg gcgacggggc ccgtgcgtcc cagcgcacat gttcggcgag 4620  
gcggggccctg cgagcgcggc caccgagaat cggacggggg tagtctcaag ctggccgc 4680  
tgctctggtg cctggccctg cgccgcgtg tatgccttc cccctggcgca aggctggc 4740  
ccggcggca ccagttgcgt gagcggaaag atggccgtt cccggccctg ctgcaggag 4800  
ctcaaaatgg aggacgcggc gctcgggaga gcggggggg gatctcacca cacaaaggaa 4860  
aaggccctt ccgtcctcag ccgtcgttc atgtgactcc acggagtacc gggcgccgtc 4920  
caggcacctc gattagtca cgagctttg gatctgcgt tccttaggtt ggggggaggg 4980  
gttttatgcg atggagtttc cccacactga gtgggtggag actgaagtta ggcagctt 5040  
gcacttgatg taattctctt tgaaatttgc ctttttgag ttggatctt gttcattct 5100  
caagcctcag acagtgggtc aaagttttt tctccattt caggtgtcgt gagagctgt 5160  
ttatgtacc gtcagatcgc ctggagacgc catccacgtt gtttgcacctt ccatagaaga 5220  
caccgggacc gatccagcct ccgcggcccc gaattcaaca tggactggac ctggagggtc 5280  
ttctgcgtgc tggctgttagc tccaggtaaa gggccaactg gttccaggc tgaggaaggg 5340  
atttttcca gtttagagga ctgtcattctt ctactgtgc ctctccgcag gtgcactc 5400  
ccaggttcag ctggcgcgtt ctggcagcga gctgaaaaaa cctggccctt ccgtgaaggt 5460  
gtccgtcaag gcttcggctt acacccttac cgacaacaac atggactggg tccgacaggc 5520  
ccctggacaa ggacttgagt ggtggcgca catcaacacc agaagcggcg gcagcatcta 5580  
caacgaagag ttccaggaca gatctcat gaccgtggac aagagcaccc acaccgccta 5640  
catggaaactg agcagcciga gaagcgagga caccgcacc tatcaactgcg ccagaagaaa 5700  
gagctacggc tactacctgg acgagtgggg cgaggaaaca ctggtcacag tgtctagcgc 5760  
cagcacaag ggccttagcg ttccctactt ggctccctgtt agcagaagca ccagcgaatc 5820

[0128]

tacagccgct ctgggctgcc tcgtgaagga ctacttcct gaggcctgtga ccgttagctg 5880  
 gaacagcgga gcactgacaa gggcggtgc cacattcca gccgtgtgc aaagcagcgg 5940  
 cctgtactct ctgagcagcg tcgtacagt gcctagcagc tctctggca cccagaccta 6000  
 cacctgtataat gtggaccaca agcctagcaa caccaagggtg gacaagcgcg tgaaatctaa 6060  
 gtacggccct ctttgtctc catgtcctgc tccagagttt ctggcgac cctccgtgtt 6120  
 cctgtttcct ccaaaggctt aggacacctt gatgatctt agaacacccg aagtgacctg 6180  
 cgtgggggtg gacgttcac aagaggaccc cgagggtgcag ttcaattggt acgtggacgg 6240  
 cgtggaaatgt cacaacgcca agaccaagcc tagagaggaa cagtacaaca gcacccatag 6300  
 agtgggttcc gtgtgtacag tgctgtacca ggattggctg aacggcaaag agtacaatgt 6360  
 caaggtgtcc aacaaggccc tgccaagcag catcgagaaa accatcagca aggccaagg 6420  
 ccagcctagg gaacccagg tttacacact gcctccaagc caagaggaaa tgaccaagaa 6480  
 ccaggtgtcc ctgacctgccc tggtaaggg cttctaccct tccgatatcg ccgtggatg 6540  
 ggagagcaat ggcgcagccag agaacaacta caagaccaca cctccgtgc tggacagcga 6600  
 cggctcatc ttccgttaca gcaagctgac tgtggataag agccgggtgc aagagggcaa 6660  
 cgtgttcagc tgtgtgtac tgacagggc cctgcacaac cactacaccc aagagggcct 6720  
 gtctctgagc cctgaacaaa aactcatctc agaagaggat ctgaatgttg tggccagga 6780  
 cacgcaggag gtcacgttgg tgccacactc ctgccttt aagggtgggg tgatctcagc 6840  
 catctggcc ctgggtgtgc tcaccatcat ctcccttatac atcctcatca tgcttgca 6900  
 gaagaagcca cgttagtaac taagtcgaca tccagacatg ataagataca ttgtatgtt 6960  
 tggacaaacc acaactagaa tgcagtgaaa aaaatgtttt atttgtgaaa ttgtgtatgc 7020  
 tattgtttta ttgttaacca ttataagctg caataaaacaa gttacaaca acaattgtat 7080  
 tcattttatg ttcaggttc agggggaggt gtgggaggtt tttaaagca agtaaaacct 7140  
 ctacaaatgt ggtatggctg attatgtatcc tgcaagcctc gtgtgttgc ccggaccacg 7200  
 ctatctgtgc aaggccccgg gccccggacg cgcgtccat gagcagagcg cccggcccg 7260  
 aggcgaagac tcggggggcg ccctggcg cccaccaggtaacaggcgtaaccggcct 7320  
 ctccatcggtt aatgcgcgcg accttcagca tgcggccat gtcccccgttgcggacggaa 7380  
 gtatgtatac 7390  
 <210> 355  
 <211> 104  
 <212> PRT  
 <213> 人工序列(Artificial Sequence)  
 <400> 355

Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys

1 5 10 15

Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg  
20 25 30  
Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn  
35 40 45  
Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser  
50 55 60  
Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys  
65 70 75 80  
Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr  
85 90 95  
Lys Ser Phe Asn Arg Gly Glu Cys  
100  
<210> 356  
<211> 112  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 356

[0129]

Arg Gly Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser  
1 5 10 15  
Val Gly Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Arg Asn Ile Glu  
20 25 30  
Arg Gln Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Glu Leu  
35 40 45  
Leu Ile Tyr Gln Ala Ser Arg Lys Glu Ser Gly Val Pro Asp Arg Phe  
50 55 60  
Ser Gly Ser Arg Tyr Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu  
65 70 75 80  
Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Asp Pro  
85 90 95  
Pro Leu Thr Phe Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Ala  
100 105 110  
<210> 357  
<211> 15  
<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 357

Met Leu Pro Ser Gln Leu Ile Gly Phe Leu Leu Leu Trp Val Pro

1 5 10 15

<210> 358

<211> 15

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 358

Met Ser Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr

1 5 10 15

<210> 359

<211> 122

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 359

[0130] Gln Val Gln Leu Val Glu Ser 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 Tyr Tyr

20 25 30

Asp Ile Gln Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val

35 40 45

Ser Ser Ile Ser Pro Ser Gly Gln Ser Thr Tyr Tyr Arg Arg Glu 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 Arg Arg Thr Gly Arg Glu Tyr Gly Gly Trp Tyr Phe Asp Tyr

100 105 110

Trp Gly Gln Gly Thr Leu Val Thr Val Ser

115 120

<210> 360

<211> 326

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 360

Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser

1 5 10 15

Arg Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp

20 25 30

Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr

35 40 45

Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr

50 55 60

Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln

65 70 75 80

Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp

85 90 95

Lys Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala

100 105 110

[0131]

Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro

115 120 125

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

130 135 140

Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val

145 150 155 160

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln

165 170 175

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln

180 185 190

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly

195 200 205

Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro

210 215 220

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Lys Glu Met Thr

225 230 235 240

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser

245 250 255

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr

260 265 270

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

275 280 285

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe

290 295 300

Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Tyr Thr Gln Lys

305 310 315 320

Ser Leu Ser Leu Ser Pro

325

<210> 361

<211> 16

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 361

Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Ala Pro Gly

[0132]

1 5 10 15

<210> 362

<211> 119

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 362

Gln Val Gln Leu Val Gln Ser Gly Ser Glu Leu Lys Lys Pro Gly Ala

1 5 10 15

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Asp Asn

20 25 30

Asn Met Asp Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met

35 40 45

Gly Asp Ile Asn Thr Arg Ser Gly Gly Ser Ile Tyr Asn Glu Glu Phe

50 55 60

Gln Asp Arg Val Ile Met Thr Val Asp Lys Ser Thr Asp Thr Ala Tyr

65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Thr Tyr His Cys

85 90 95

Ala Arg Arg Lys Ser Tyr Gly Tyr Tyr Leu Asp Glu Trp Gly Glu Gly

100 105 110

Thr Leu Val Thr Val Ser Ser

115

<210> 363

<211> 325

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 363

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg

1 5 10 15

Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr

20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser

35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser

50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr

65 70 75 80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys

85 90 95

Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Pro Cys Pro Ala Pro

100 105 110

Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys

115 120 125

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val

130 135 140

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp

145 150 155 160

Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr

165 170 175

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp

180 185 190

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu  
195 200 205  
Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg  
210 215 220  
Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Met Thr Lys  
225 230 235 240  
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp  
245 250 255  
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys  
260 265 270  
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser  
275 280 285  
Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser  
290 295 300  
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Glu Ser  
305 310 315 320  
Leu Ser Leu Ser Pro  
[0134] 325  
<210> 364  
<211> 110  
<212> PRT  
<213> 人工序列(Artificial Sequence)  
<400> 364  
Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly  
1 5 10 15  
Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Arg Asn Ile Glu Arg Gln  
20 25 30  
Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Glu Leu Leu Ile  
35 40 45  
Tyr Gln Ala Ser Arg Lys Glu Ser Gly Val Pro Asp Arg Phe Ser Gly  
50 55 60  
Ser Arg Tyr Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro  
65 70 75 80  
Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Asp Pro Pro Leu

85 90 95

Thr Phe Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Ala

100 105 110

<210> 365

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 365

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Arg Asn Ile Tyr Arg Gln

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Glu Leu Leu Ile

35 40 45

Tyr Gln Ala Ser Arg Lys Glu Ser Gly Val Pro Asp Arg Phe Ser Gly

50 55 60

Ser Arg Tyr Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro

[0135] 65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Asp Pro Pro Leu

85 90 95

Thr Phe Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Ala

100 105 110

<210> 366

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 366

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Arg Asn Ile Tyr Arg Gln

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Glu Leu Leu Ile

35 40 45

Tyr Gln Ala Ser Arg Lys Tyr Ser Gly Val Pro Asp Arg Phe Ser Gly

50 55 60

Ser Arg Tyr Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro

65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Asp Pro Pro Leu

85 90 95

Thr Phe Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Ala

100 105 110

<210> 367

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 367

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Arg Asn Ile Tyr Arg Gln

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Glu Leu Leu Ile

[0136]

35 40 45

Tyr Gln Ala Ser Arg Lys Tyr Ser Gly Val Pro Asp Arg Phe Ser Gly

50 55 60

Ser Arg Tyr Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro

65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Ser Pro Pro Leu

85 90 95

Thr Phe Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Ala

100 105 110

<210> 368

<211> 110

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 368

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

1 5 10 15

Asp Arg Val Thr Ile Thr Cys Lys Ala Ser Arg Asn Ile Tyr Arg Gln

20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Glu Leu Leu Ile

35 40 45

Tyr Gln Ala Ser Arg Arg Tyr Ser Gly Val Pro Asp Arg Phe Ser Gly

50 55 60

[0137] Ser Arg Tyr Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro

65 70 75 80

Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Ser Pro Pro Leu

85 90 95

Thr Phe Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Ala

100 105 110

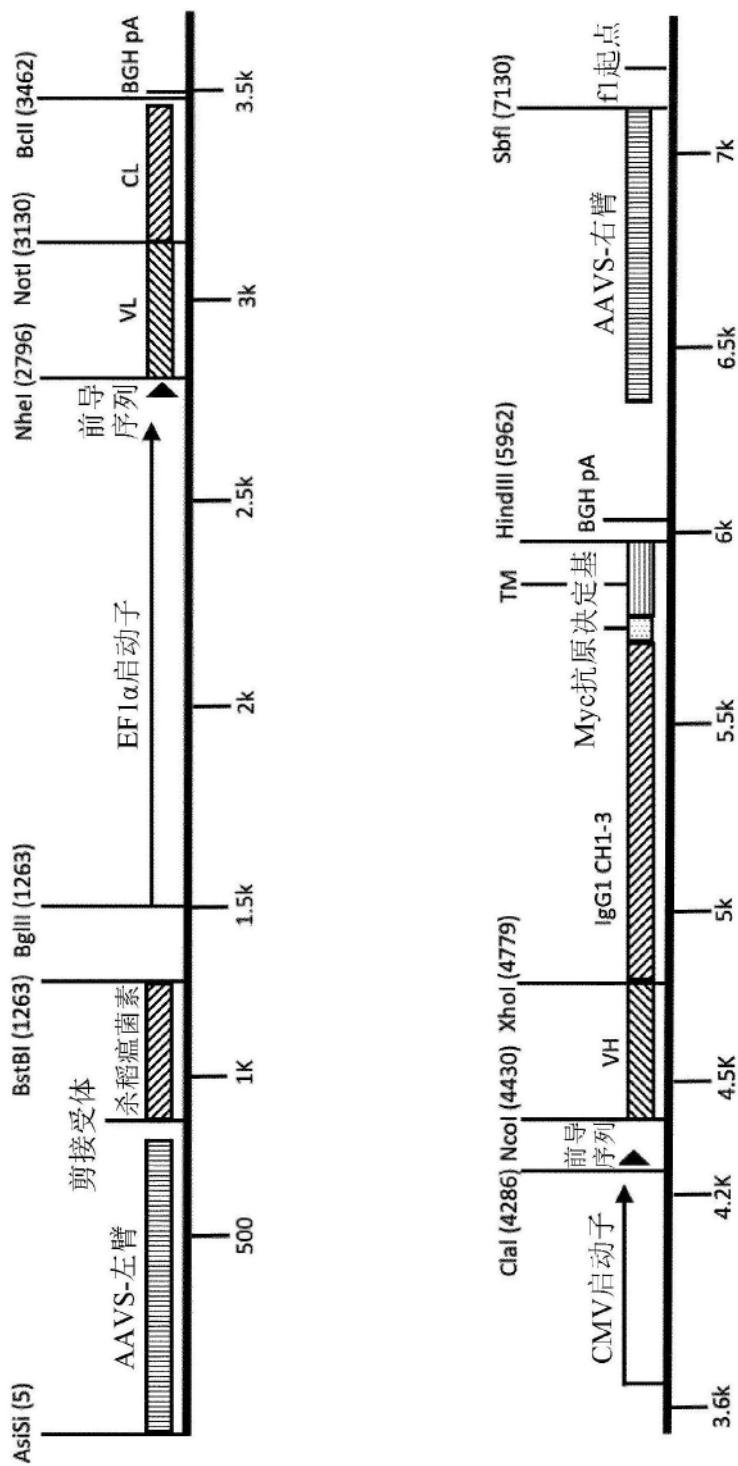


图1a

序列: pINT17-杀稻瘟菌素\_a 范围: 1到10380

>AAVS-左臂

10	20	30	40	50	60
GCG	ATC	GCT	TTC	TCT	GAC
CTG	CAT	CTG	CCC	TGG	GCC
CAG	CTT	GTC	TGG	GCC	GTC
GGT	TCC	GTC	TTC	CTC	ACC
GCT	CTT	TCC	CCC	TTG	CGC
TCC	CCG	TGT	GGT	CTC	CGG
CAC	TCG	GGT	TCC	CTT	CGT
GAT	GGC	CTT	CTC	CGA	CGG
TCA	TCA	CCG	TTT	TTC	TGG
CCA	TCC	TCT	GGG	GCA	ACA
ACT	CCT	TCC	GGT	CGG	CGC
CCA	GCC	CCC	TGT	GCA	GCT
ACC	CGG	GCC	CCT	ATG	TTC
CCC	GAG	CTG	GGA	ACT	GGT
TTT	TAT	CTG	TCC	TCA	GGG

>剪接受体

850	860	870	880	890	900
CTC	CCA	CAG	GGC	ATG	GCA
M	A	K	P	L	S
890	900	910	920	930	940
GAA	AGA	GAC	ATC	CCC	ATC
E	E	E	I	S	D
940	950	960	970	980	990
TCT	TCT	GCT	GCT	ACC	GTC
T	A	T	I	N	S
990	1000	1010	1020	1030	1040
GTG	AGC	GAC	GGG	AGA	ATC
G	S	G	R	I	F
1040	1050	1060	1070	1080	1090
GTC	CTG	GGG	ACT	GCA	GCC
G	C	N	G	G	V
1090	1100	1110	1120	1130	1140
GTG	GCC	ATT	GGC	AAC	GAA
V	A	I	G	N	E
1110	1120	1130	1140	1150	1160
GTC	TCC	GGT	AGG	GGC	ATC
G	S	C	T	G	H
1160	1170	1180	1190	1200	1210
ACC	TCT	GAC	TCT	GAT	GTC
A	C	A	G	C	G
1210	1220	1230	1240	1250	1260
GCC	GTC	GGT	ATC	CGT	GAA
A	V	I	R	E	L
1220	1230	1240	1250	1260	1270
CTG	CTG	CTG	CTG	CTG	CTG
G	G	G	G	G	G
1270	1280	1290	1300	1310	1320
CGC	CTT	CTA	TGA	AAG	GTT
C	C	T	A	A	GGG
1280	1290	1300	1310	1320	1330
CGG	CTT	CTA	TGA	AAG	GTT
G	G	A	A	A	GGG
1330	1340	1350	1360	1370	1380
GGG	CTT	CTA	TGA	AAG	GTT
1340	1350	1360	1370	1380	1390
GGG	CTT	CTA	TGA	AAG	GTT
1390	1400	1410	1420	1430	1440
GGG	CTT	CTA	TGA	AAG	GTT
1400	1410	1420	1430	1440	1450
GGG	CTT	CTA	TGA	AAG	GTT
1450	1460	1470	1480	1490	1500
GGG	CTT	CTA	TGA	AAG	GTT
1460	1470	1480	1490	1500	1510
GGG	CTT	CTA	TGA	AAG	GTT
1510	1520	1530	1540	1550	1560
GGG	CTT	CTA	TGA	AAG	GTT
1520	1530	1540	1550	1560	1570
GGG	CTT	CTA	TGA	AAG	GTT
1570	1580	1590	1600	1610	1620
GGG	CTT	CTA	TGA	AAG	GTT
1580	1590	1600	1610	1620	1630
GGG	CTT	CTA	TGA	AAG	GTT
1630	1640	1650	1660	1670	1680
GGG	CTT	CTA	TGA	AAG	GTT
1640	1650	1660	1670	1680	1690
GGG	CTT	CTA	TGA	AAG	GTT
1690	1700	1710	1720	1730	1740
GGG	CTT	CTA	TGA	AAG	GTT
1700	1710	1720	1730	1740	1750
GGG	CTT	CTA	TGA	AAG	GTT
1750	1760	1770	1780	1790	1800
GGG	CTT	CTA	TGA	AAG	GTT
1760	1770	1780	1790	1800	1810
GGG	CTT	CTA	TGA	AAG	GTT
1810	1820	1830	1840	1850	1860
GGG	CTT	CTA	TGA	AAG	GTT
1820	1830	1840	1850	1860	1870
GGG	CTT	CTA	TGA	AAG	GTT
1870	1880	1890	1900	1910	1920
GGG	CTT	CTA	TGA	AAG	GTT
1880	1890	1900	1910	1920	1930
GGG	CTT	CTA	TGA	AAG	GTT
1930	1940	1950	1960	1970	1980
GGG	CTT	CTA	TGA	AAG	GTT
1940	1950	1960	1970	1980	1990
GGG	CTT	CTA	TGA	AAG	GTT
1990	2000	2010	2020	2030	2040
GGG	CTT	CTA	TGA	AAG	GTT
2000	2010	2020	2030	2040	2050
GGG	CTT	CTA	TGA	AAG	GTT
2050	2060	2070	2080	2090	2100
GGG	CTT	CTA	TGA	AAG	GTT
2060	2070	2080	2090	2100	2110
GGG	CTT	CTA	TGA	AAG	GTT
2110	2120	2130	2140	2150	2160
GGG	CTT	CTA	TGA	AAG	GTT
2120	2130	2140	2150	2160	2170
GGG	CTT	CTA	TGA	AAG	GTT
2170	2180	2190	2200	2210	2220
GGG	CTT	CTA	TGA	AAG	GTT
2180	2190	2200	2210	2220	2230
GGG	CTT	CTA	TGA	AAG	GTT
2230	2240	2250	2260	2270	2280
GGG	CTT	CTA	TGA	AAG	GTT
2240	2250	2260	2270	2280	2290
GGG	CTT	CTA	TGA	AAG	GTT
2290	2300	2310	2320	2330	2340
GGG	CTT	CTA	TGA	AAG	GTT
2300	2310	2320	2330	2340	2350
GGG	CTT	CTA	TGA	AAG	GTT
2350	2360	2370	2380	2390	2400
GGG	CTT	CTA	TGA	AAG	GTT
2360	2370	2380	2390	2400	2410
GGG	CTT	CTA	TGA	AAG	GTT
2410	2420	2430	2440	2450	2460
GGG	CTT	CTA	TGA	AAG	GTT
2420	2430	2440	2450	2460	2470
GGG	CTT	CTA	TGA	AAG	GTT
2470	2480	2490	2500	2510	2520
GGG	CTT	CTA	TGA	AAG	GTT
2480	2490	2500	2510	2520	2530
GGG	CTT	CTA	TGA	AAG	GTT
2530	2540	2550	2560	2570	2580
GGG	CTT	CTA	TGA	AAG	GTT
2540	2550	2560	2570	2580	2590
GGG	CTT	CTA	TGA	AAG	GTT
2590	2600	2610	2620	2630	2640
GGG	CTT	CTA	TGA	AAG	GTT
2600	2610	2620	2630	2640	2650
GGG	CTT	CTA	TGA	AAG	GTT
2650	2660	2670	2680	2690	2700
GGG	CTT	CTA	TGA	AAG	GTT
2660	2670	2680	2690	2700	2710
GGG	CTT	CTA	TGA	AAG	GTT
2710	2720	2730	2740	2750	2760
GGG	CTT	CTA	TGA	AAG	GTT
2720	2730	2740	2750	2760	2770
GGG	CTT	CTA	TGA	AAG	GTT
2770	2780	2790	2800	2810	2820
GGG	CTT	CTA	TGA	AAG	GTT
2780	2790	2800	2810	2820	2830
GGG	CTT	CTA	TGA	AAG	GTT
2830	2840	2850	2860	2870	2880
GGG	CTT	CTA	TGA	AAG	GTT
2840	2850	2860	2870	2880	2890
GGG	CTT	CTA	TGA	AAG	GTT
2890	2900	2910	2920	2930	2940
GGG	CTT	CTA	TGA	AAG	GTT
2900	2910	2920	2930	2940	2950
GGG	CTT	CTA	TGA	AAG	GTT
2950	2960	2970	2980	2990	3000
GGG	CTT	CTA	TGA	AAG	GTT
2960	2970	2980	2990	3000	3010
GGG	CTT	CTA	TGA	AAG	GTT
3010	3020	3030	3040	3050	3060
GGG	CTT	CTA	TGA	AAG	GTT
3020	3030	3040	3050	3060	3070
GGG	CTT	CTA	TGA	AAG	GTT
3070	3080	3090	3100	3110	3120
GGG	CTT	CTA	TGA	AAG	GTT
3080	3090	3100	3110	3120	3130
GGG	CTT	CTA	TGA	AAG	GTT
3130	3140	3150	3160	3170	3180
GGG	CTT	CTA	TGA	AAG	GTT
3140	3150	3160	3170	3180	3190
GGG	CTT	CTA	TGA	AAG	GTT
3190	3200	3210	3220	3230	3240
GGG	CTT	CTA	TGA	AAG	GTT
3200	3210	3220	3230	3240	3250
GGG	CTT	CTA	TGA	AAG	GTT
3250	3260	3270	3280	3290	3300
GGG	CTT	CTA	TGA	AAG	GTT
3260	3270	3280	3290	3300	3310
GGG	CTT	CTA	TGA	AAG	GTT
3310	3320	3330	3340	3350	3360
GGG	CTT	CTA	TGA	AAG	GTT
3320	3330	3340	3350	3360	3370
GGG	CTT	CTA	TGA	AAG	GTT
3370	3380	3390	3400	3410	3420
GGG	CTT	CTA	TGA	AAG	GTT
3380	3390	3400	3410	3420	3430
GGG	CTT	CTA	TGA	AAG	GTT
3430	3440	3450	3460	3470	3480
GGG	CTT	CTA	TGA	AAG	GTT
3440	3450	3460	3470	3480	3490
GGG	CTT	CTA	TGA	AAG	GTT
3490	3500	3510	3520	3530	3540
GGG	CTT	CTA	TGA	AAG	GTT
3500	3510	3520	3530	3540	3550
GGG	CTT	CTA	TGA	AAG	GTT
3550	3560	3570	3580	3590	3600
GGG	CTT	CTA	TGA	AAG	GTT
3560	3570	3580	3590	3600	3610
GGG	CTT	CTA	TGA	AAG	GTT
3610	3620	3630	3640	3650	3660
GGG	CTT	CTA	TGA	AAG	GTT
3620	3630	3640	3650	3660	3670
GGG	CTT	CTA	TGA	AAG	GTT
3670	3680	3690	3700	3710	3720
GGG	CTT	CTA	TGA	AAG	GTT
3680	3690	3700	3710	3720	3730
GGG	CTT	CTA	TGA	AAG	GTT
3730	3740	3750	3760	3770	3780
GGG	CTT	CTA	TGA	AAG	GTT
3740	3750	3760	3770	3780	3790
GGG	CTT	CTA	TGA	AAG	GTT
3790	3800	3810	3820	3830	3840
GGG	CTT	CTA	TGA	AAG	GTT
3800	3810	3820	3830	3840	3850
GGG	CTT	CTA	TGA	AAG	GTT
3850	3860	3870	3880	3890	3900
GGG	CTT	CT			

>EF1a启动子

1510                  1520                  1530                  1540                  1550                  1560  
 GCA TTC TAG TTG TGG AGA TCT CGT GAG GCT CCG GTG CCC GTC AGT GGG CAG AGC GCA CAT  
 CGC CCA CAG TCC CCG AGA AGT TGG GGG GAG GGG TCG GCA ATT GAA CCG GTG CCT AGA GAA  
 GGT GGC GCG GGG TAA ACT GGG AAA GTG ATG TCG TGT ACT GGC TCC GCC TTT TTC CCG AGG  
 GTG GGG GAG AAC CGT ATA TAA GTG CAC TAG TCG CCG TGA ACG TTC TTT TTC GCA ACG GGT  
 TTG CCG CCA GAA CAC AGG TAA GTG CCG TGT GTG GTC CCC GCG GGC CTG GCC TCT TTA CGG  
 GTT ATG GCC CTT GCG TGC CTT GAA TTA CTT CCA CCT GGC TGC AGT ACG TGA TTC TTG ATC  
 CCG AGC TTC GGG TTG GAA GTG GGT GGG AGA GTT CGT GGC CTT GCG CCT TAA AGG GAG CCC CTT  
 CGC CTC GTG CTT GAG TTG TGG CCT GGC CTG GGC GCT GGG GCC GCC GCG TGC GAA TCT GGT  
 GGC ACC TTC GCG CCT GTC TCG CTG CTT TCG ATA AGT CTC TAG CCA TTT AAA ATT TTT GAT  
 GAC CTG CTG CGA CGC TTT TTT TCT GGC AAG ATA GTC TTG TAA ATG CGG GCC AAG ATC AGC  
 ACA CTG GTA TTT CGG TTT TTG GGG CCG CGG GCG GCG ACG GGG CCC GTG CGT CCC AGC GCA  
 CAT GTT CGG CGA GGC GGG GCC TGC GAG CGC GGC CAC CGA GAA TCG GAC GGG GGT AGT CTC  
 AAG CTG CCC GGC CTG CTC TGG TGC CTG CGC TCG CGC CGC CGT GTC TGA TCG CCC CGC CCT GGG  
 CGG CAA GGC TGG CCC GGT CGG CAC CAG TTG CGT GAG CGG AAA GAT GGC CGC TTC CCG GCC  
 CTG CTG CAG GGA GCA CAA AAT GGA GGA CGC GGC GCT CGG GAG AGC GGG CGG GTG AGT CAC  
 CCA CAC AAA GGA AAA GGG CCT TTC CGT CCT CAG CCG TCG CTT CAT GTG ACT CCA CGG AGT  
 ACC GGG CGC CGT CCA GGC ACC TCG ATT AGT TCT CCA GCT TTT GGA GTA CGT CGT CTT TAG  
 GTT GGG GGG AGG GGT TTT ATG CGA TGG AGT TTC CCC ACA CTG AGT GGG TGG AGA CTG AAG  
 TTA GGC CAG CTT GGC ACT TGA TGT AAT TCT CCT TGG AAT TTG CCC TTT TTG AGT TTG GAT  
 CTT GGT TCA TTC TCA AGC CTC AGA CAG TGG TTC AAA GTT TTT TTC CAT TTC AGG TGT

2710                  2720                  2730                  2740                  2750                  2760  
 CGT GAA AAC TAC CCC TAA AAG CCA AAA GAT CCG GAG TGG CCA CCA TGA GGG CCT GGA TCT  
 M R A W I >  
b b 前导 b b >  
 序列  
 NheI  
 |  
 2770                  2780                  2790                  2800                  2810                  2820  
 TCT TTC TCC TTT GCC TGG CCG GGA GGG CTC TGG CAG CTA GCG ACA TCC AGA TGA CCC AGA  
 F F L L C L A G R A L A A >  
b b b b b 前导 b b b b >  
 序列  
 S D I Q M T Q >  
c c c VL c c c >

2830                  2840                  2850                  2860                  2870                  2880  
 GCC CAA GCA GCC TGA GCG CCA GCG TGG GTG ACA GAG TGA CCA TCA CCT GTA GAG CCA CGC  
 S P S S L S A S V G D R V T I T C R A S >  
c c c c c c VL c c c c c c >

2890                  2900                  2910                  2920                  2930                  2940  
 GTC ACA TCC ACA ACT ACC TGG CTT GGT ACC AGC AGA AGC CAG GTC AGG CTC CAA AGC TGC  
 G N I H N Y L A W Y Q Q K P G K A P K L >  
c c c c c c c VL c c c c c c >

2950                  2960                  2970                  2980                  2990                  3000  
 TGA TCT ACT ACA CCA CCA CCC TGG CTG ACG GTG TGC CAA GCA GAT TCA GCG GTC GCG GTC  
 L I Y Y T T T L A D G V P S R F S G S G >  
c c c c c c c VL c c c c c c >

3010                  3020                  3030                  3040                  3050                  3060  
 GCG GTA CCG ACT ACA CCT TCA CCA TCA GCA GCC TCC AGC CAG AGG ACA TCG CCA CCT ACT  
 S G T D Y T F T I S S L Q P E D I A T Y >  
c c c c c c c VL c c c c c c >

3070                  3080                  3090                  3100                  3110                  3120  
 ACT GCC AGC ACT TCT GGA GCA CCC CAA GGA CGT TCG GCC AAG GGA CCA AGG TGG AAA TCA  
 Y C Q H F W S T P R T F G Q G T K V E I >  
c c c c c c c VL c c c c c c >

图1b(续)

>NotI

3130 3140 3150 3160 3170 3180  
 AAC GTA CCG CGG CCC CTT CCG TGT TCA TCT TCC CTC CCT CCG ACG AGC AGC TGA AGT  
 K R T>  
c c c c >

P S V F I F P P S D E Q L K>  
d d d d d d d CL d d d d d d d d >

3190 3200 3210 3220 3230 3240  
 CCG GCA CCG CCT CTG TGG TGT GCC TGC TGA ACA ACT TCT ACC CTC GGG AGG CCA AGG TGC  
 S G T A S V V C L L N N F Y P R E A K V>  
d d d d d d d d CL d d d d d d d d >

3250 3260 3270 3280 3290 3300  
 AGT GGA AGG TGG ACA ACG CCC TGC AGT CCG GCA ACT CCC AGG AAT CCG TCA CCG AGC AGG  
 Q W K V D N A L Q S G N S Q E S V T E Q>  
d d d d d d d d CL d d d d d d d d >

3310 3320 3330 3340 3350 3360  
 ACT CCA AGG ACT CTA CCT ACT CCC TGT CCT CCA CCC TGA CCC TGT CCA AGG CCG ACT ACG  
 D S K D S T Y S L S S T L T L S K A D Y>  
d d d d d d d d CL d d d d d d d d >

3370 3380 3390 3400 3410 3420  
 AGA AGC ACA AGC TGT ACG CCT GCG AAG TGA CCC ACC AGG GCC TGT CCT CTC CCG TGA CCA  
 E K H K L Y A C E V T H Q G L S S P V T>  
d d d d d d d d CL d d d d d d d d >

>BGH pA

3430 3440 3450 3460 3470 3480  
 AGT CCT TCA ACC GGG GCG AGT GCT AAT AAG GAT CCA CGA CGT GAT CAG CCT CGA CTG TGC  
 K S F N R G E C>  
d d d CL d d d >

3490 3500 3510 3520 3530 3540  
 CTT CTA GTT GCC AGC CAT CTG TTG TTT GCC CCT CCC CCG TGC CTT CCT TGA CCC TGG AAG  
 GTG CCA CTC CCA CTG TCC TTT CCT AAT AAA ATG AGG AAA TTG CAT CGC ATT GTC TGA GTA  
 GGT GTC ATT CTA TTC TGG GGG GTG GGG TGG GGC AGG ACA GCA AGG GGG AGG ATT GGG AAG

>CMV启动子

3670 3680 3690 3700 3710 3720  
 ACA ATA GCA GGC ATG CTG GGG ACG ATC GTC AGC TGG ATC TAG TAA TCA ATT ACG GGG TCA  
 TTA GTT CAT AGC CCA TAT ATG GAG TTC CGC GTT ACA TAA CTT ACG GTA ATT GGC CGG CCT  
 GGC TGA CGG CCC AAC GAC CCC CGC CCA TTG ACY TCA ATA ATG ACG TAT GTT CCC ATA GTA  
 ACG CCA ATA GGG ACT TTC CAT TGA CGT CAA TGG GTG GAG TAT TTA CGG TAA ACT GCC CAC  
 TTG GCA GTA CAT CAA GTG TAT CAT ATG CCA AGT ACG CCC CCT ATT GAC GTC ATT GAC GGT  
 AAA TGG CCC GCC TGG CAT TAT GCC CAG TAC ATG ACC TTA TGG GAC TTT CCT ACT TGG CAG  
 TAC ATC TAC GTA TTA GTC ATC GCT ATT ACC ATG CTG ATG CGG TTT TGG CAG TAC ATC AAT  
 GGG CGT GGA TAG CGG TTT GAC TCA CGG GGA TTT CCA AGT CTC CAC CCC ATT GAC GTC AAT  
 GGG AGT TTG TTT TGG CAC CAA AAT CAA CGG GAC TTT CCA AAA TGT CGT AAC AAC TCC GCC  
 CCA TTG ACG CAA ATG GGC GGT AGG CGT GTA CGG TGG GAG GTC TAT ATA AGC AGA GCT GGT

>Clal

4270 4280 4290 4300 4310 4320  
 TTA GTG AAC CGT CAG ATC AGA TCC ATC GAT TGG CCA CCA TGA GTT GGA GCT GTA TCA TCC  
M S W S C I I >  
\_e\_e\_e\_前导\_e\_e\_e\_>  
序列

图1b(续)

>内含子

4330	4340		4350	4360	4370	4380
TCT TCT TGG TAG CAA CAG CTA CAG GTA AGG GGT TAA CAG TAG CAG GCT TGA GGT CTG GAC						
L F L V A T A T>						
—e—e—LEAD—e—e—e—>						
>NcoI						
4390	4400	4410	4420	4430	4440	
ATA TAT ATG GGT GAC AAT GAC ATC CAC TTT GCC TCC CTC TCC ACA GGC GCC ATG GCC CAG						
M A Q>						
—f—f—>						
4450	4460	4470	4480	4490	4500	
GTC CAA CTG CAG GAG AGC GGT CCA GGT CTT GTG AGA CCT AGC CAG ACC CTG AGC CTG ACC						
V Q L Q E S G P G L V R P S Q T L S L T>						
—f—f—f—f—f—f—f—f—VH—f—f—f—f—f—f—f—f—f—f—f—f—f—f—f—>						
4510	4520	4530	4540	4550	4560	
TGC ACC GTG TCT GGC AGC ACC TTC AGC GGC TAT GGT GTA AAC TGG GTG AGA CAG CCA CCT						
C T V S G S T F S G Y G V N W V R Q P P>						
—f—f—f—f—f—f—f—f—f—VH—f—f—f—f—f—f—f—f—f—f—f—f—f—f—>						
4570	4580	4590	4600	4610	4620	
GGA CGA GGT CTT GAG TGG ATT GGA ATG ATT TGG GGT GAT GGA AAC ACA GAC TAT AAT TCA						
G R G L E W I G M I W G D G N T D Y N S>						
—f—f—f—f—f—f—f—f—f—VH—f—f—f—f—f—f—f—f—f—f—f—f—f—>						
4630	4640	4650	4660	4670	4680	
GCT CTC AAA TCC AGA GTG ACA ATG CTG GTA GAC ACC AGC AAG AAC CAG TTC AGC CTG AGA						
A L K S R V T M L V D T S K N Q F S L R>						
—f—f—f—f—f—f—f—f—f—VH—f—f—f—f—f—f—f—f—f—f—f—f—>						
4690	4700	4710	4720	4730	4740	
CTC AGC AGC GTG ACA GCC GCC GAC ACC GCG GTC TAT TAT TGT GCA AGA GAG AGA GAT TAT						
L S S V T A A D T A V Y Y C A R E R D Y>						
—f—f—f—f—f—f—f—f—f—VH—f—f—f—f—f—f—f—f—f—f—f—>						
>XbaI						
4750	4760	4770	4780	4790	4800	
AGG CTT GAC TAC TGG GGT CAA GGC AGC CTC GTC ACA GTC TCG AGT GCC TCC ACC AAG GGC						
R L D Y W G Q G S L V T V>						
—f—f—f—f—f—VH—f—f—f—f—f—f—f—>						
S S A S T K G>						
—g—I GG1 CH1-3_g—g—>						
4810	4820	4830	4840	4850	4860	
CCT AGC GTC TTT CCT CTG GCC CCT TCC AAG TCT ACC TCT GGC GGC ACC GCT GCT CTG						
P S V F P L A P S S K S T S G G T A A L>						
—g—g—g—g—g—g—I GG1 CH1-3_g—g—g—g—g—g—g—g—>						
4870	4880	4890	4900	4910	4920	
GGC TGC CTG GTG AAG GAC TAC TTC CCT GAG CCT GTG ACC GTG TCC TGG AAC TCT GGC GCC						
G C L V K D Y F P E P V T V S W N S G A>						
—g—g—g—g—g—g—I GG1 CH1-3_g—g—g—g—g—g—g—g—>						
4930	4940	4950	4960	4970	4980	
CTG ACC TCC GGC GTG CAT ACC TTC CCT GCC GTC CTC CAG TCC TCC GGC CTG TAC TCC CTG						
L T S G V H T F P A V L Q S S G L Y S L>						
—g—g—g—g—g—g—I GG1 CH1-3_g—g—g—g—g—g—g—>						
4990	5000	5010	5020	5030	5040	
TCC TCC GTG GTG ACC GTG CCT TCC TCC TCT CTG GGC ACC CAG ACC TAC ATC TGC AAC GTG						
S S V V T V P S S S L G T Q T Y I C N V>						
—g—g—g—g—g—g—I GG1 CH1-3_g—g—g—g—g—g—g—>						

图1b(续)

5050 5060 5070 5080 5090 5100  
 AAC CAC AAG CCT TCC AAC ACC AAG GTG GAC AAG AAG GTG GAG CCT AAG TCC TGC GAC AAG  
 N H K P S N T K V D K K V E P K S C D K>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g g>  
  
 5110 5120 5130 5140 5150 5160  
 ACC CAC ACC TGC CCT CCA TGT CCT GCC CCT GAG CTG CTG GGC GGA CCC TCC GTG TTC CTG  
 T H T C P P C P A P E L L G G P S V F L>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5170 5180 5190 5200 5210 5220  
 TTC CCT CCT AAG CCT AAG GAC ACC CTG ATG ATC TCC CGG ACC CCT GAA GTG ACC TGC GTG  
 F P P K P K D T L M I S R T P E V T C V>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5230 5240 5250 5260 5270 5280  
 GTG GTG GAC GTG TCC CAC GAA GAT CCT GAA GTG AAG TTC AAT TGG TAC GTG GAC GGC GTG  
 V V D V S H E D P E V K F N W Y V D G V>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5290 5300 5310 5320 5330 5340  
 GAG GTG CAC AAC GCC AAG ACC AAG CCT CGG GAG GAA CAG TAC AAC TCC ACC TAC CGG GTG  
 E V H N A K T K P R E E Q Y N S T Y R V>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5350 5360 5370 5380 5390 5400  
 GTG TCT GTG CTG ACC GTG CTG CAC CAG GAC TGG CTG AAC GGC AAA GAA TAC AAG TGC AAG  
 V S V L T V L H Q D W L N G K E Y K C K>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5410 5420 5430 5440 5450 5460  
 GTG TCC AAC AAG GCC CTG CCT GCC CCT ATC GAA AAG ACC ATC TCC AAG GCT AAG GGC CAG  
 V S N K A L P A P I E K T I S K A K G Q>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5470 5480 5490 5500 5510 5520  
 CCA CGG GAA CCT CAG GTC TAC ACA CTG CCT CCT AGC CGG GAC GAG CTG ACC AAG AAC CAG  
 P R E P Q V Y T L P P S R D E L T K N Q>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5530 5540 5550 5560 5570 5580  
 GTG TCC CTG ACC TGT CTG GTG AAG GGC TTC TAC CCT TCC GAT ATC GCC GTG GAG TGG GAG  
 V S L T C L V K G F Y P S D I A V E W E>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5590 5600 5610 5620 5630 5640  
 TCT AAC GGC CAG CCT GAG AAC AAC TAC TAA GAG ACC ACC CCT CCT GTG CTG GAC TCC GAC GGC  
 S N G Q P E N N Y K T T P P V L D S D G>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5650 5660 5670 5680 5690 5700  
 TCC TTC CTG TAC TCC AAG CTG ACC GTG GAC AAG TCC CGG TGG CAG CAG GGC AAC GTG  
 S F F L Y S K L T V D K S R W Q Q G N V>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>  
  
 5710 5720 5730 5740 5750 5760  
 TTC TCC TGC TCC GTG ATG CAC GAG GCC CTG CAC AAC CAC TAC ACC CAG AAG TCC CTG TCC  
 F S C S V M H E A L H N H Y T Q K S L S>  
 g g g g g g g g g g IGG1 CH1-3 g g g g g g g g g g>

图1b(续)

5770 5780 5790 5800 5810 5820  
 CTG TCT CCT GGC AAG GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT GCT GTG GGC CAG  
 L S P G K>  
IGG1 CH1-3\_g > E Q K L I S E E D L>  
h h h MYC抗原 h h h >  
 决定基 N A V G Q>  
i TM i i >  
 5830 5840 5850 5860 5870 5880  
 GAC ACG CAG GAG GTC ATC GTG GTG CCA CAC TCC TTG CCC TTT AAG GTG GTG GTG ATC TCA  
 D T Q E V I V V P H S L P F K V V V I S>  
i i i i i i i TM i i i i i i i i i i i i i i >  
 5890 5900 5910 5920 5930 5940  
 GCC ATC CTG GCC CTG GTG GTG CTC ACC ATC ATC TCC CTT ATC ATC CTC ATC ATG CTT TGG  
 A I L A L V V L T I I S L I I L I M L W>  
i i i i i i i TM i i i i i i i i i i i i >  
 >HindIII >前聚腺苷酸  
 5950 5960 | 5970 5980 | 5990 6000  
 CAG AAG AAG CCA CGT TAG TAA AAG CTT GTC ACT TGG AAA GTA ATA GTT TTT CCT GCA CGG  
 Q K K P R \* \* >  
i i TM i i i >  
 >BGH\_pA  
 6010 | 6020 6030 6040 6050 6060  
 GTA GTA ATC AGC CTC GAC TGT GCC TTC TAG TTG CCA GCC ATC TGT TGT TTG CCC CTC CCC  
 CGT GCC TTC CTT GAC CCT GGA AGG TGC CAC TCC CAC TGT CCT TTC CTA ATA AAA TGA GGA  
 AAT TGC ATC GCA TTG TCT GAG TAG GTG TCA TTC TAT TCT GGG GGG TGG GGT GGG GCA GGA  
 <BGH\_聚腺苷酸 >loxP  
 6190 6200 6210 6220 | 6230 6240  
 CAG CAA GGG GGA GGA TTG GGA AGA CAA TAG CAG GCA TGC TGG GGA TGG CCC GGG CAT GAT  
 >AAVS右臂  
 6250 6260 6270 6280 6290 6300  
 AAC TTC GTA TAA TGT ATG CTA TAC GAA GTT ATG TAT ACG GCG CGC CCA CTA GGG ACA GGA  
 TTG GTG ACA GAA AAG CCC CAT CCT TAG GCC TCC TCC TTC CTA GTC TCC TGA TAT TGG GTC  
 TAA CCC CCA CCT CCT GTT AGG CAG ATT CCT TAT CTG GTG ACA CAC CCC CAT TTC CTG GAG  
 CCA TCT CTC TCC TTG CCA GAA CCT CTA AGG TTT GCT TAC GAT GGA GCC AGA GAG GAT CCT  
 GGG AGG GAG AGC TTG GCA GGG GGT GGG AGG GAA GGG GGG GAT GCG TGA CCT GCC CGG TTC  
 TCA GTG GCC ACC CTG CGC TAC CCT CTC CCA GAA CCT GAG CTG CTC TGA CGC GGC TGT CTG  
 GTG CGT TTC ACT GAT CCT GGT GCT GCA GCT TCC TTA CAC TTC CCA AGA GGA GAA GCA GTT  
 TGG AAA AAC AAA ATC AGA ATA AGT TGG TCC TGA GTT CTA ACT TTG GCT CTT CAC CTT TCT  
 AGT CCC CAA TTT ATA TTG TTC CTC CGT GCG TCA GTT TTA CCT GTG AGA TAA GGC CAG TAG  
 CCA GCC CCG TCC TGG CAG GGC TGT GGT GAG GAG GGG GGT GTC CGT GTG GAA AAC TCC CTT  
 TGT GAG AAT GGT GCG TCC TAG GTG TTC ACC AGG TCG TGG CCG CCT CTA CTC CCT TTC TCT  
 TTC TCC ATC CTT TCC TTA AAG AGT CCC CAG TGC TAT CTG GGA CAT ATT CCT CCG CCC  
 AGA GCA GGG TCC CGC TTC CCT AAG GCC CTG CTC TGG GCT TCT GGG TTT GAG TCC TTG GCA  
 AGC CCA GGA GAG GCG CTC AGG CTT CCC CCT TGT CCC TCC TCG TCC ACC ATC TCA TGC CCC  
 >Sbf1  
 7090 7100 7110 7120 | 7130 7140  
 TGG CTC TCC TGC CCC TTC CCT ACA GGG GTT CCT GGC TCT GCT CTC CTG CAG GCG ATC TCT  
 >β球蛋白隔离子  
 7150 7160 | 7170 7180 7190 7200  
 CGA TCT CTC GAT TTC GAT CAA GAC ATT CCT TTA ATG GTC TTT TCT GGA CAC CAC TAG GGG  
 TCA GAA GTA GTT CAT CAA ACT TTC TTC CCT CCC TAA TCT CAT TGG TTA CCT TGG GCT ATC

图1b(续)

>fl起点

7270	7280	7290	7300	7310	7320
GAA ACT TAA TTA AGC CAC CTG ACG CGC CCT GTA GCG GCG CAT TAA GCG CGG CGG GTG TGG	TGG TTA CGC GCA CGC TGA CCG CTA CAC TTG CCA GCG CCC TAG CGC CCG CTC CTT TCG CTT	TCT TCC CTT CCT TTC TCG CCA CGT TCG CGG GCT TTC CCC GTC AAG CTC TAA ATC GGG GGC	TCC CTT TAG GGT TCC GAT TTA GTG CTT TAC GGC ACC TCG ACC CCA AAA AAC TTG ATT AGG	GTG ATG GTT CAC GTA GTG GGC CAT CGC CCT GAT AGA CGG TTT TTC GCC CTT TGA CGT TGG	AGT CCA CGT TCT TTA ATA GTG GAC TCT TGT TCC AAA CTG GAA CAA CAC TCA ACC CTA TCT
CGG TCT ATT CTT TTG ATT TAT AAG GGA TTT TGC CGA TTT CGG CCT ATT GGT TAA AAA ATG	AGC TGA TTT AAC AAA ATT TTA AGC CGA ATT TTA ACA AAA TAT TAA CGC TTA CAA TTT ACG	CGT TAA GAT ACA TTG ATG AGT TTG GAC AAA CCA CAA CTA GTT AAT TAA CCA GTC AAG TCA	GCT ACT TGG CGA GAT CGA CTT GTC TGG GTT TCG ACT ACG CTC AGA ATT GCG TCA GTC AAG		

>pUC起点

7870	7880	7890	7900	7910	7920
TTC GAT CTG GTC CTT GCT ATT GCA CCC GTT CTC CGA TTA CGA GTT TCA TTT AAA TCA TGT	GAG CAA AAG GCC AGC AAA AGG CCA GGA ACC GTA AAA AGG CGG CGT TGC TGG CGT TTT TCC	ATA GGC TCC GCC CCC CTG ACG AGC ATC ACA AAA ATC GAC GCT CAA GTC AGA GGT GGC GAA	ACC CGA CAG GAC TAT AAA GAT ACC AGG CGT TTC CCC CTG GAA GCT CCC TCG TGC GCT CTC	CTG TTC CGA CCC TGC CGC TTA CGG GAT ACC TGT CCG CCT TTC TCC CTT CGG GAA GCG TGG	CGC TTT CTC ATA GCT CAC GCT GTA GGT ATC TCA GTT CGG TGT AGG TCG TTC GCT CCA AGC
TGG GCT GTG TGC ACG AAC CCC CCG TTC AGC CCG ACC GCT CGG CCT TAT CCG GTC ACT ATC	GTC TTG AGT CCA ACC CGG TAA GAC ACG ACT TAT CGC CAC TGG CAG CAG CCA CTG GTC ACA	GGA TTA GCA GAG CGA GGT ATG TAG CGC GTG CTA CAG AGT TCT TGA AGT GGT GGC CTA ACT	ACG GCT ACA CTA GAA GAA CAG TAT TTG GTA TCT GCG CTC TGC TGA AGC CAG TTA CCT TCG	GAA AAA GAG TTG GTA GCT CTT GAT CCG GCA AAC AAA CCA CCG CTG GTC GAA GCG GTT TTT	TTG TTT GCA AGC AGC AGA TTA CGC GCA AAA AAG GAT CTC AAG AAG ATC CTT TGA TCT
TTT CTA CGG CGT CTG AGC CTC AGT GGA AGC AAA ACT CAC GTT AAG GGA TTT TGG TCA TGA	GAT TAT CAA AAA GGA TCT TCA CCT AGA TCC TTT TAA ATT AAA AAT GAA GTT TTA AAT CAA	TCT AAA GTA TAT ATG AGT AAA CTT GGT CTG ACA GTT ACC AAT GCT TAA TCA GTG AGG CAC	CTA TCT CAG CGA TCT GTC TAT TTC GTT CAT CCA TAG TTG CAT TTA AAT TTC CGA ACT CTC	CAA GGC CCT CGT CGG AAA ATC TTC AAA CCT TCC GTC CGA TCC ATC TTG CAG GCT ACC TCT	

>Fse1

8890	8900	8910	8920	8930	8940
CGA ACG AAC TAT CGC AAG TCT CTT GGC CGG CCT TGC GCC TTG GCT ATT GCT TGG CAG CGC	CTA TCG CCA GGT ATT ACT CCA ATC CCG AAT ATC CGA GAT CGG GAT CAC CCG AGA GAA GTT				

>Ascl

9010	9020	9030	9040	9050	9060
CAA CCT ACA TCC TCA ATC CCG ATC TAT CCG AGA TCC GAG GAA TAT CGA AAT CGG GGC GCG	CCT GGT GTA CCG AGA ACG ATC CTC TCA GTG CGA GTC TCG AGC ATC CAT ATC GTT GCT TGG	CAG TCA GCC AGT CGG AAT CCA GCT TGG GAC CCA GGA AGT CCA ATC GTC AGA TAT TGT ACT	CAA GCC TGG TCA CGG CAG CGT ACC GAT CTG TTT AAA CCT AGA TAT TGA TAG TCT GAT CGG	TCA ACG TAT AAT CGA GTC CTA GCT TTT GCA AAC ATC TAT CAA GAG ACA GGA TCA GCA GGA	

>Kan-R

9310	9320	9330	9340	9350	9360
GGC TTT CGC ATG ATT GAA CAA GAT GGA TTG CAC GCA GGT TCT CCG GCG GCT TGG GTG GAG	M I E Q D G L H A G S P A A W V E>				
	_____j_____j_____j_____j_____j_____j_____KANR_____j_____j_____j_____j_____j_____j_____j_____>				
9370	9380	9390	9400	9410	9420
AGG CTA TTC GGC TAT GAC TGG GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC	R L F G Y D W A Q Q T I G C S D A A V F>				
	_____j_____j_____j_____j_____j_____j_____j_____KANR_____j_____j_____j_____j_____j_____j_____j_____>				
9430	9440	9450	9460	9470	9480
CGG CTG TCA GCG CAG GGG CGT CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT GCC CTG	R L S A Q G R P V L F V K T D L S G A L>				
	_____j_____j_____j_____j_____j_____j_____j_____KANR_____j_____j_____j_____j_____j_____j_____j_____>				

图1b(续)

图1b(续)

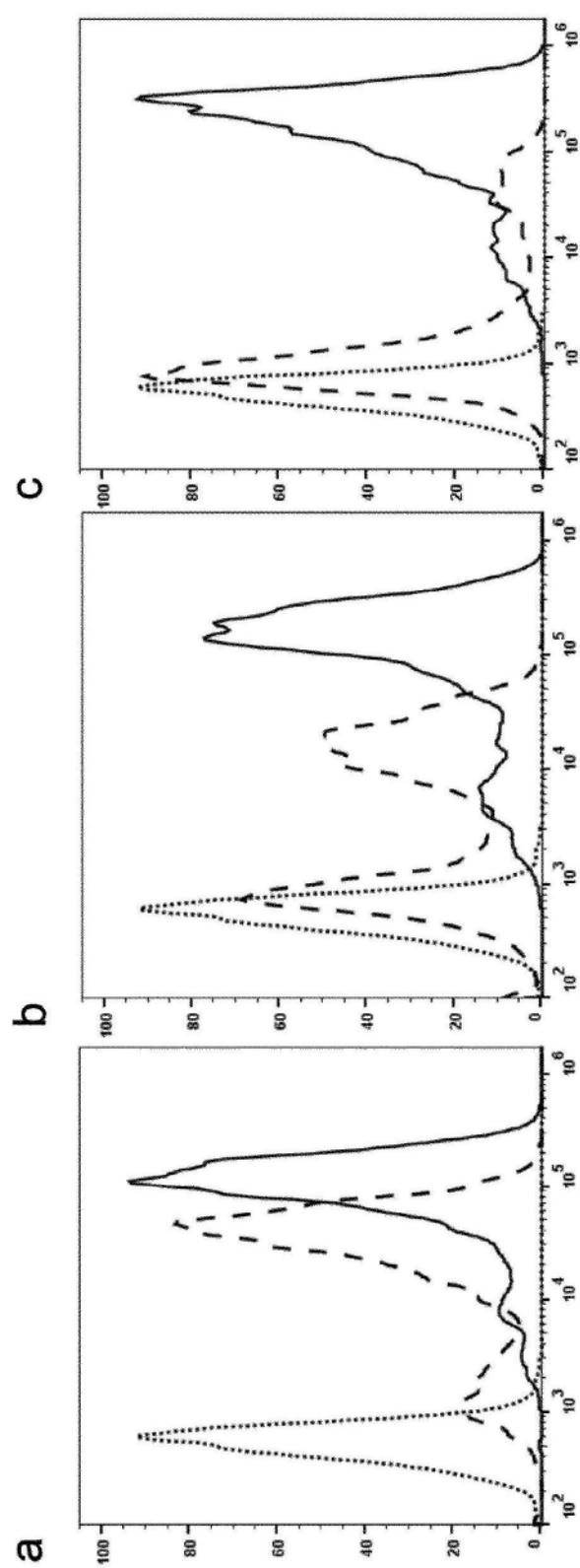


图2

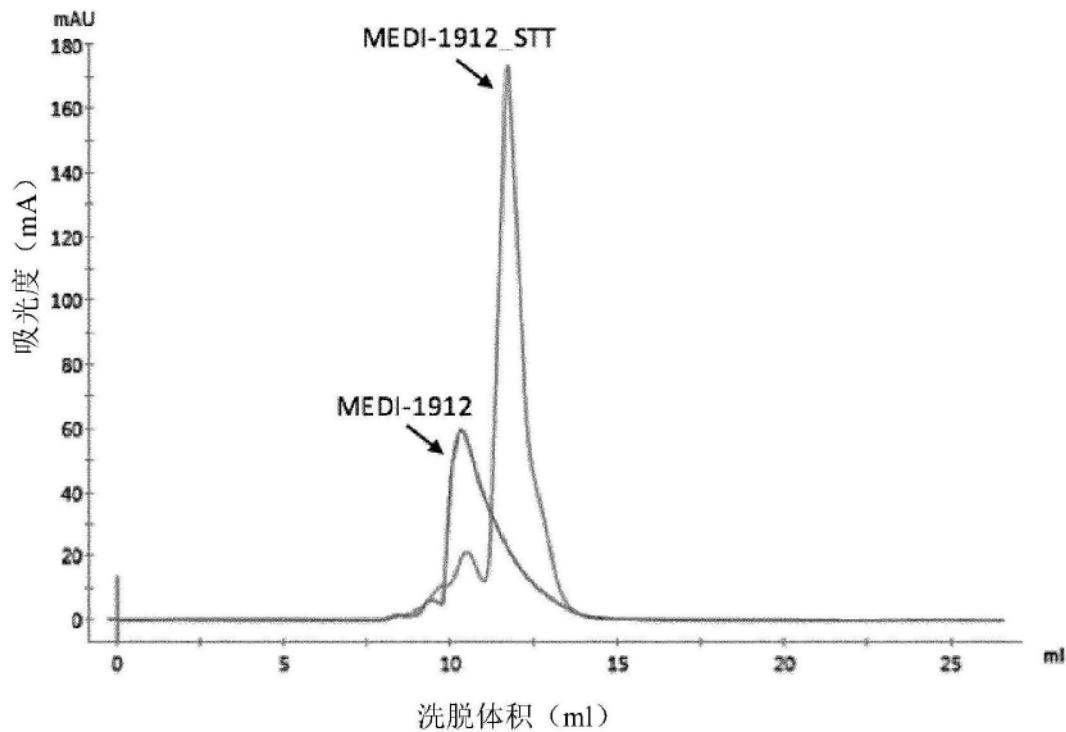


图3

序列: VH-MEDI1912-基因块-opt 范围: 1到392

冬 4

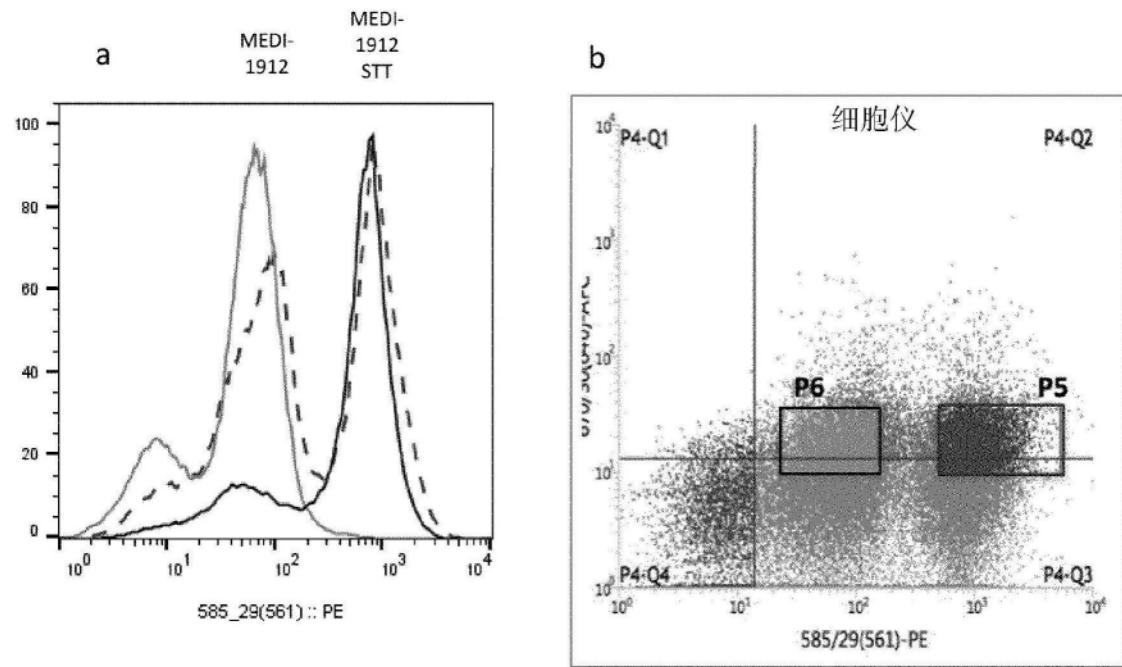


图5

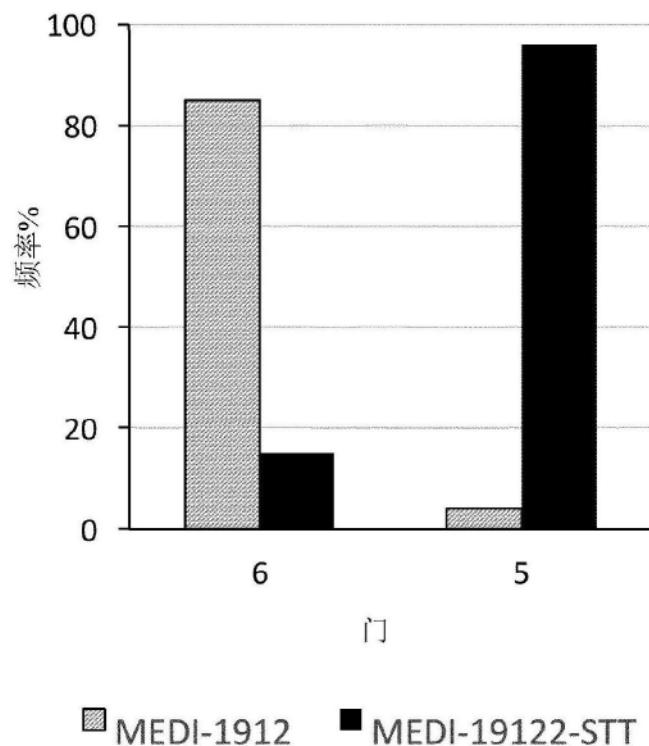


图6

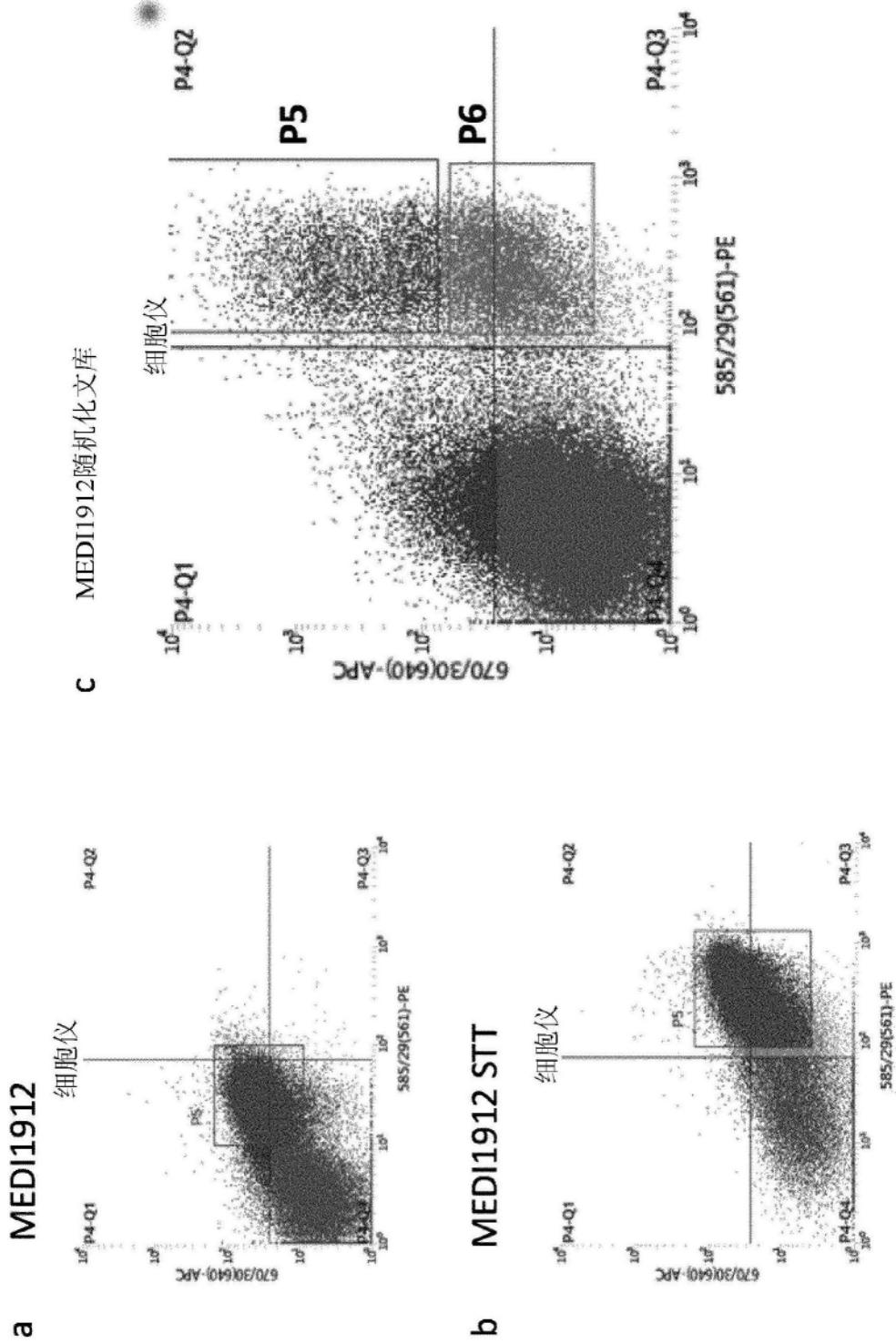


图7

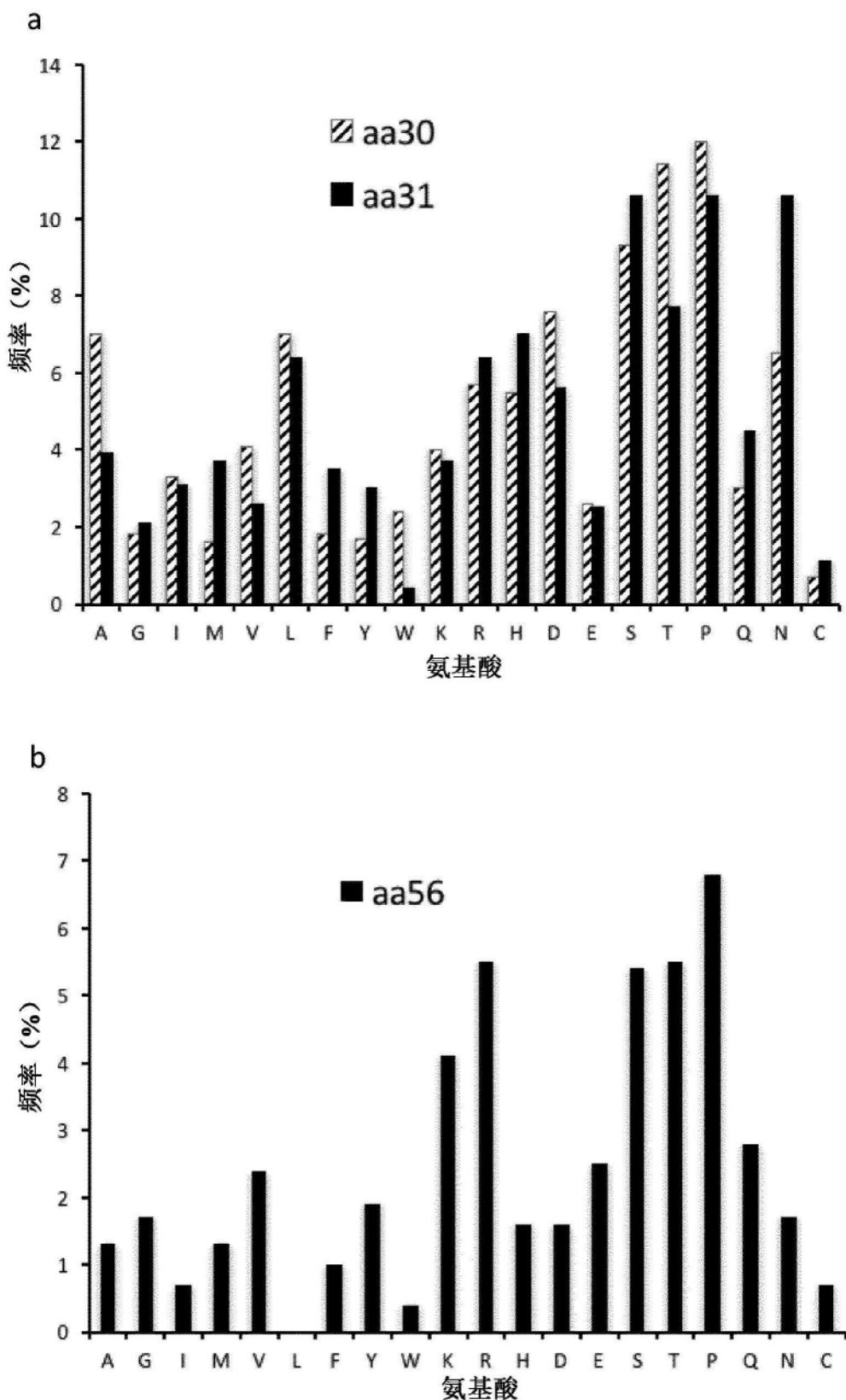


图8

## A. 5A10 VH

5A10 (小鼠)  
5A10-i  
博可珠单抗

10	20	30	<u>CDR1</u>		40	50	<u>CDR2</u>		60
QVQLQQPGAE <del>L</del> VKPG	ASVKLSCKASGYTFT	SYWMHWVKQRPGQGL	EWIGEINPSNGRTNY						
...V.S...VK....	V.....	Y...R.A....	M.....G....						
...V.S...VK....	V.....	Y...R.A....	M...S.FG....						

5A10 (小鼠)  
5A10-i  
博可珠单抗

<u>CDR2</u>	70	80	90	<u>CDR3</u>		100	110
NEKF <del>K</del> SATLTVDKS	SSTAYMQQLSSLTSED	SAVYYCARERPLYAM	DYGQGTTSVTVSS				
....RV.M.R.T.	T..V..E...R...	T.....	.....T....				
....RV.M.R.T.	T..V..E...R...	T.....	S.L....T....				

## B. 5A10 VL

5A10-VL (小鼠)  
5A10-i-VL  
博可珠单抗 VL

10	20	30	<u>CDR1</u>		40	50	<u>CDR2</u>		60	
DIVMTQSHKF <del>M</del> STSV	GDRV <del>S</del> ITCKASQDVS	TAVAWYQQKPGQSPK	LLIYSAS <del>Y</del> RYTGVPD							
...Q....PSSL.A...	....T.....	.....KA...	.....S							
...Q....PSSL.A...	....T...R...GI.	S.L.....KA...	.....S							

5A10-VL (小鼠)  
5A10-i-VL  
博可珠单抗 VL

70	80	90	<u>CDR3</u>		100
RFTGSGSGTDFTFTI	SSVQAEDLAVYYCQQ	RYSTPRTFGGGTKLE	IK		
...S.....	...L.P..I.T....	.....Q.....	..		
...S.....	...L.P..I.T....	...LW....Q.....	..		

图9

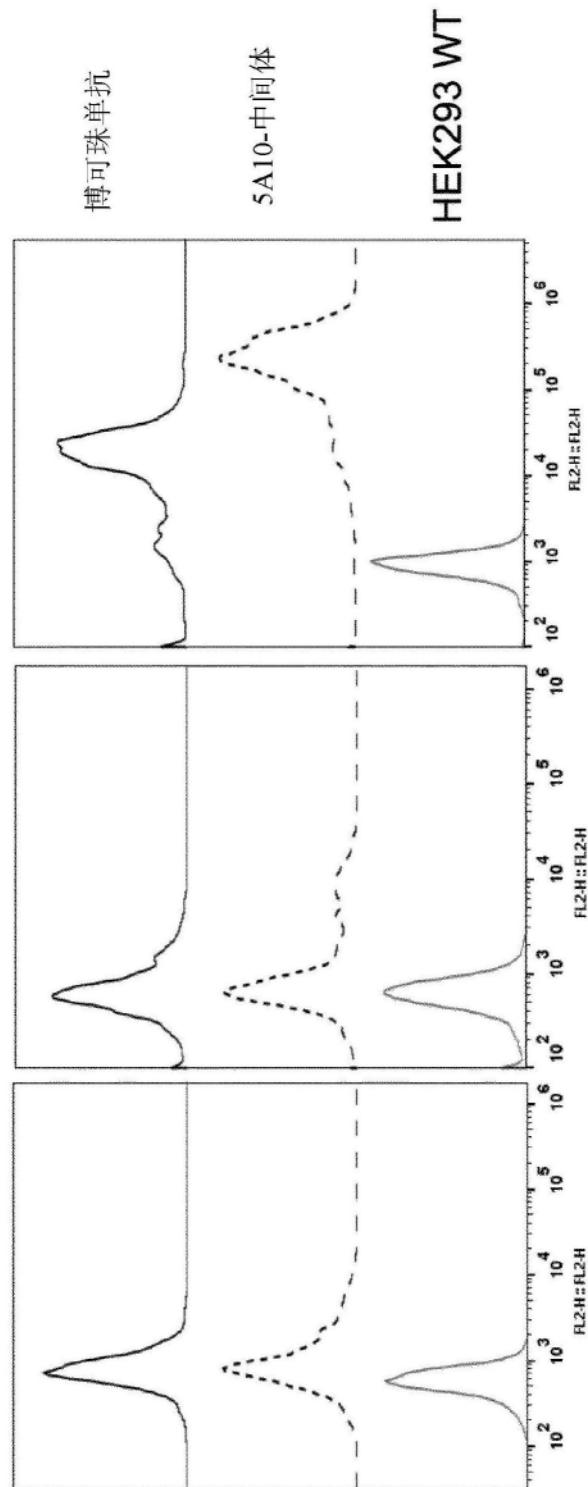


图10

<-----FR1-IMGT-----><-----FR2-IMGT-----><-----FR3-IMGT----->	
查询_1	1
IGHV1-46*01	1
IGHV1-46*01	1
IGHV1-46*03	1
IGHV1-46*03	1
IGHV1-46*02	1
IGHV1-2*02	1
IGHV1-2*05	1
IGHV1-2*01	1
IGHV1-2*04	1
IGHV1-2*03	1
IGHV1-3*01	1
IGHV1-18*01	1
IGHV1-18*04	1
IGHV1-18*03	1
IGHV1-45*01	1
IGHV1-45*02	1
IGHV1-8*01	1

VH	DNA序列
a.Y33A	TTTTTGCCATGGCCCAGGTTCAGCTGGTTCAGTCTGGCGCCGAAGTGAAGAAACCT GGCGCCTCTGTGAAGGTGCTCTGCAAGGCCAGCGGCTACACCTTACCAAGCTAC <b>GCT</b> ATGCACTGGGTCCGACAGGCCCTGGACAAGGACTTGAGTGGATGGCGAGATCAGC CCATTGGCGGCAGGACCAACTACAACGAGAAGTTCAAGAGCCGCGTGACCATGACC AGAGACACCAGCACCTCACCGTGTACATGGAACGTGAGCAGCCTGAGAAGCGAGGAC ACCGCCGTGTACTACTGTGCCAGAGAGAGGCCACTGTACGCCCTGATCTTGGGC CAGGGCACCACCGTGACAGTCTCGAGTTTTTT
b.Y33D	TTTTTGCCATGGCCCAGGTTCAGCTGGTTCAGTCTGGCGCCGAAGTGAAGAAACCT GGCGCCTCTGTGAAGGTGCTCTGCAAGGCCAGCGGCTACACCTTACCAAGCTAC <b>GAT</b> ATGCACTGGGTCCGACAGGCCCTGGACAAGGACTTGAGTGGATGGCGAGATCAGC CCATTGGCGGCAGGACCAACTACAACGAGAAGTTCAAGAGCCGCGTGACCATGACC AGAGACACCAGCACCTCACCGTGTACATGGAACGTGAGCAGCCTGAGAAGCGAGGAC ACCGCCGTGTACTACTGTGCCAGAGAGAGGCCACTGTACGCCCTGATCTTGGGC CAGGGCACCACCGTGACAGTCTCGAGTTTTTT
c. S52N, F54S, R57S	TTTTTGCCATGGCCCAGGTTCAGCTGGTTCAGTCTGGCGCCGAAGTGAAGAAACCT GGCGCCTCTGTGAAGGTGCTCTGCAAGGCCAGCGGCTACACCTTACCAAGCTACTAC ATGCACTGGGTCCGACAGGCCCTGGACAAGGACTTGAGTGGATGGCGAGATC <b>AAC</b> <b>CCATCTGGCGGCAGC</b> ACCAACTACAACGAGAAGTTCAAGAGCCGCGTGACCATGACC AGAGACACCAGCACCTCACCGTGTACATGGAACGTGAGCAGCCTGAGAAGCGAGGAC ACCGCCGTGTACTACTGTGCCAGAGAGAGGCCACTGTACGCCCTGATCTTGGGC CAGGGCACCACCGTGACAGTCTCGAGTTTTTT
d. Y33A, S52N, F54S, R57S	TTTTTGCCATGGCCCAGGTTCAGCTGGTTCAGTCTGGCGCCGAAGTGAAGAAACCT GGCGCCTCTGTGAAGGTGCTCTGCAAGGCCAGCGGCTACACCTTACCAAGCTAC <b>GCT</b> ATGCACTGGGTCCGACAGGCCCTGGACAAGGACTTGAGTGGATGGCGAGATC <b>AAC</b> <b>CCATCTGGCGGCAGC</b> ACCAACTACAACGAGAAGTTCAAGAGCCGCGTGACCATGACC AGAGACACCAGCACCTCACCGTGTACATGGAACGTGAGCAGCCTGAGAAGCGAGGAC ACCGCCGTGTACTACTGTGCCAGAGAGAGGCCACTGTACGCCCTGATCTTGGGC CAGGGCACCACCGTGACAGTCTCGAGTTTTTT
e. Y33D, S52N, F54S, R57S	TTTTTGCCATGGCCCAGGTTCAGCTGGTTCAGTCTGGCGCCGAAGTGAAGAAACCT GGCGCCTCTGTGAAGGTGCTCTGCAAGGCCAGCGGCTACACCTTACCAAGCTAC <b>GAT</b> ATGCACTGGGTCCGACAGGCCCTGGACAAGGACTTGAGTGGATGGCGAGATC <b>AAC</b> <b>CCATCTGGCGGCAGC</b> ACCAACTACAACGAGAAGTTCAAGAGCCGCGTGACCATGACC AGAGACACCAGCACCTCACCGTGTACATGGAACGTGAGCAGCCTGAGAAGCGAGGAC ACCGCCGTGTACTACTGTGCCAGAGAGAGGCCACTGTACGCCCTGATCTTGGGC CAGGGCACCACCGTGACAGTCTCGAGTTTTTT
f. W.T.	TTTTTGCCATGGCCCAGGTTCAGCTGGTTCAGTCTGGCGCCGAAGTGAAGAAACCT GGCGCCTCTGTGAAGGTGCTCTGCAAGGCCAGCGGCTACACCTTACCAAGCTACTAC ATGCACTGGGTCCGACAGGCCCTGGACAAGGACTTGAGTGGATGGCGAGATCAGC CCATTGGCGGCAGGACCAACTACAACGAGAAGTTCAAGAGCCGCGTGACCATGACC AGAGACACCAGCACCTCACCGTGTACATGGAACGTGAGCAGCCTGAGAAGCGAGGAC ACCGCCGTGTACTACTGTGCCAGAGAGAGGCCACTGTACGCCCTGATCTTGGGC CAGGGCACCACCGTGACAGTCTCGAGTTTTTT
博可珠 单抗VL 加终止 密码子	TTTTTGCTAGCGACATCCAGATGACACAGAGCCCTAGCAGCCTGTCTGCCAGCGTG GGAGACAGAGTGCACCATCACCTGTAGAGCCAGGCCAGGGCATCTCTGCCCTGGCA TGGTATCAGCAGAACGCCCTGGCAAGGCCCTAAGCTGCTGATCTACAGCGCCAGC <b>TAA</b> AGATACACCGCGCTGCCAGCAGATTCTGGCAGCGCTCTGGCACCGACTCACC TTCACCATAAGCAGCCTGAGCCTGAGGATATGCCACCTACTACTGCCAGCAGCGG TACTCT <b>TAGTAAC</b> GGACATTGGCCAGGGACCAAGCTGGAAATCAAGCGTACCGCG GCCGCTTTTT

图12

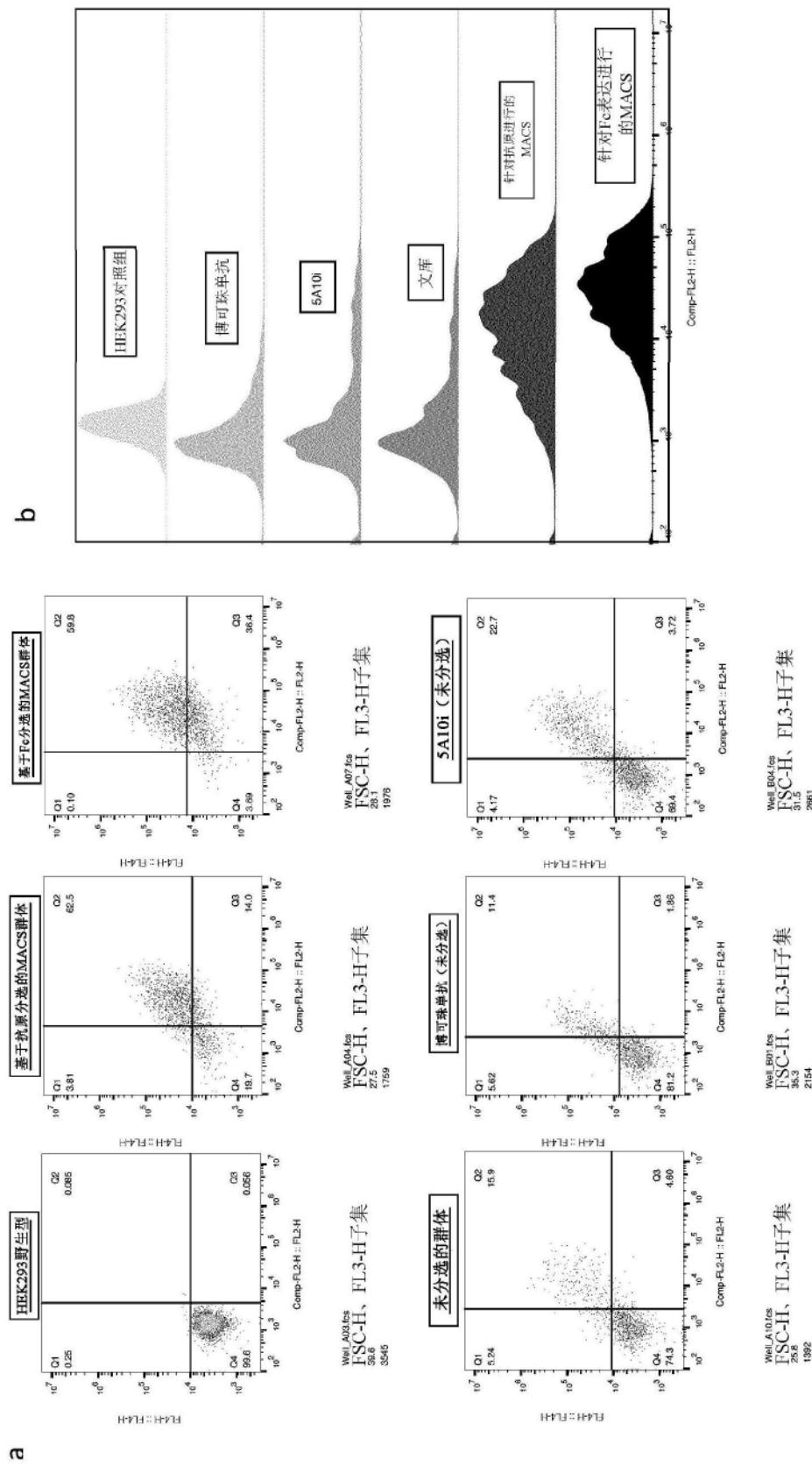


图13

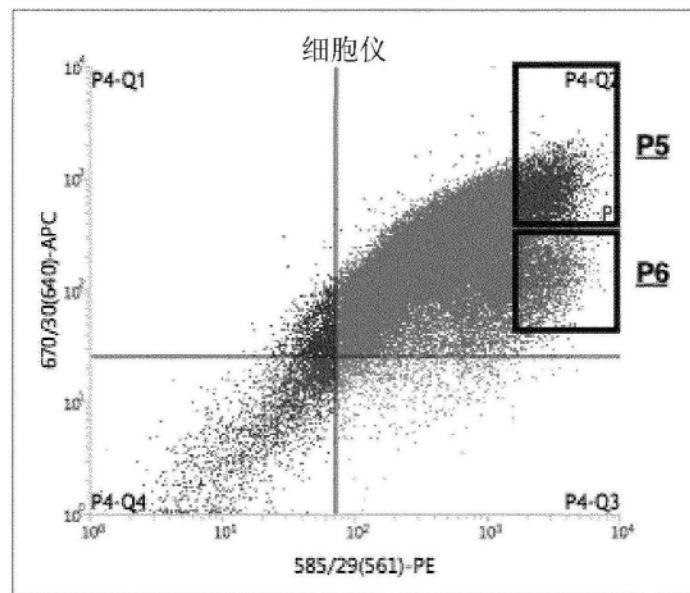
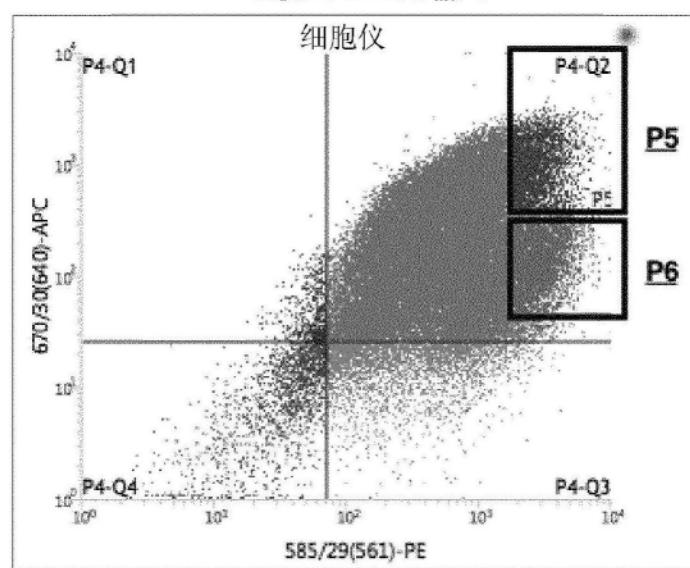
**a. 抗原MACS输入****b. 抗Fc MACS输入**

图14

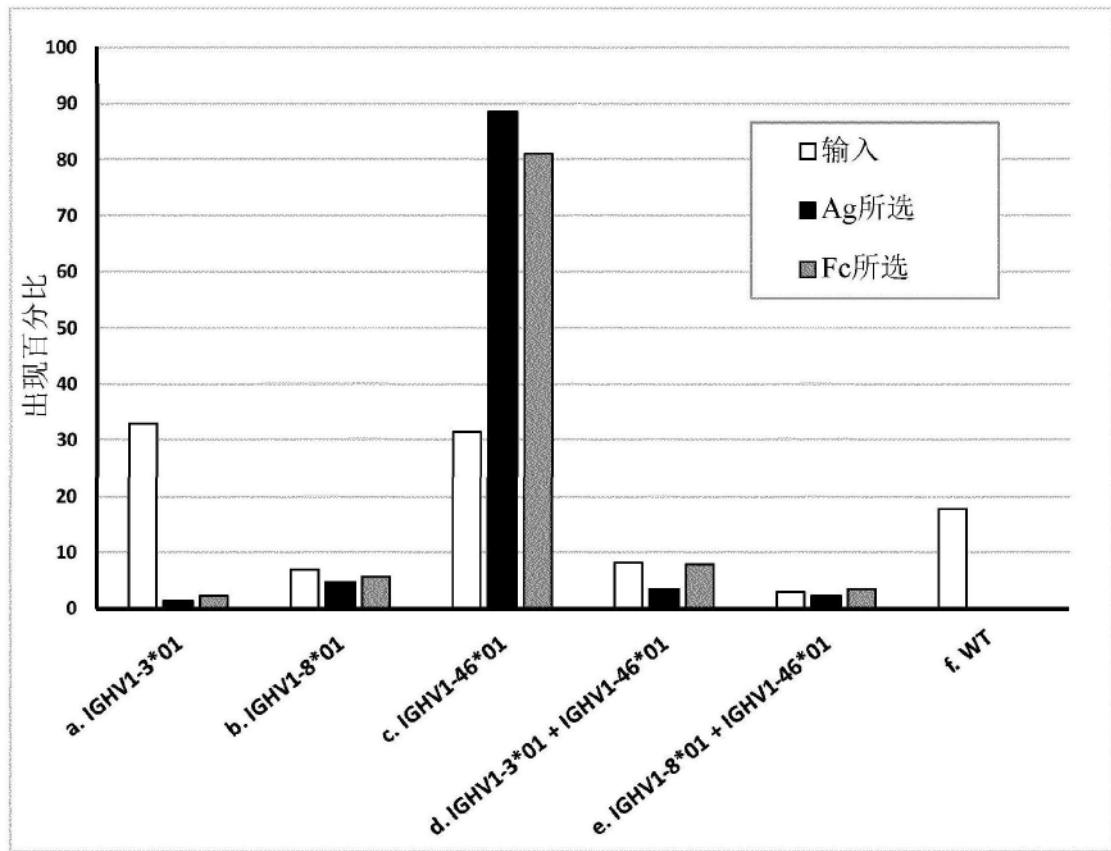


图15

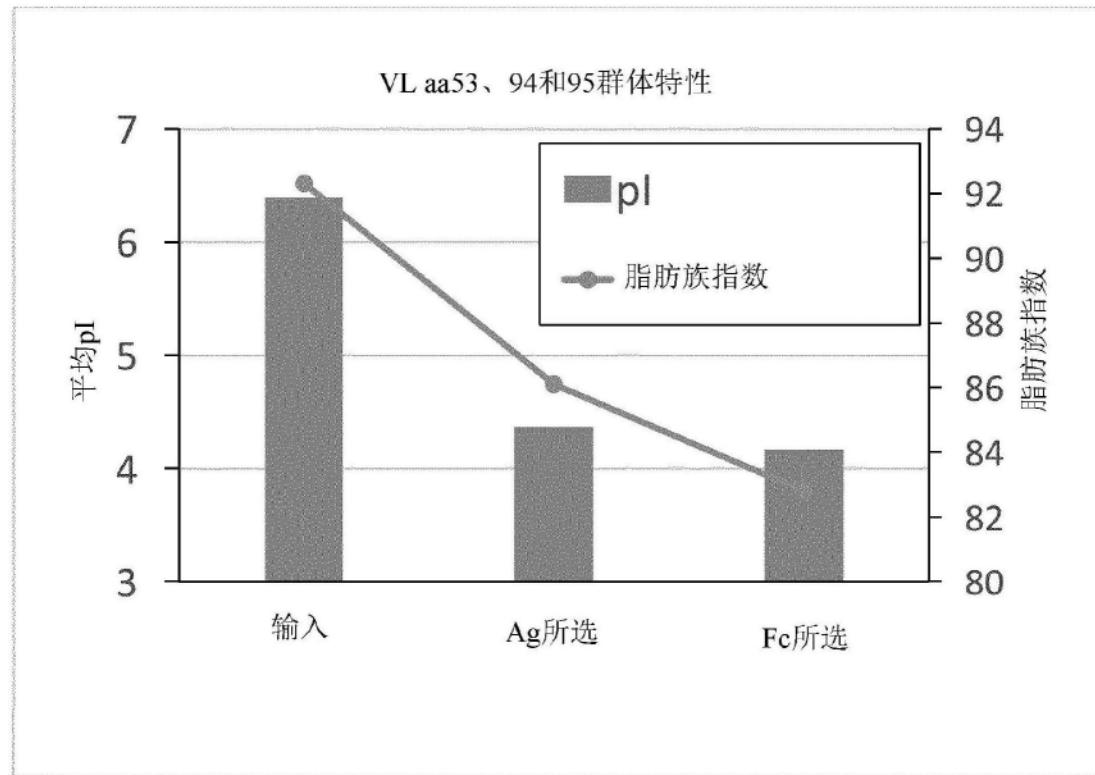


图16

克隆名称	抗原结合 (捕捉 ELISA)	AC- SINS $\Delta\lambda$ (nm)	VH CDR1	VH CDR2	VH	VL CDR2	VL CDR3
884_01_A01	6174	12	GYTFTSYY	INPSGGSTNYN	c	SRYT	QQRYSLOQRT
884_01_A02	3570	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSPWRT
884_01_A03	3478	12	GYTFTSYY	INPSGGSTNYN	c	SQRYT	QQRYSNSRT
884_01_A04	5005	16	GYTFTSYY	INPSGGSTNYN	c	SSRYT	QQRYSLKRT
884_01_A05	1639	12	GYTFTSYS	INPSGGRTNYN		SQRYT	QQRYSPPRT
884_01_A06	2912	10	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSNNRT
884_01_A07	189	12	GYTFTSYD	INPSGGSTNYN	e	SPRYT	QQRYSGRT
884_01_A08	4338	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSLART
884_01_A09	1852	12	GYTFTSYY	INPSGGSTNYN	c	SDRYT	QQRYSTLRT
884_01_A10	124	12	nd	nd	nd	nd	nd
884_01_A11	2477	10	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYSLART
884_01_A12	4703	10	GYTFTSYY	INPSGGSTNYN	c	SRYT	QQRYSLLRT
884_01_B01	323	14	GYTFTSYD	ISPFGGRTNYN	b	SQRYT	QQRYSHTRT
884_01_B02	4756	12	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSFART
884_01_B03	2180	12	GYTFTSYY	INPSGGSTNYN	c	SDRYT	QQRYSIRRT
884_01_B04	147	10	GYTFTSYA	INPSGGSTNYN	d	STRYT	QQRYSVCRT
884_01_B05	3038	10	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSNNRT
884_01_B06	449	24	GYTFTSYD	ISPFGGRTNYN	b	SLRYT	QQRYSWART
884_01_B07	2864	12	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSHTRT
884_01_B08	2576	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSQMRT
884_01_B09	926	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSTLRT
884_01_B10	4752	10	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSADRT
884_01_B11	2534	12	GYTFTSYY	INPSGGSTNYN	c	SSRYT	QQRYSFNRT
884_01_B12	3481	14	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSLKRT
884_01_C01	3002	12	GYTFTSYY	INPSGGSTNYN	c	SSRYT	QQRYSYDRT
884_01_C02	2645	10	GYTFTSYY	INPSGGSTNYN	c	SDRYT	QQRYSQRT
884_01_C03	3011	12	GYTFTSYY	INPSGGSTNYN	c	SSRYT	QQRYSYVRT
884_01_C04	343	12	GYTFTSYD	ISPFGGRTNYN	b	SDRYT	QQRYSLYRT
884_01_C05	4155	10	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSADRT
884_01_C06	1293	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSMRT
884_01_C07	4638	12	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSYYRT
884_01_C08	5271	10	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSLIRT
884_01_C09	3735	12	GYTFTSYY	INPSGGSTNYN	c	SPRYT	QQRYSERRT
884_01_C10	1413	10	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYS
884_01_C11	4435	10	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSQMRT
884_01_C12	2317	12	GYTFTSYY	INPSGGSTNYN	c	nd	nd
884_01_D01	2396	10	GYTFTSYY	INPSGGSTNYN	c	SSRYT	QQRYSVNRT
884_01_D02	928	12	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYSFFRT
884_01_D03	2707	12	GYTFTSYY	INPSGGSTNYN	c	SQRYT	QQRYSNVRT
884_01_D04	4086	10	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSFART
884_01_D05	2361	12	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSAVRT
884_01_D06	2560	10	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSDPRT
884_01_D07	1152	0	nd	nd	nd	SARYT	QQRYSMTRT
884_01_D08	1226	10	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYSISRT
884_01_D09	291	12	GYTFTSYD	ISPFGGRTNYN	b	SDRYT	QQRYSLYRT
884_01_D10	2331	10	GYTFTSYY	INPSGGSTNYN	c	SDRYT	QQRYSNMRT
884_01_D11	2363	10	GYTFTSYY	INPSGGSTNYN	c	SDRYT	QQRYSQMRT
884_01_D12	3715	10	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSADRT
884_01_E01	796	10	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYSTIRT
884_01_E02	2790	12	GYTFTSYY	INPSGGSTNYN	c	SDRYT	QQRYSCLART
884_01_E03	111	22	GYTFTSYA	ISPFGGRTNYN	a	SERYT	QQRYSNHRT
884_01_E04	2565	10	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYSAKRT
884_01_E05	1149	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSRRRT

图17

884_01_E06	1382	12	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSCTRT
884_01_E07	105	12	nd	nd	nd	nd	nd
884_01_E08	2009	12	GYTFTSYY	INPSGGSTNYN	c	SNRYT	QQRYSYMRT
884_01_E09	1012	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSHFRT
884_01_E10	2897	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSSART
884_01_E11	2294	12	GYTFTSYY	INPSGGSTNYN	c	nd	nd
884_01_E12	4168	24	GYTFTSYY	INPSGGSTNYN	c	SMRYT	QQRYSNRRT
884_01_F01	162	26	GYTFTSYY	INPSGGSTNYN	c	SYLES	QHSRDLPLT
884_01_F02	4741	22	GYTFTSYY	INPSGGSTNYN	c	SYRYT	QQRYSTGRT
884_01_F03	2655	10	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSDETRT
884_01_F04	3780	12	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSETRT
884_01_F05	494	12	GYTFTSYA	INPSGGSTNYN	d	SYRYT	QQRYSNSRT
884_01_F06	1241	8	GYTFTSYY	INPSGGSTNYN	c	SNRYT	QQRYSPPRT
884_01_F07	4532	12	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSSART
884_01_F08	3007	8	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSLERT
884_01_F09	1203	10	GYTFTSYY	INPSGGSTNYN	c	STRYT	QQRYSHSRT
884_01_F10	1339	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSVART
884_01_F11	160	10	GYTFTSYD	INPSGGSTNYN	e	STRYT	QQRYSLRT
884_01_F12	4294	10	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSHPRT
884_01_G01	5237	12	GYTFTSYY	INPSGGSTNYN	c	SHRYT	QQRYSDETRT
884_01_G02	2727	10	GYTFTSYY	INPSGGSTNYN	c	SQRYT	QQRYSNMRT
884_01_G03	2719	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSSKRT
884_01_G04	1228	12	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSHFRT
884_01_G05	1477	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSPPRT
884_01_G06	3557	12	GYTFTSYY	INPSGGSTNYN	c	SSRYT	QQRYSESRT
884_01_G07	939	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSVYRT
884_01_G08	233	0	GYTFTSYA	INPSGGSTNYN	d	SMRYT	QQRYSCLRT
884_01_G09	4788	12	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSATRT
884_01_G10	4511	10	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSATRT
884_01_G11	892	10	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSDCRT
884_01_G12	2133	12	GYTFTSYY	INPSGGSTNYN	c	SVRYT	QQRYSMART
884_01_H01	3075	10	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSGERT
884_01_H02	2688	10	GYTFTSYY	INPSGGSTNYN	c	SERYT	QQRYSIRRT
884_01_H03	5182	12	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSLMRT
884_01_H04	4873	12	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSLDRT
884_01_H05	1997	12	GYTFTSYY	INPSGGSTNYN	c	SARYT	QQRYSYVRT
884_01_H06	4607	10	GYTFTSYY	INPSGGSTNYN	c	SLRYT	QQRYSSART
884_01_H07	3066	10	GYTFTSYY	INPSGGSTNYN	c	SPRYT	QQRYSLNRT
博可珠单抗	4604	26	GYTFTSYY	ISPFGGRTN	f	SYRYT	QQRYSLWRT
5A10i	4748	10	GYTFTSYW	INPSNGRTN		SYRYT	QQRYSTPRT
未转染	200	0	-	-	-	-	-
未转染	141	0	-	-	-	-	-
未转染	nd	0	-	-	-	-	-

图17(续)

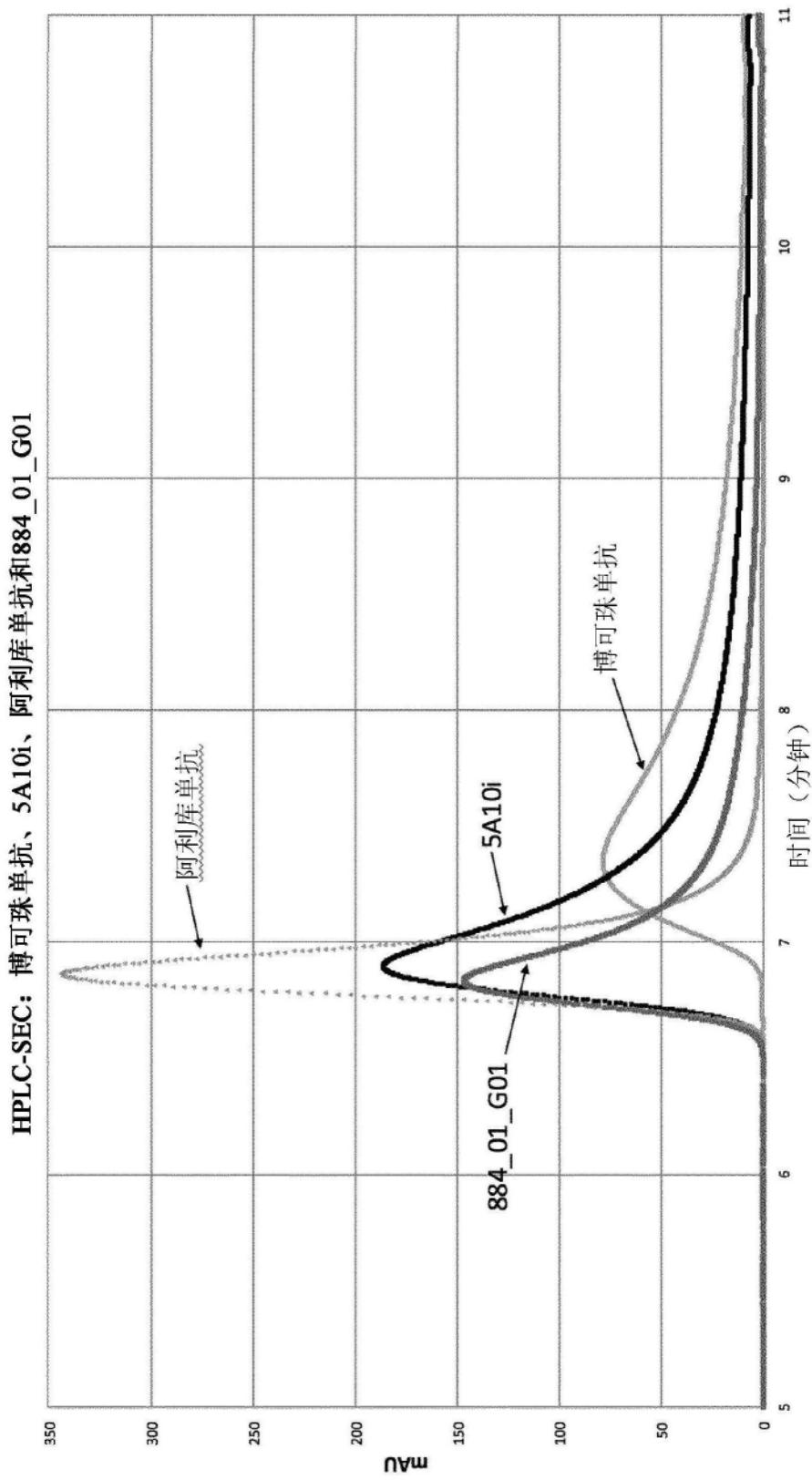


图18

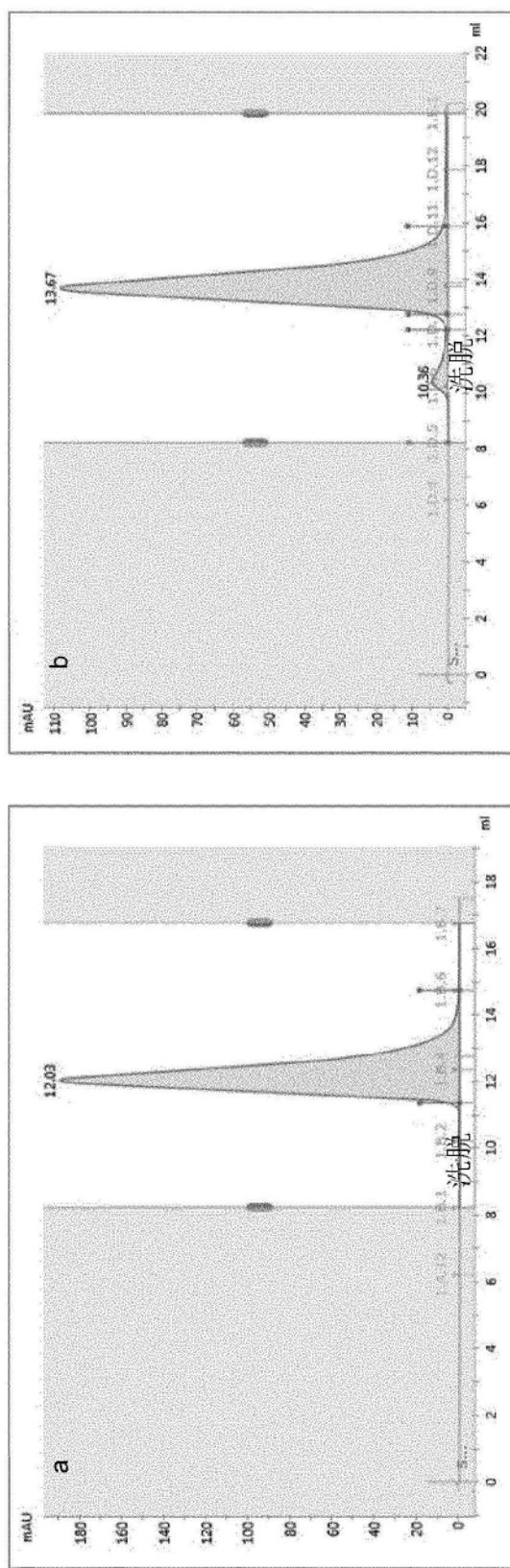


图19

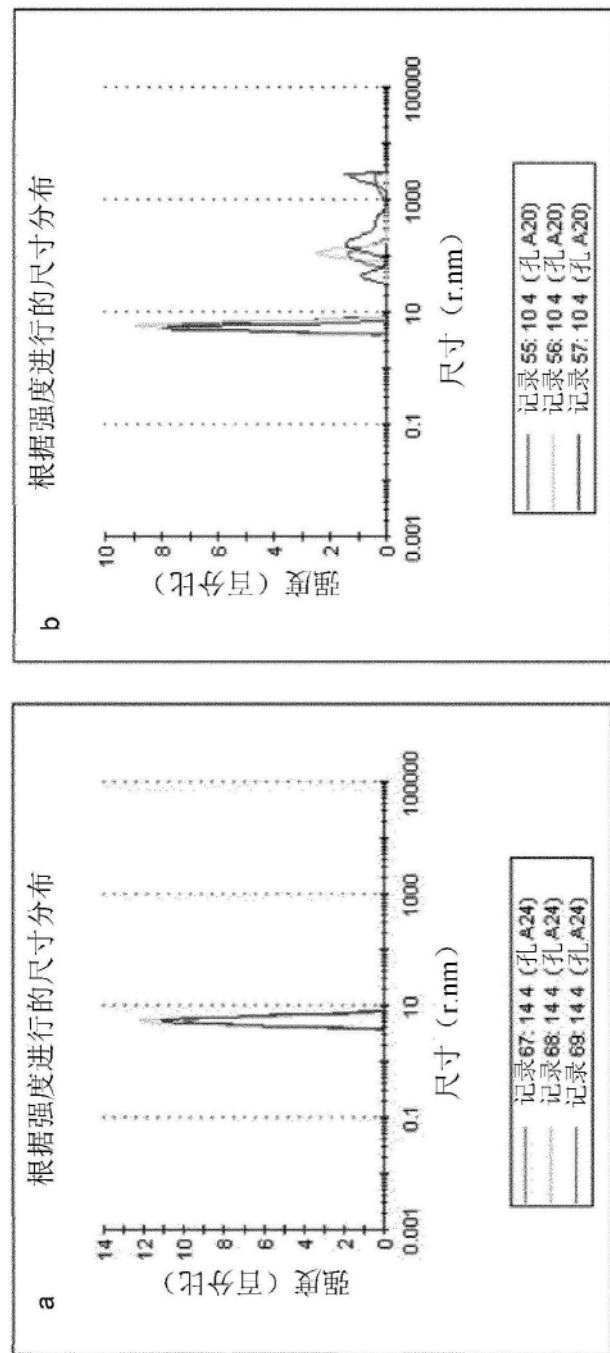


图20

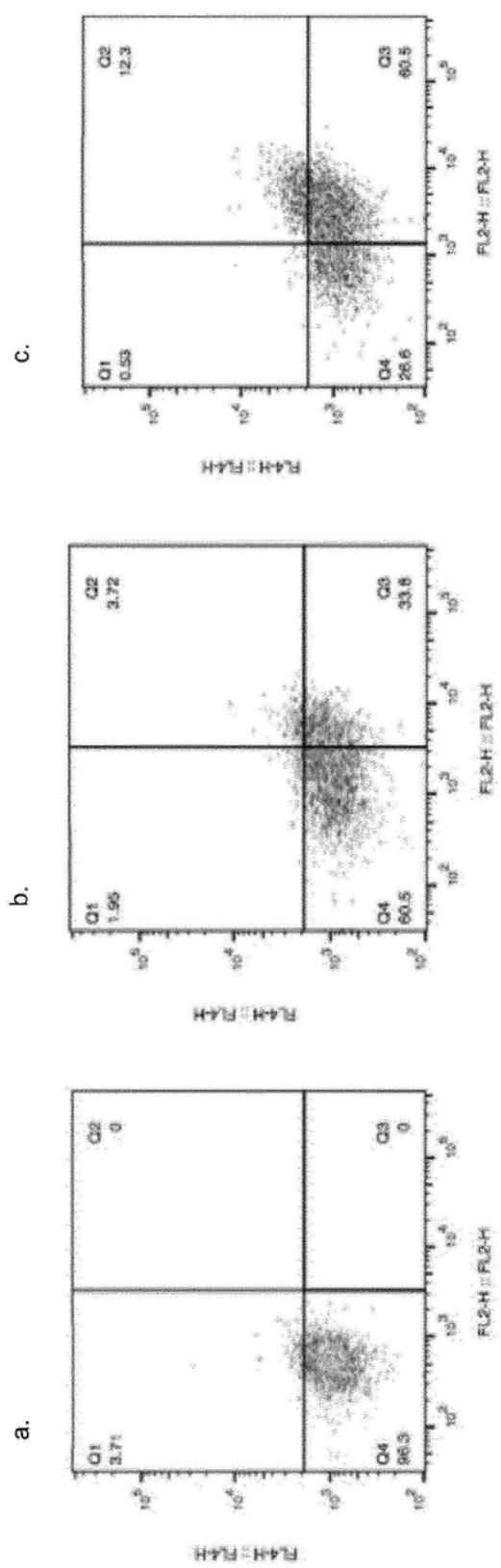


图21

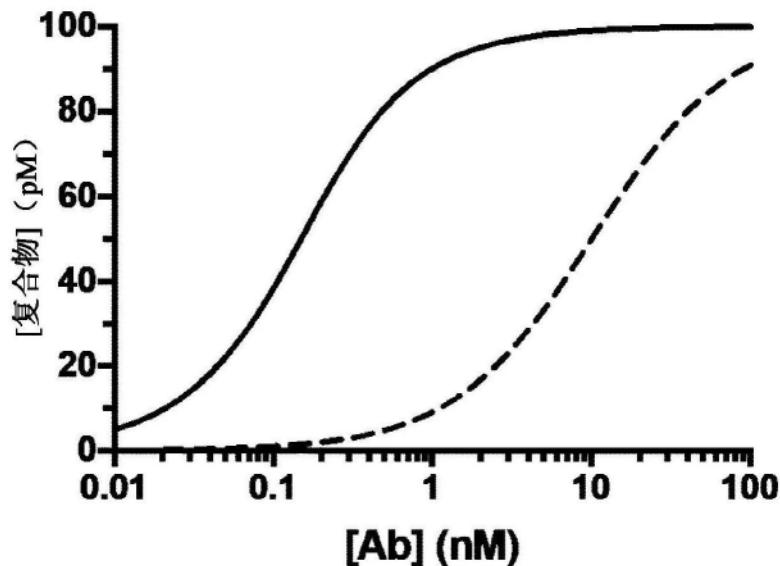


图22a

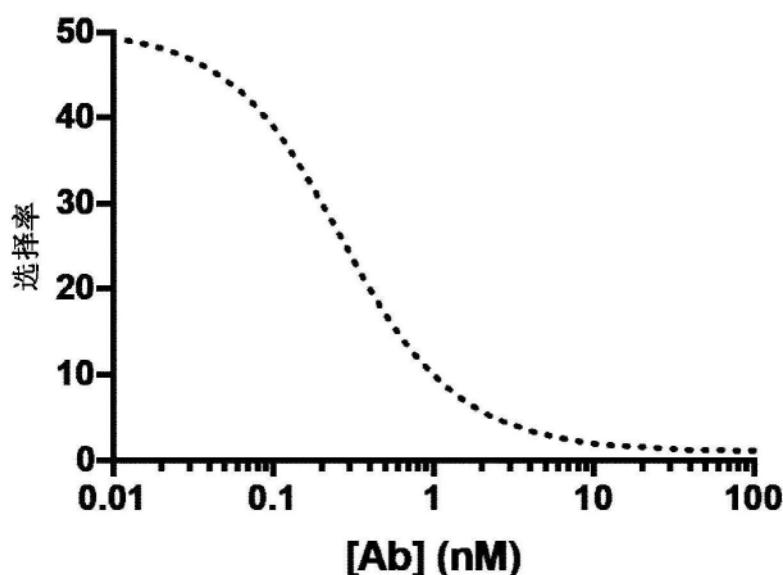


图22b

>剪接接合点  
>HindIII

J9	K	S>	L	S	L	S	P	G	K	*
J10	AAG	TCC	CTA	AGC	TTG	TCT	CC	GG/GTAAA	TGA	GT
J29	AAG	TCC	CTA	AGC	TTG	TCT	CC	AG/GTAAA	TGA	GT
J30	AAG	TCC	CTA	AGC	TTG	TCT	CC	AG/GTAAG	TGA	GT

GCC ACG GGC AAG CCC CCG CTC  
GCC ACG GGC AAG CCC CCG CTC  
GCC ACG GGC AAG CCC CCG CTC  
GCC ACG GGC AAG CCC CCG CTC

终止

图23

>序列: pInt17-J30 范围: 4801到8500

```

>XhoI
|
| 4810      4820      4830      4840      4850      4860
AGT CTC GAG TGC CTC CAC CAA GGG CCC TAG CGT CTT TCC TCT GGC CCC TTC CTC CAA GTC
S>
_a_>
| 4870      4880      4890      4900      4910      4920
TAC CTC TGG CGG CAC CGC TGC TCT GGG CTG CCT GGT GAA GGA CTA CTT CCC TGA GCC TGT
T S G G T A A L G C L V K D Y F P E P V>
| 4930      4940      4950      4960      4970      4980
GAC CGT GTC CTG GAA CTC TGG CGC CCT GAC CTC CGG CGT GCA TAC CTT CCC TGC CGT CCT
T V S W N S G A L T S G V H T F P A V L>
| 4990      5000      5010      5020      5030      5040
CCA GTC CTC CGG CCT GTA CTC CCT GTC CTC CGT GGT GAC CGT GCC TTC CTC CTC TCT GGG
Q S S G L Y S L S S V V T V P S S S L G>
| 5050      5060      5070      5080      5090      5100
CAC CCA GAC CTA CAT CTG CAA CGT GAA CCA CAA GCC TTC CAA CAC CAA GGT GGA CAA GAA
T Q T Y I C N V N H K P S N T K V D K K>
| 5110      5120      5130      5140      5150      5160
GGT GGA GCC TAA GTC CTG CGA CAA GAC CCA CAC CTG CCC TCC ATG TCC TGC CCC TGA GCT
V E P K S C D K T H T C P P C P A P E L>
| 5170      5180      5190      5200      5210      5220
GCT GGG CGG ACC CTC CGT GTT CCT GTT CCC TCC TAA GCC TAA GGA CAC CCT GAT GAT CTC
L G G P S V F L F P P K P K D T L M I S>
| 5230      5240      5250      5260      5270      5280
CCG GAC CCC TGA AGT GAC CTG CGT GGT GGA CGT GTC CCA CGA AGA TCC TGA AGT GAA

```

图24

图24(续)

内含子  
6130 6140 6150 6160 6170 6180  
GCC CTC CCT CCA GCA GCA GCT GCC CTG GGC TGG GCC ACG GGA AGC CCT AGG AGC CCC TGG  
内含子  
6190 6200 6210 6220 6230 6240  
GGA CAG ACA CAC AGC CCC TGC CTC TGT AGG AGA CTG TCC TGT CCT GTG AGC GCC CTG TCC  
内含子  
6250 6260 6270 6280 6290 6300  
TCC GAC CTC CAT GCC CAC TCG GGG GCA TGC CTA GTC CAT GTG CGT AGG GAC AGG CCC TCC  
内含子  
6310 6320 6330 6340 6350 6360  
CTC ACC CAT CTA CCC CCA CGG CAC TAA CCC CTG GCT GCC CTG CCC AGC CTC GCA CCC GCA  
内含子  
6370 6380 6390 6400 6410 6420  
TGG GGA AAC CAC CGA CTC CGG GGA CAT GCA CTC TCG GGC CCT GTG GAG GGA CTG GTC CAG  
内含子  
6430 6440 6450 6460 6470 6480  
ATG CCC ACA CAC ACA CTC AGC CCA GAC CCG TTC AAC AAA CCC CGC GCT GAG GTT GGC CGG  
内含子  
6490 6500 6510 6520 6530 6540  
CCA CAC GGC CAC CAC ACA CAC ACG TGC ACG CCT CAC ACA CGG AGC CTC ACC CGG GCG AAC  
内含子  
6550 6560 6570 6580 6590 6600  
CGC ACA GCA CCC AGA CCA GAG CAA GGT CCT CGC ACA CGT GAA CAC TCC TCA GAC ACA GGC  
内含子  
6610 6620 6630 6640 6650 6660  
CCC CAC GAG CCC CAC GCG GCA CCT CAA GGC CCA CGA GCC GCT CGG CAG CTT CTC CAC ATG  
内含子  
6670 6680 6690 6700 6710 6720  
CTG ACC TGC TCA GAC AAA CCC AGC CCT CCT CTC ACA AGG TGC CCC TGC AGC CGC CAC ACA  
内含子  
6730 6740 6750 6760 6770 6780  
CAC ACA GGC CCC CAC ACA CAG GGG AAC ACA CGC CAC GTC GCG TCC CTG GCA CTG GCC CAC  
内含子  
6790 6800 6810 6820 6830 6840  
TTC CCA ATG CCG CCC TTC CCT GCA GCT GAG GTC ACA TGA GGT GTG GGC TTC ACC ATC CTC  
内含子  
6850 6860 6870 6880 6890 6900  
CTG CCC TCT GGG CCT CAG GGA GGG ACA CAG GAG ATG GGG AGC GGG TCC TGC TGA GGG CCA  
内含子  
6910 6920 6930 6940 6950 6960  
GGT CGC TAT CTA GGG CTG GGT GTC TGG CTG AGT CCC GGG GCC AAA GCT GGT GCC CAG GGC  
内含子  
6970 6980 6990 7000 7010 7020  
AGG CAG CTG TGG GGA GCT GAC CTC AGG ACA CTG CTG GCC CAT CCC GGC CGG GCC CTA CAT  
内含子  
7030 7040 7050 7060 7070 7080  
CCT GGG TCC TGC CAC AGA GGG AAT CAC CCC CAG AGG CCC GAG CCC AGC AGG ACA CAG TAT  
内含子  
7090 7100 7110 7120 7130 7140  
TGA CCA CCC ACT TCC TGT CCA GAG CTG CAA CTG GAG GAG AGC TGT GCG GAG GCG CAG GAC  
L Q L E E S C A E A Q D>  
c c 内含子 c c c d d d M1外显子 d d d d d d >

图24(续)

7150            7160            7170            7180            7190            7200  
 GGG GAG CTG GAC GGG CTG TGG ACC ACC ATC ACC ATC TTC ATC ACA CTC TTC CTG CTA AGC  
 G E L D G L W T T I T F I T L F L L S>  
 \_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_MI外显子d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_>

7210            7220            7230            7240            7250            7260  
 GTG TGC TAC AGT GCC ACC ATC ACC TTC TTC AAG GTT AAA TAA TAG CTT ACG ACG TGA  
 V C Y S A T I T F F K V K>  
 \_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_M1外显子d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_d\_\_\_\_\_>

>BGH\_pA

7270            7280            7290            7300            7310            7320  
 TCA GCC TCG ACT GTG CCT TCT AGT TGC CAG CCA TCT GTT GTT TGC CCC TCC CCC GTG CCT  
 TCC TTG ACC CTG GAA GGT GCC ACT CCC ACT GTC CTT TCC TAA TAA AAT GAG GAA ATT GCA  
 TCG CAT TGT CTG AGT AGG TGT CAT TCT ATT CTG GGG GGT GGG GTG GGG CAG GAC AGC AAG  
 GGG GAG GAT TGG GAA GAC AAT AGC AGG CAT GCT GGG GAA TTC TGG CCC GGG CAT GAT AAC

>Asc1            >AAVS1\_右同源臂

7510            7520            7530            7540            7550            7560  
 TTC GTA TAA TGT ATG CTA TAC GAA GTT ATG TAT ACG GCG CGC CCA CTA GGG ACA GGA TTG  
 GTG ACA GAA AAG CCC CAT CCT TAG GCC TCC TCC CTA GTC TCC TGA TAT TGG GTC TAA  
 CCC CCA CCT CCT GTT AGG CAG ATT CCT TAT CTG GTG ACA CAC CCC CAT TTC CTG GAG CCA  
 TCT CTC TCC TTG CCA GAA CCT CTA AGG TTT GCT TAC GAT GGA GCC AGA GAG GAT CCT GGG  
 AGG GAG AGC TTG GCA GGG GGT GGG AGG GAA GGG GGT GCG TGA CCT GCC CGG TTC TCA  
 GTG GCC ACC CTG CGC TAC CCT CTC CCA GAA CCT GAG CTG CTC TGA CGC GGC TGT CTG GTG  
 CGT TTC ACT GAT CCT GGT GCT GCA GCT TCC TTA CAC TTC CCA AGA GGA GAA GCA GTT TGG  
 AAA AAC AAA ATC AGA ATA AGT TGG TCC TGA GTT CTA ACT TTG GCT CTT CAC CTT TCT AGT  
 CCC CAA TTT ATA TTG TTC CTC CGT GCG TCA GTT TTA CCT GTG AGA TAA GGC CAG TAG CCA  
 GCC CCG TCC TTG CAG GGC TGT GGT GAG GAG GGG GGT GTC CGT GTG GAA AAC TCC CTT TGT  
 GAG AAT GGT GCG TCC TAG GTG TTC ACC AGG TCG TGG CCG CCT CTA CTC CCT TTC TCT TTC  
 TCC ATC CTT CTT TCC TTA AAG AGT CCC CAG TGC TAT CTG GGA CAT ATT CCT CCG CCC AGA  
 GCA GGG TCC CGC TTC CCT AAG GCC CTG CTC TGG GCT TCT GGG TTT GAG TCC TTG GCA AGC  
 CCA GGA GAG GCG CTC AGG CTT CCC CCT TCC TCG TCC ACC ATC TCA TGC CCC TGG

>Sbf1

8350            8360            8370            8380            8390            8400  
 CTC TCC TGC CCC TTC CCT ACA GGG GTT CCT GGC TCT GCT CTC CTG CAG GCG ATC TCT CGA

>3' β 球蛋白隔离子

8410            8420            8430            8440            8450            8460  
 TCT CTC GAT TTC GAT CAA GAC ATT CCT TTA ATG GTC TTT TCT GGA CAC CAC TAG GGG TCA  
 GAA GTA GTT CAT CAA ACT TTC CCT CCC TAA TCT CAT T

图24(续)

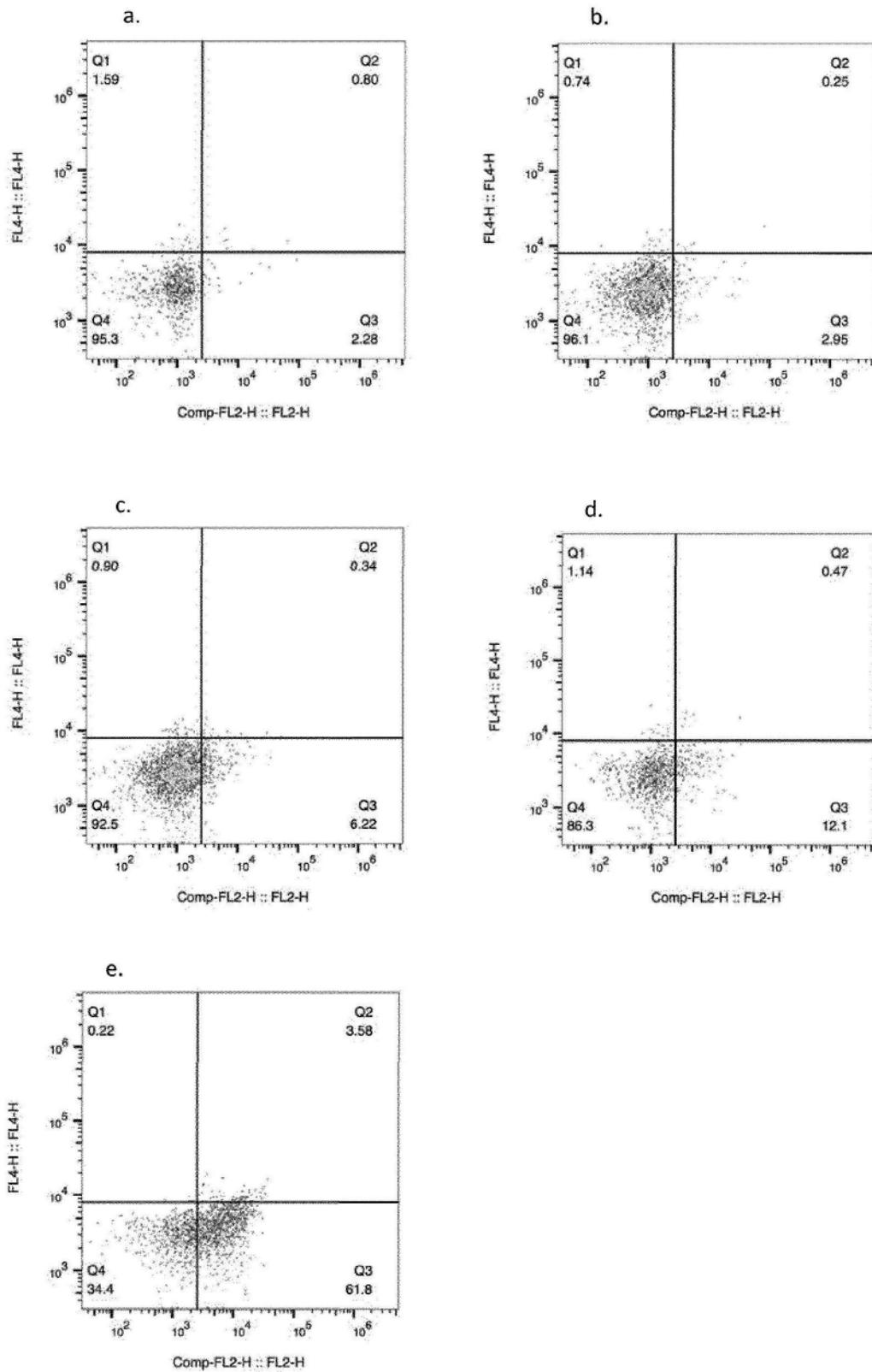


图25

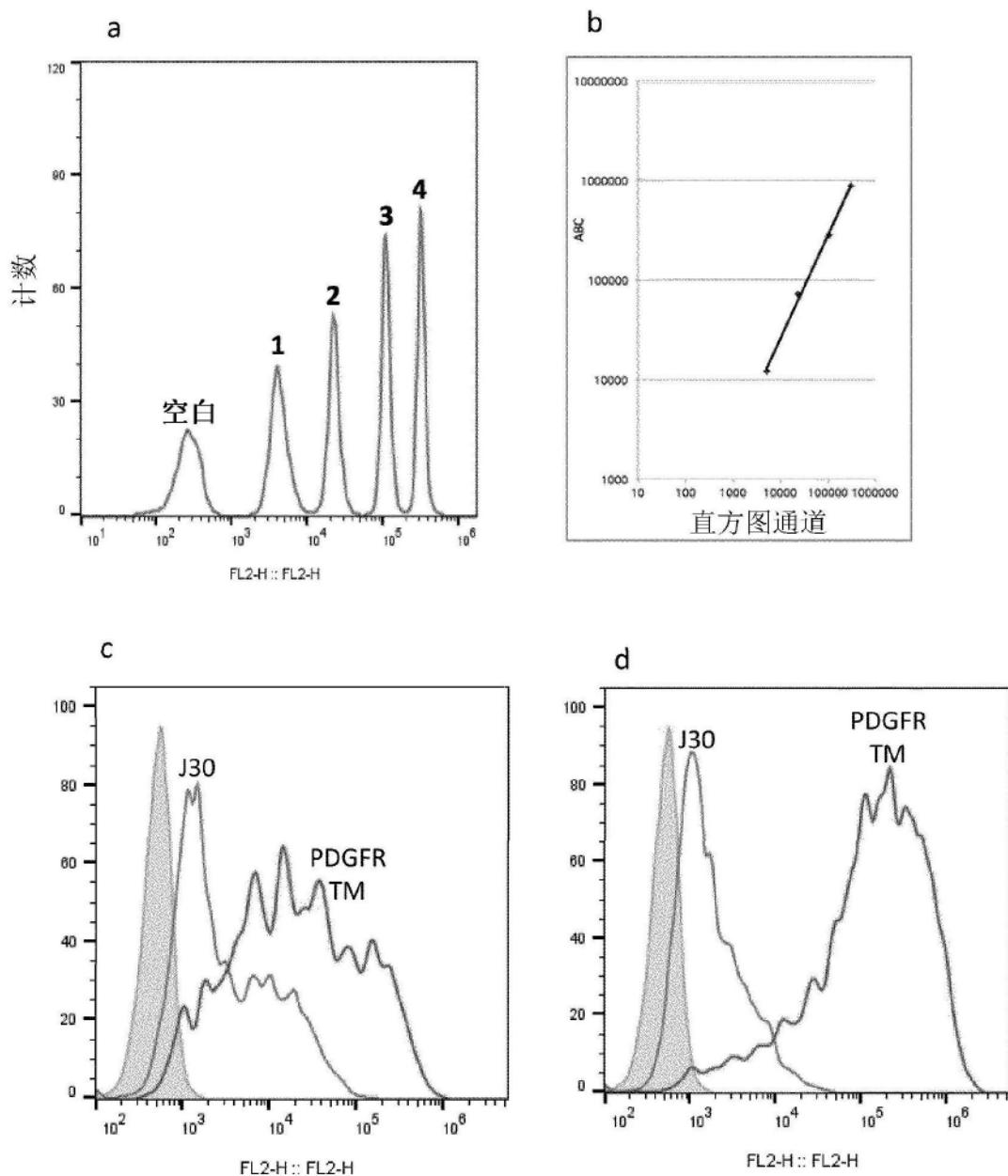


图26

## pINT17-J30

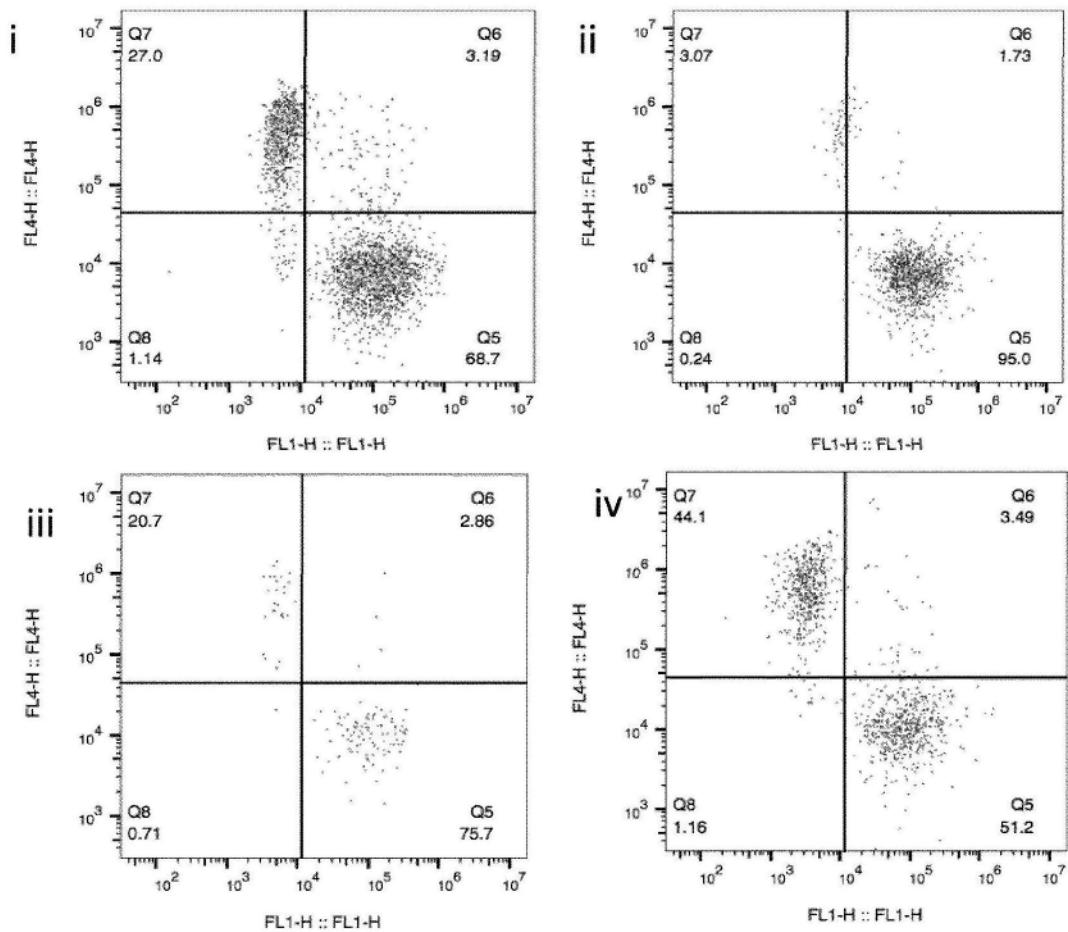


图27a

## pINT17-BSD

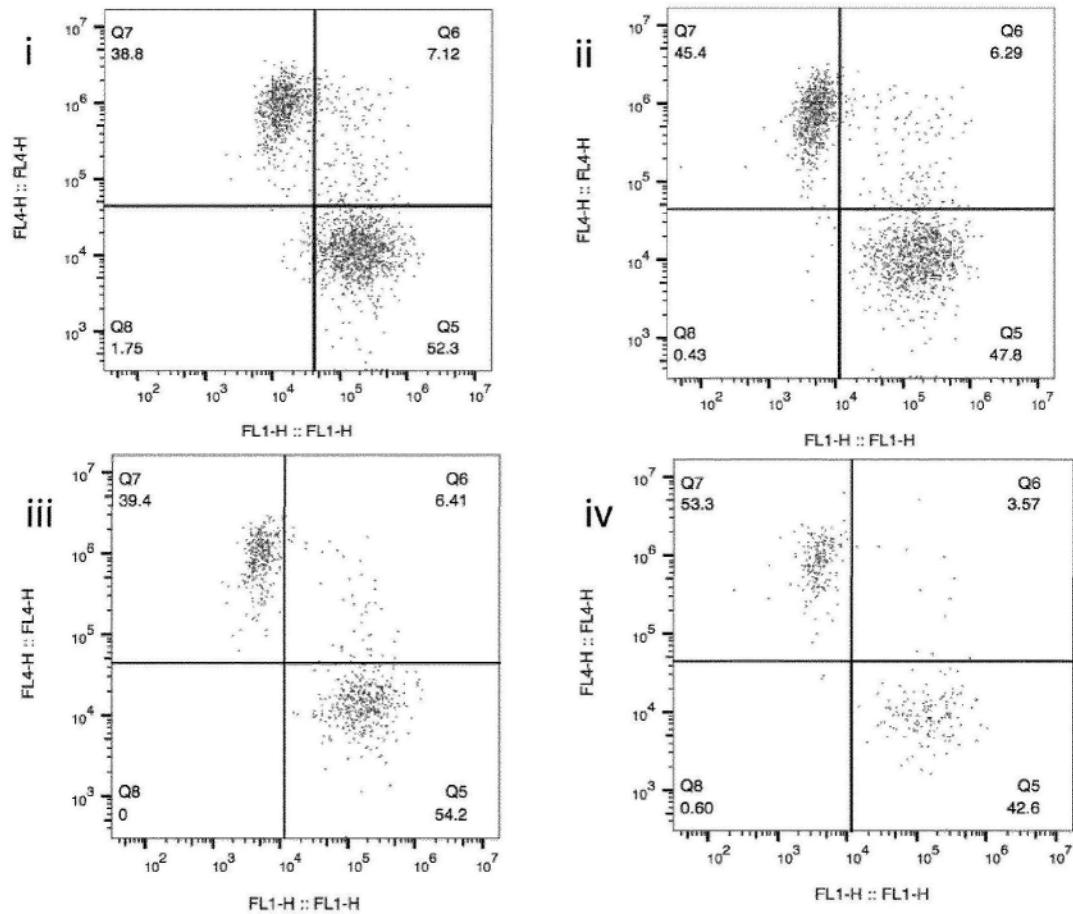


图27b

序列: pINT 18-Tet1 范围: 1至7680

>AsiSI >AAVS1\_左同源臂

```

|   |   |   |   |   |
|   10    20    30    40    50    60
GCG ATC GCT TTC TCT GAC CTG CAT TCT CTC CCC TGG GCC TGT GCC GCT TTC TGT CTG
CAG CTT GTG GCC TGG GTC ACC TCT ACG GCT GGC CCA GAT CCT TCC CTG CCG CCT CCT TCA
GGT TCC GTC TTC CTC CAC TCC CCC TTG CTC TCT GCT GTG TTG CTG CCC AAG GAT
GCT CTT TCC GGA GCA CTT CCT TCT CGG CGC TGC ACC ACG TGA TGT CCT CTG AGC GGA TCC
TCC CCG TGT CTG GGT CCT CTC CGG GCA TCT CTC CCT CAC CCA ACC CCA TGC CGT CTT
CAC TCG CTG GGT TCC CTT TTC CTT CTC CCT CGT GGG CCT GTG CCA TCT CTC GTT TCT TAG
GAT GGC CTT CTC CGA CGG ATG TCT CCC TTG CGT CCC GCC TCC CCT TCT TGT AGG CCT GCA
TCA TCA CGG TTT TTC TGG ACA ACC CCA AAG TAC CCC GTC TCC CTG GCT TTA GCC ACC TCT
CCA TCC TCT TGC TTT CCT TGC CTG GAC ACC CCG TTC TCC TGT GGA TTC GGG TCA CCT CTC
ACT CCT TTC ATT TGG GCA GCT CCC CTA CCC TTA CCT CTC TAG TCT GTG CAA GCT CTT
CCA GCC CCC TGT CAT GGC ATC TTC CAG GGG TCC GAG AGC TCA GCT AGT CTT CTT CCT CCA
ACC CGG GCC CCT ATG TCC ACT TCA GGA CAG CAT GTT TGC TGC CTC CAG GGA TCC TGT GTC
CCC GAG CTG GGA CCA CCT TAT ATT CCC AGG GCC GGT TAA TGT GGC TCT GGT TCT GGG TAC

```

790            800            810            820            830            840  
 TTT TAT CTG TCC CCT CCA CCC CAC AGT GGG GCA AGA TGC ATC TTC TGA CCT CTT CTC TTC  
a a a a a a SYN内含子 a a a a a a >

>剪接受体

```

|   |   |   |   |   |
|   850    860    870    880    890    900
CTC CCA CAG GGC ATG GCA AAA CCT CTG AGC CAG GAA GAC AGC ACA CTG ATT GAA AGA GCA
M A K P L S Q E E S T L I E R A>
a a a a > b b b b b b 杀稻瘟菌素 b b b b b b b b b b >

|   |   |   |   |   |
|   910    920    930    940    950    960
ACC GCT ACT ATC AAC AGC ATC CCC ATC TCC GAA GAC TAT TCT GTG GCT AGT GCC GCT CTG
T A T I N S I P I S E D Y S V A S A A L>
b b b b b b b b b b 杀稻瘟菌素 b b b b b b b b b b >

|   |   |   |   |   |
|   970    980    990    1000   1010   1020
TCC AGC GAC GGG AGA ATC TTC ACC GGT GTG AAC GTC TAC CAC TTT ACA GGC GGA CCA TGC
S S D G R I F T G V N V Y H F T G G P C>
b b b b b b b b b b 杀稻瘟菌素 b b b b b b b b b b >

|   |   |   |   |   |
|   1030   1040   1050   1060   1070   1080
GCA GAG CTG GTG GTC CTG GGG ACT GCA GCC GCT GCA GGC GCT GGT AAT CTG ACC TGT ATC
A E L V V L G T A A A A A G N L T C I>
b b b b b b b b b b 杀稻瘟菌素 b b b b b b b b b b >

|   |   |   |   |   |
|   1090   1100   1110   1120   1130   1140
GTG GCC ATT GGC AAC GAA AAT AGG GGC ATC CTG TCC CCA TGC GGC AGG TGT CGG CAG GTG
V A I G N E N R G I L S P C G R C R Q V>
b b b b b b b b b b 杀稻瘟菌素 b b b b b b b b b b >

|   |   |   |   |   |
|   1150   1160   1170   1180   1190   1200
CTG CTG GAT CTG CAT CCT GGC ATC AAG GCA ATT GTC AAA GAC TCT GAT GGA CAG CCT ACC
L L D L H P G I K A I V K D S D G Q P T>
b b b b b b b b b b 杀稻瘟菌素 b b b b b b b b b b >

|   |   |   |   |   |
|   1210   1220   1230   1240   1250   1260
GCC GTC GGT ATC CGT GAA CTG CTG CCT AGC GGC TAT GTC TGG GAG GGA TAA TGA GCT TGG
A V G I R E L L P S G Y V W E G * * >
b b b b b b b b b b 杀稻瘟菌素 b b b b b b b b b b >
```

>BGH\_pA

```

|   |   |   |   |   |
|   1270   1280   1290   1300   1310   1320
CTT CGA AAC ACA CAG ATC TAC ACG GTA CCA GCT TAC GAC GTG ATC AGC CTC GAC TGT GCC
TTC TAG TTG CCA GCC ATC TGT TGT TTG CCC CTC CCC CGT GCC TTC CTT GAC CCT GGA AGG
TGC CAC TCC CAC TGT CCT CTA ATA AAA TGA GGA AAT TGC ATC GCA TTG TCT GAG TAG

```

图28

GTG TCA TTC TAT TCT GGG GGG TGG GGT GGG GCA GGA CAG CAA GGG GGA GGA TTG GGA AGA

图28(续)

图28(续)

TGA TAG AGA ACG TAT GTC GAG TTT ACT CCC TAT CAG TGA TAG AGA ACG TAT GTC GAG TTT  
ACT CCC TAT CAG TGA TAG AGA ACG TAT GTC GAG TTT ATC CCT ATC AGT GAT AGA GAA CGT  
ATG TCG AGT TTA CTC CCT ATC AGT GAT AGA GAA CGT ATG TCG AGG TAG GCG TGT ACG GTG

3970	3980	3990	4000	4010	4020
GGA GGC CTA TAT AAG CAG AGC TCG TTT AGT GAA CCG TCA GAT CGC CTG GAT TCG AAT GAG					
				M R>	<u>d</u> >
4030	4040	4050	4060	4070	4080
GGC CTG GAT CTT CTT TCT CCT TTG CCT GGC CGG GAG GGC TCT GGC AGC TAG CGA GAT TGT					
A W I F F L L C L A G R A L A >	<u>d</u> <u>d</u> <u>d</u> <u>d</u> BM40前导序列 <u>d</u> <u>d</u> <u>d</u> <u>d</u> <u>d</u> >				
				L A R L >	<u>e</u> MK3475-VL <u>e</u> >
4090	4100	4110	4120	4130	4140
CCT GAC CCA GAG CCC CGC AAC ACT GTC CCT GTC CCC CGG AGA AAG AGC AAC CCT GTC CTG					
S * P R A P Q H C P C P E K E Q P C P >	<u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> MK3475-VL <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> >				
4150	4160	4170	4180	4190	4200
TAG AGC ATC AAA AGG TGT GTC TAC CAG TGG GTA CAG CTA TCT GCA CTG GTA CCA GCA GAA					
V E H Q K V C L P V G T A I C T G T S R >	<u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> MK3475-VL <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> >				
4210	4220	4230	4240	4250	4260
GCC CGG ACA GCA CCC TAG GCT GCT GAT CTA CCT GGC TTC TTA TCT GGA GAG TGG AGT GCC					
S P D R P L G C * S T W L L I W R V E C >	<u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> MK3475-VL <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> >				
4270	4280	4290	4300	4310	4320
AGC ACG GTT CTC AGG TTC CGG CAG CGG AAC AGA CTT TAC CCT GAC AAT TTC CAG CCT GGA					
Q H G S Q V P A A E Q T L P * Q F P A W >	<u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> MK3475-VL <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> >				
4330	4340	4350	4360	4370	4380
GCC AGA AGA CTT CGC CGT GTA CTA TTG CCA GCA TTC CAG AGA TCT GCC CCT GAC TTT TGG					
S Q K T S P C T I A S I P E I C P * L L >	<u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> MK3475-VL <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> <u>e</u> >				
4390	4400	4410	4420	4430	4440
CGG AGG GAC CAA GGT CGA AAT CAA AAG GAC TGC GGC CGC AAC CGT GGC TGC CCC TTC CGT					
A E G P R S K S K G L >	<u>e</u> <u>e</u> <u>e</u> MK3475-VL <u>e</u> <u>e</u> <u>e</u> <u>e</u> >				
			T V A A P S V >	<u>f</u> <u>f</u> <u>CL-K</u> <u>f</u> <u>f</u> >	
4450	4460	4470	4480	4490	4500
GTT CAT CTT CCC TCC CTC CGA CGA GCA GCT GAA GTC CGG CAC CGC CTC TGT GGT GTG CCT					
F I F P P S D E Q L K S G T A S V V C L >	<u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> CL-K <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> >				
4510	4520	4530	4540	4550	4560
GCT GAA CAA CTT CTA CCC TCG GGA GGC CAA GGT GCA GTG GAA GGT GGA CAA CGC CCT GCA					
L N N F Y P R E A K V Q W K V D N A L Q >	<u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> CL-K <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> >				
4570	4580	4590	4600	4610	4620
GTC CGG CAA CTC CCA GGA ATC CGT CAC CGA GCA GGA CTC CAA GGA CTC TAC CTA CTC CCT					
S G N S Q E S V T E Q D S K D S T Y S L >	<u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> CL-K <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> <u>f</u> >				

图28(续)

图28(续)

5290 5300 5310 5320 5330 5340  
 TGG GCT TTG ACT ATT GGG GGC AGG GAA CTA CCG TCA CAG TCT CGA GTG CCT CCA CCA AGG  
 M G F D Y W G Q G T T V T V>  
 j j j j j MK3475 VH j j j j j >  
 S S A S T K>  
 k IGG1 CH1-3 k >  
  
 5350 5360 5370 5380 5390 5400  
 GCC CTA GCG TCT TTC CTC TGG CCC CTT CCT CCA AGT CTA CCT CTG GCG GCA CCG CTG CTC  
 G P S V F P L A P S S K S T S G G T A A>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5410 5420 5430 5440 5450 5460  
 TGG GCT GCC TGG TGA AGG ACT ACT TCC CTG AGC CTG TGA CCG TGT CCT GGA ACT CTG GCG  
 L G C L V K D Y F P E P V T V S W N S G>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5470 5480 5490 5500 5510 5520  
 CCC TGA CCT CCG GCG TGC ATA CCT TCC CTG CCG TCC TCC AGT CCT CCG GCC TGT ACT CCC  
 A L T S G V H T F P A V L Q S S G L Y S>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5530 5540 5550 5560 5570 5580  
 TGT CCT CCG TGG TGA CCG TGC CTT CCT CTC TGG GCA CCC AGA CCT ACA TCT GCA ACG  
 L S S V V T V P S S S L G T Q T Y I C N>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5590 5600 5610 5620 5630 5640  
 TGA ACC ACA AGC CTT CCA ACA CCA AGG TGG ACA AGA AGG TGG AGC CTA AGT CCT GCG ACA  
 V N H K P S N T K V D K K V E P K S C D>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
 1 1 铰链 1 1 >  
  
 5650 | 5660 5670 5680 5690 5700  
 AGA CCC ACA CCT GCC CTC CAT GTC CTG CCC CTG AGC TGC TGG GCG GAC CCT CCG TGT TCC  
 K T H T C P P C P A P E L L G G P S V F>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
 K T H T C P P C P A P E L L G G P >  
 1 1 1 1 1 1 铰链 1 1 1 1 1 1 >  
  
 5710 5720 5730 5740 5750 5760  
 TGT TCC CTC CTA AGC CTA AGG ACA CCC TGA TGA TCT CCC GGA CCC CTG AAG TGA CCT GCG  
 L F P P K P K D T L M I S R T P E V T C>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5770 5780 5790 5800 5810 5820  
 TGG TGG TGG ACG TGT CCC ACG AAG ATC CTG AAG TGA AGT TCA ATT GGT ACG TGG ACG GCG  
 V V V D V S H E D P E V K F N W Y V D G>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5830 5840 5850 5860 5870 5880  
 TGG AGG TGC ACA ACG CCA AGA CCA AGC CTC GGG AGG AAC AGT ACA ACT CCA CCT ACC GGG  
 V E V H N A K T K P R E E Q Y N S T Y R>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5890 5900 5910 5920 5930 5940  
 TGG TGT CTG TGC TGA CCG TGC ACC AGG ACT GGC TGA AGC GCA AAG AAT ACA AGT GCA  
 V V S V L T V L H Q D W L N G K E Y K C>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >  
  
 5950 5960 5970 5980 5990 6000  
 AGG TGT CCA ACA AGG CCC TGC CTG CCC CTA TCG AAA AGA CCA TCT CCA AGG CTA AGG GCC  
 K V S N K A L P A P I E K T I S K A K G>  
 k k k k k k k IGG1 CH1-3 k k k k k k k >

图28(续)

图28(续)

AGC CAT CTC TCT CCT TGC CAG AAC CTC TAA GGT TTG CTT ACG ATG GAG CCA GAG AGG ATC  
 CTG GGA GGG AGA GCT TGG CAG GGG GTG GGA GGG AAG GGG GGG ATG CGT GAC CTG CCC GGT  
 TCT CAG TGG CCA CCC TGC GCT ACC CTC TCC CAG AAC CTG AGC TGC TCT GAC GCG GCT GTC  
 TGG TGC GTT TCA CTG ATC CTG GTG CTG CAG CTT CCT TAC ACT TCC CAA GAG GAG AAG CAG  
 TTT GGA AAA ACA AAA TCA GAA TAA GTT GGT CCT GAG TTC TAA CTT TGG CTC TTC ACC TTT  
 CTA GTC CCC AAT TTA TAT TGT TCC TCC GTG CGT CAG TTT TAC CTG TGA GAT AAG GCC AGT  
 AGC CAG CCC CGT CCT GGC AGG GCT GTG GTG AGG AGG GGG GTG TCC GTG TGG AAA ACT CCC  
 TTT GTG AGA ATG GTG CGT CCT AGG TGT TCA CCA GGT CGT GGC CGC CTC TAC TCC CTT TCT  
 CTT TCT CCA TCC TTC TTT CCT TAA AGA GTC CCC AGT GCT ATC TGG GAC ATA TTC CTC CGC  
 CCA GAG CAG CGT CCC CCT TCC CTA AGG CCC TGC TCT GGG CTT CTG CGT TTG AGT CCT TGG  
 CAA GCC CAG GAG CGC TCA GGC TTC CCT GTC CCC CTT CGT CCA CCA TCT CAT GCC

>Sbf1

7630            7640            7650            7660            7670            7680  
 CCT GGC TCT CCT GCC CCT TCC CTA CAG GGG TTC CTG GCT CTG CTC TCC TGC AGG CGA TCT

图28(续)

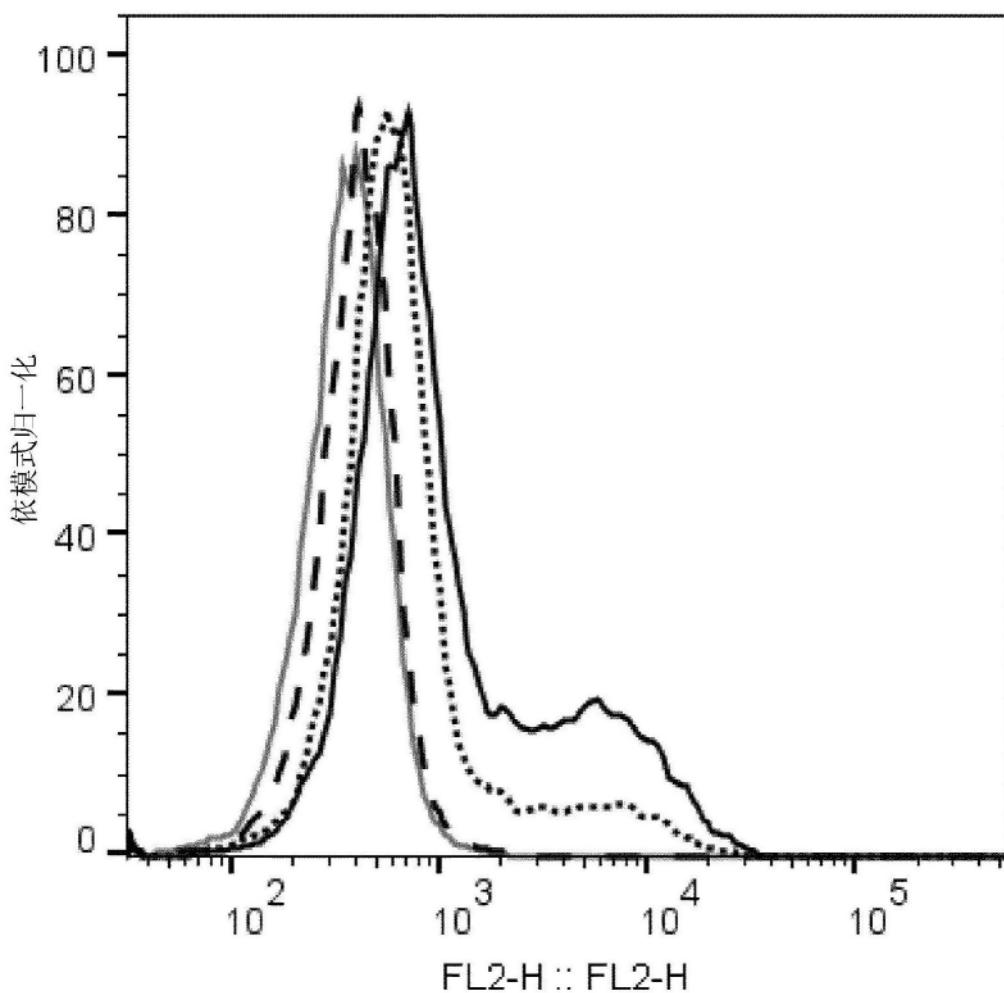


图29

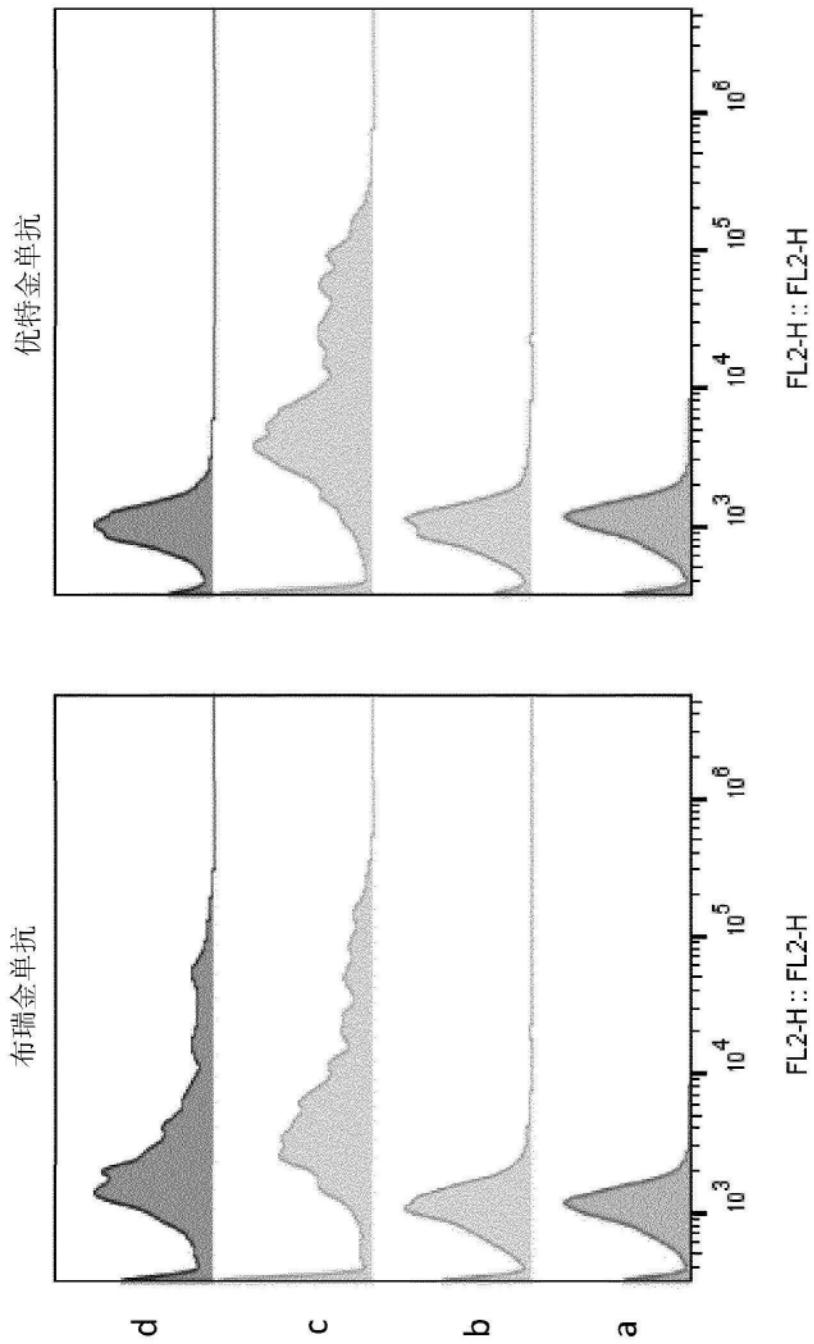


图30

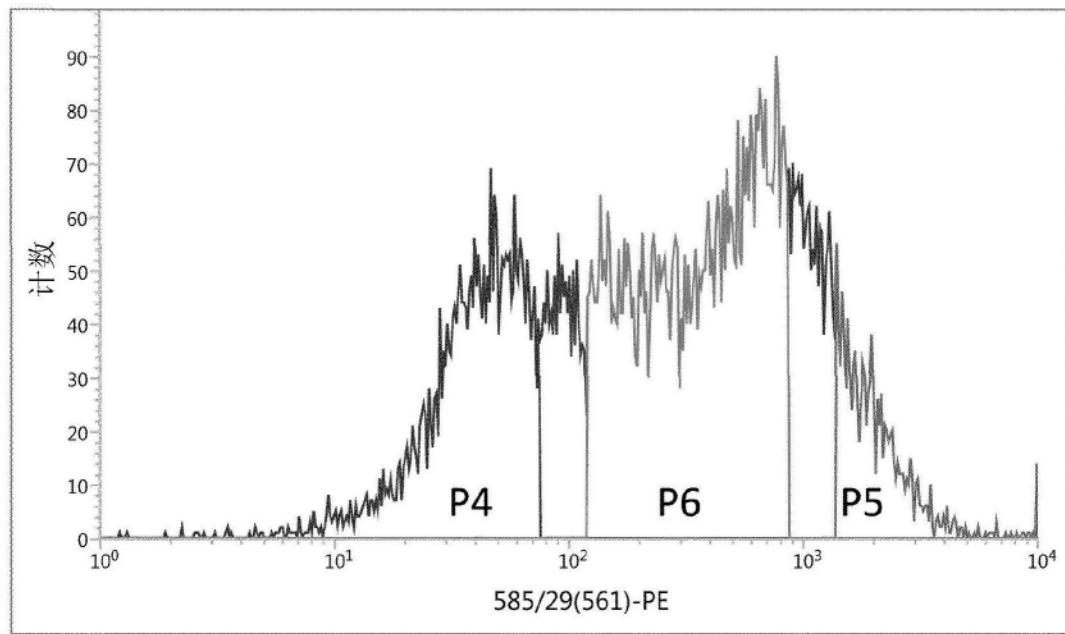


图31

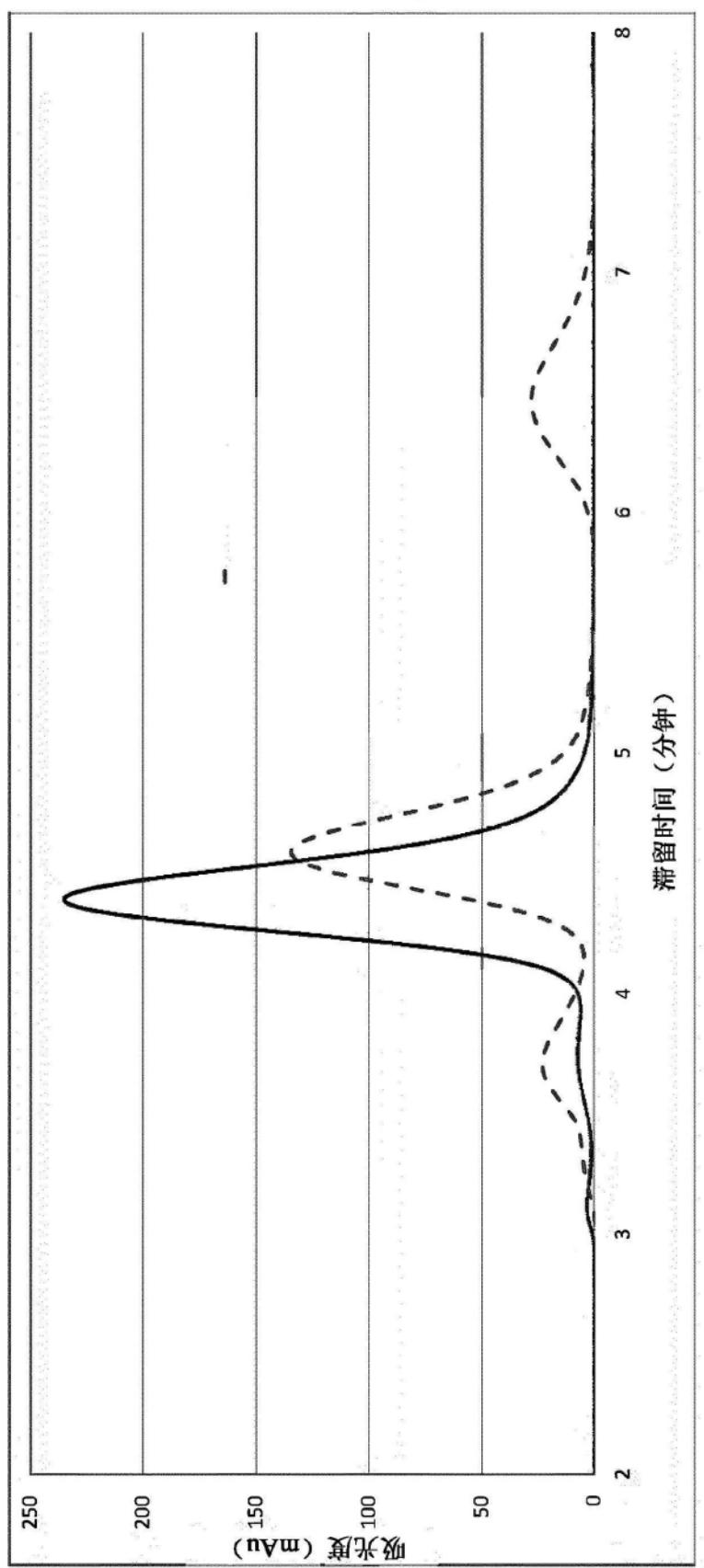


图32

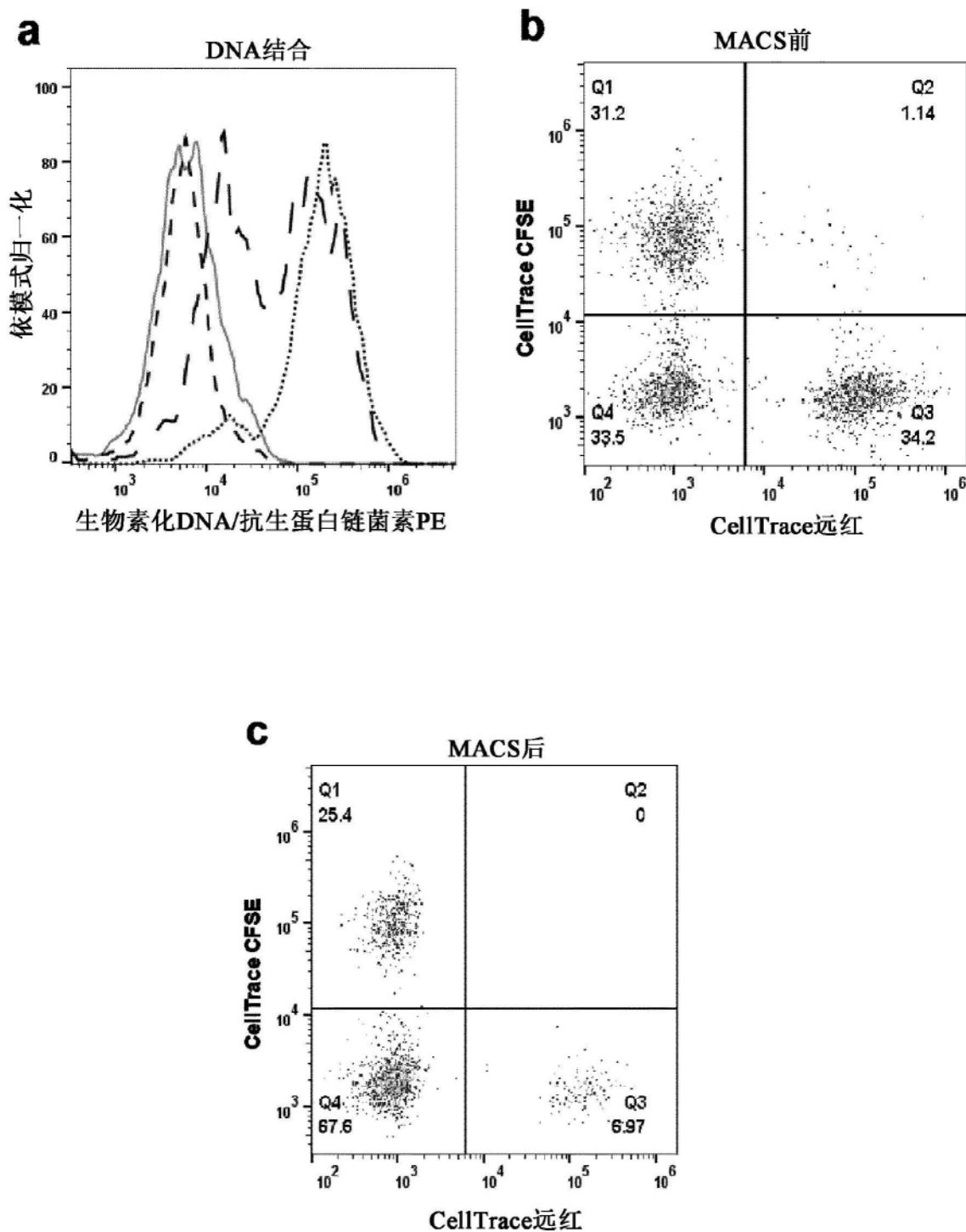


图33

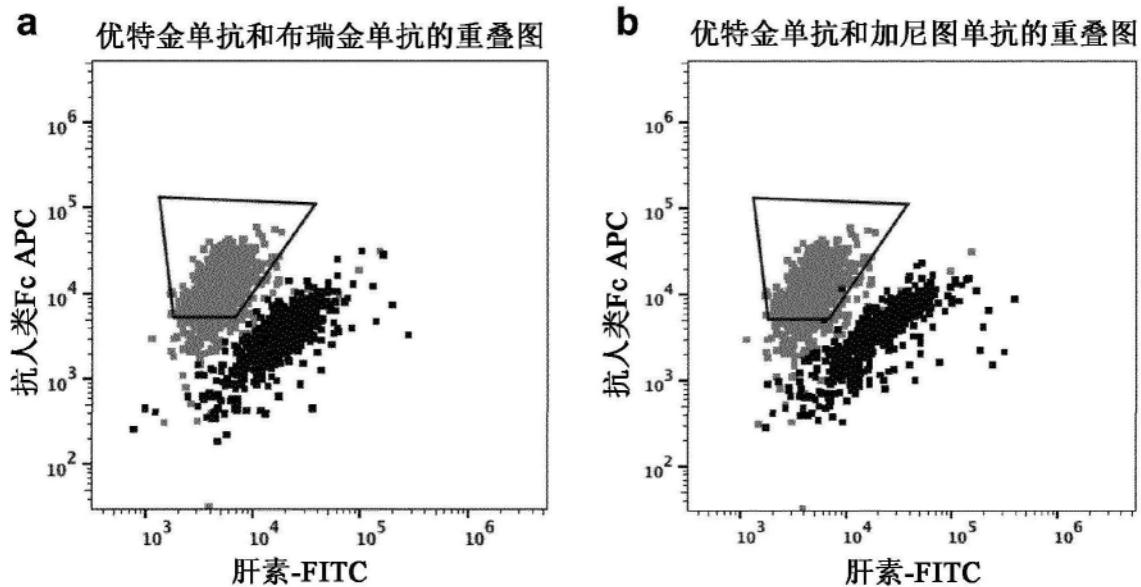


图34

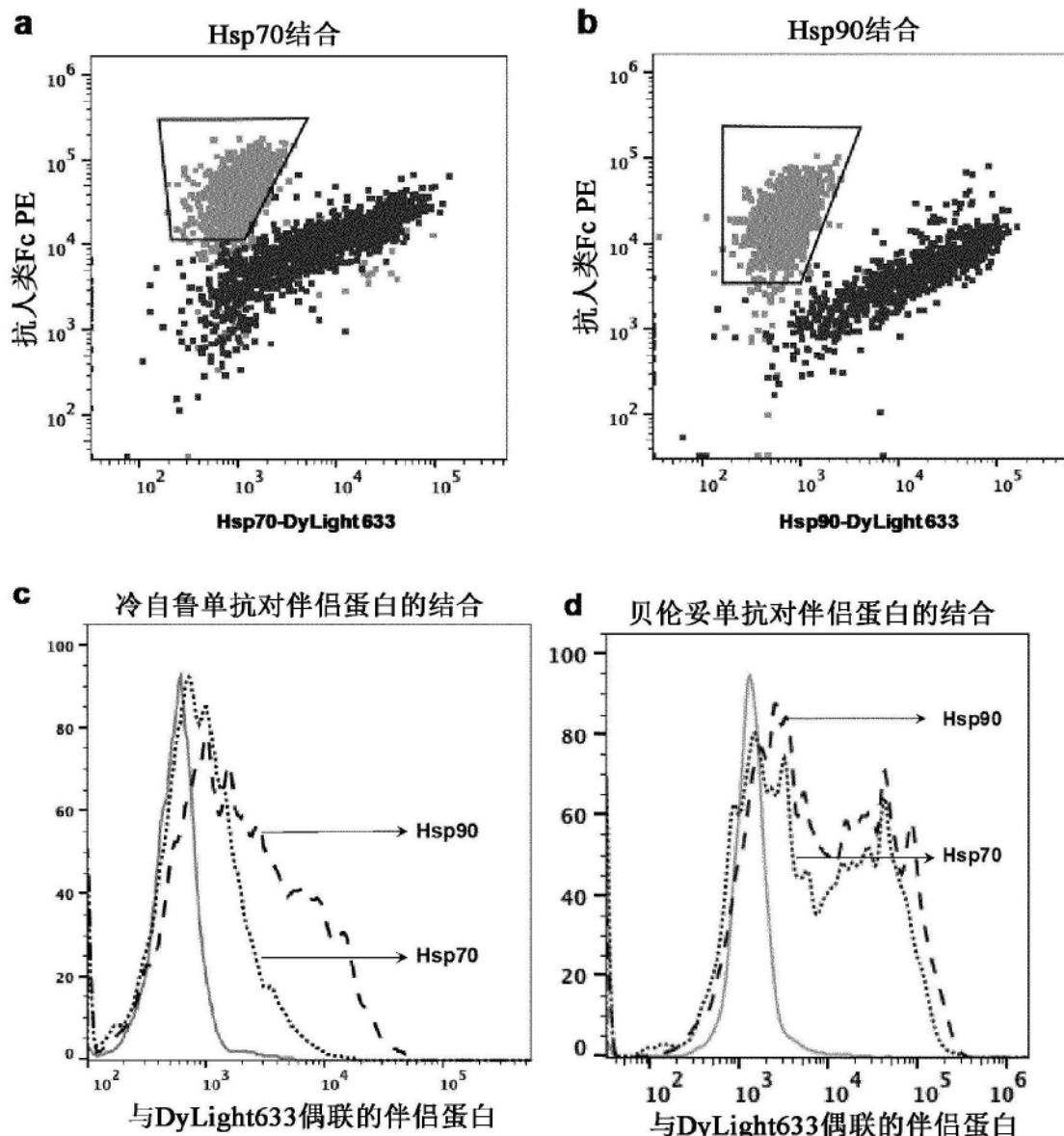


图35

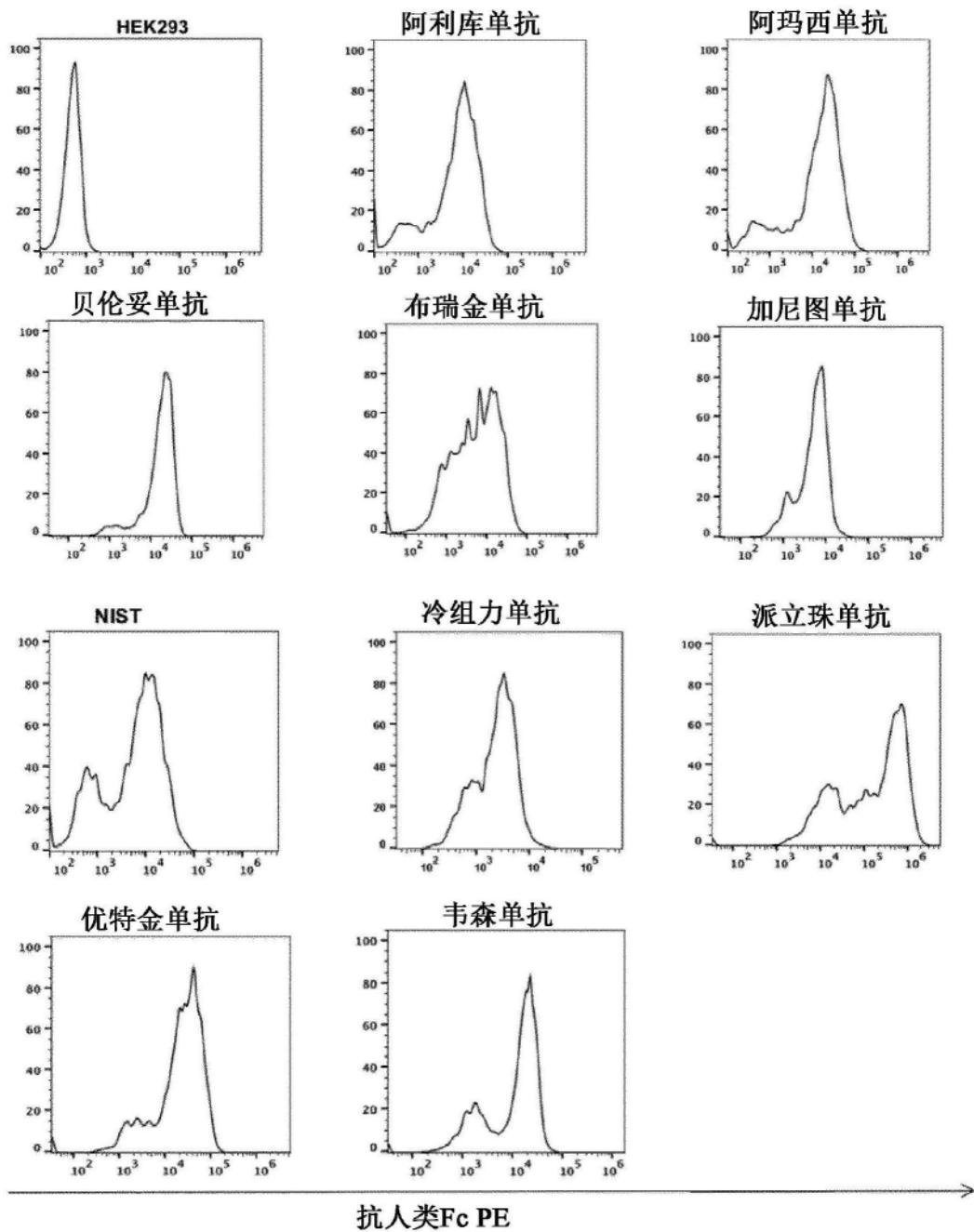


图36a

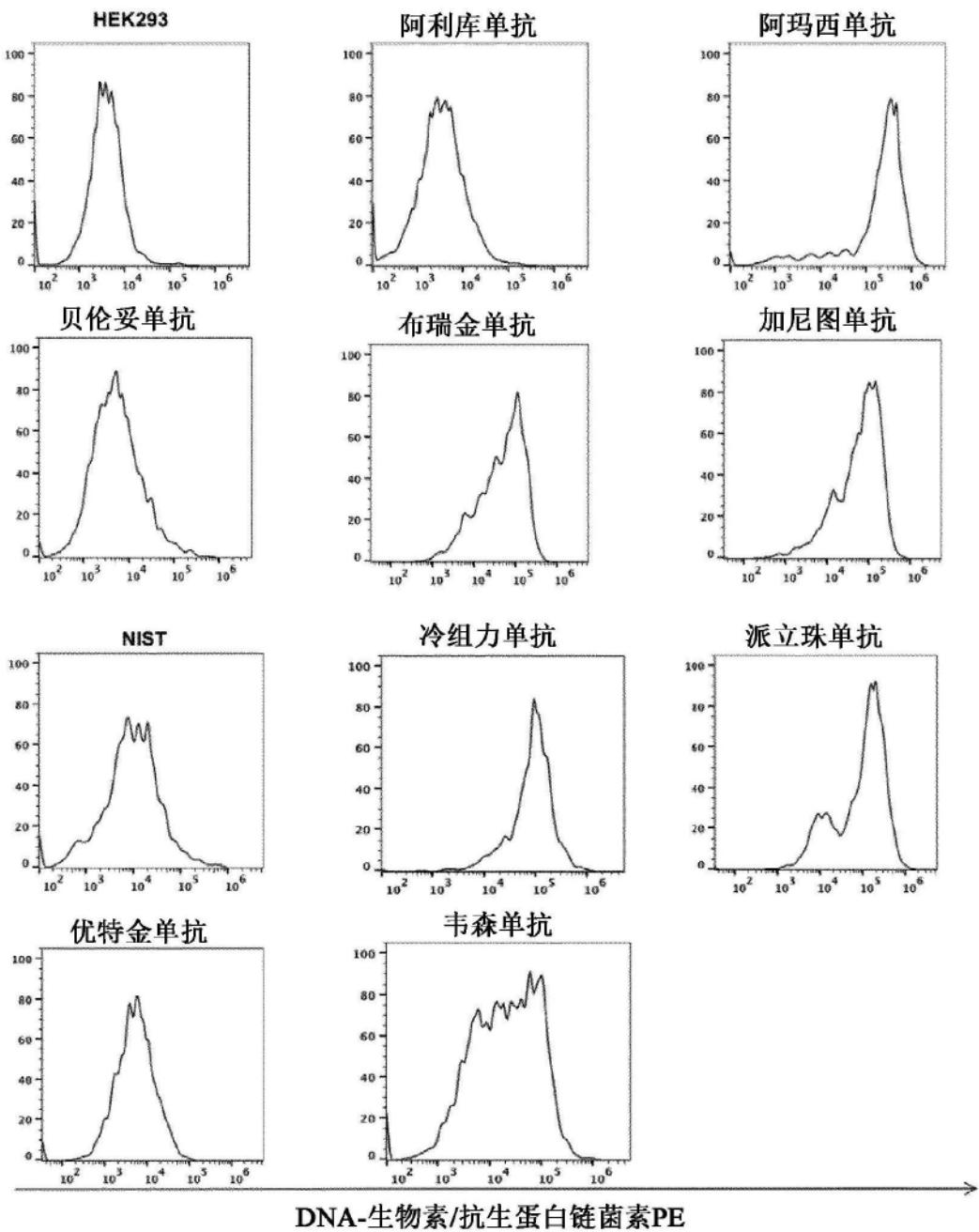


图36b

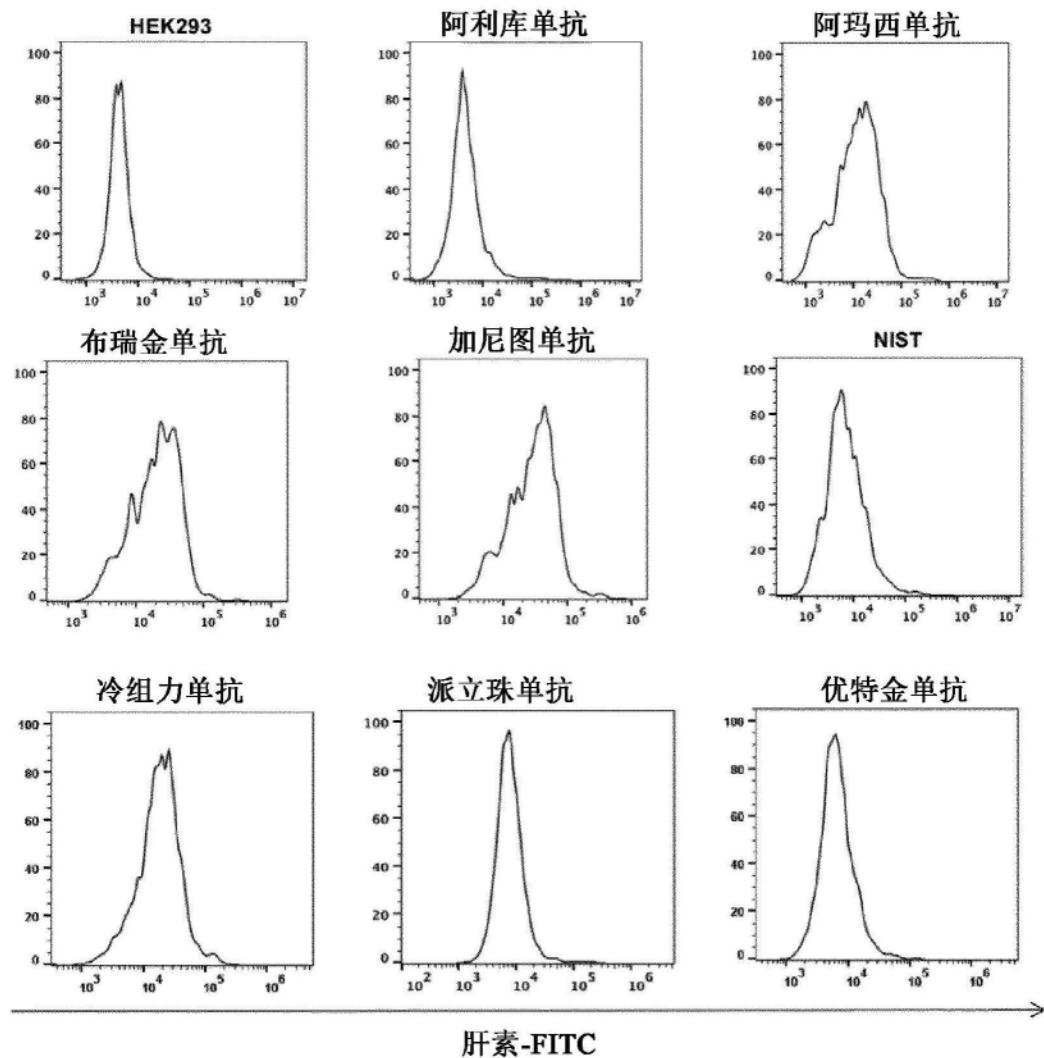


图36c

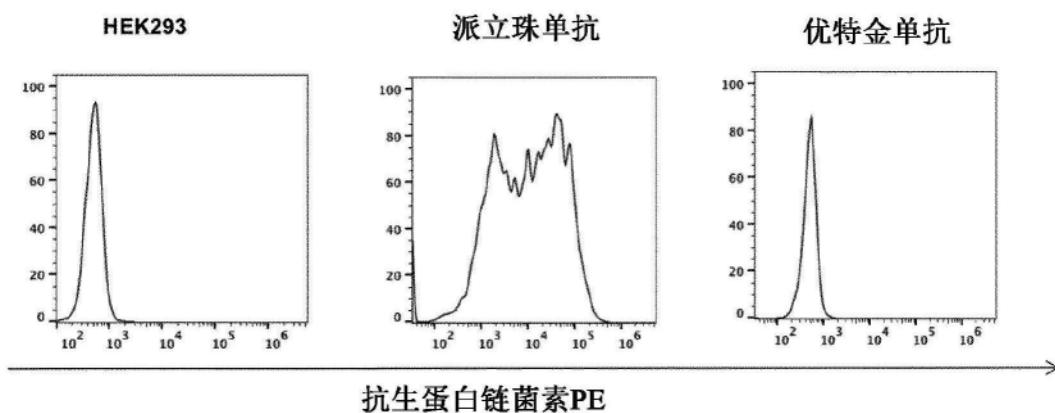


图36d

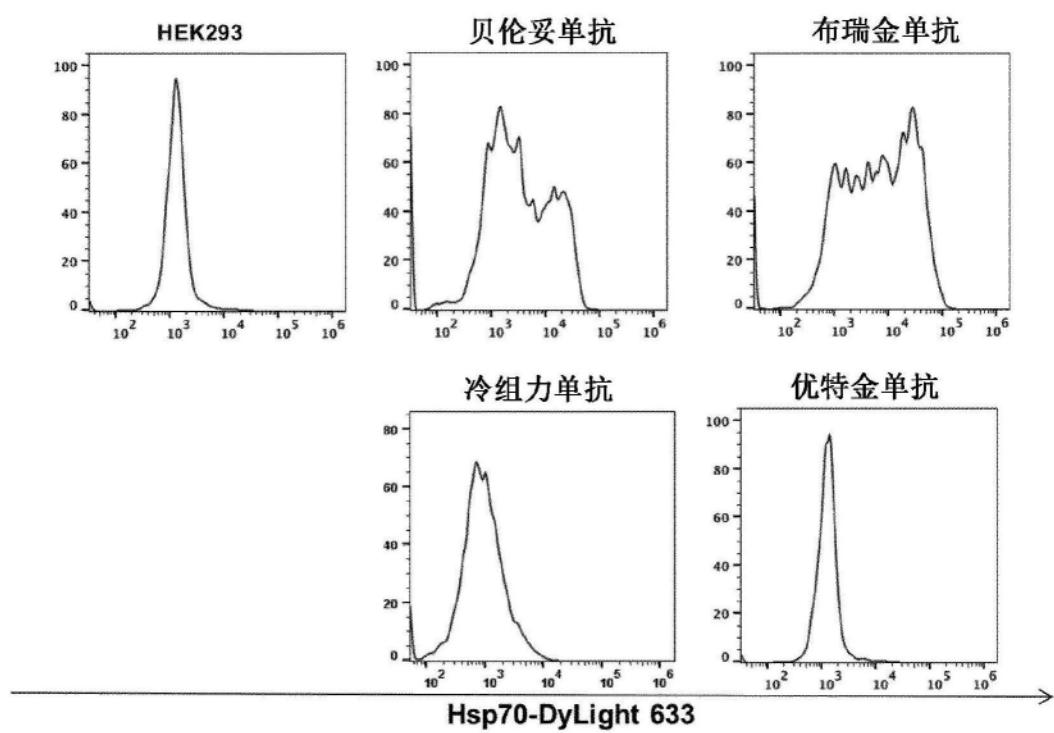
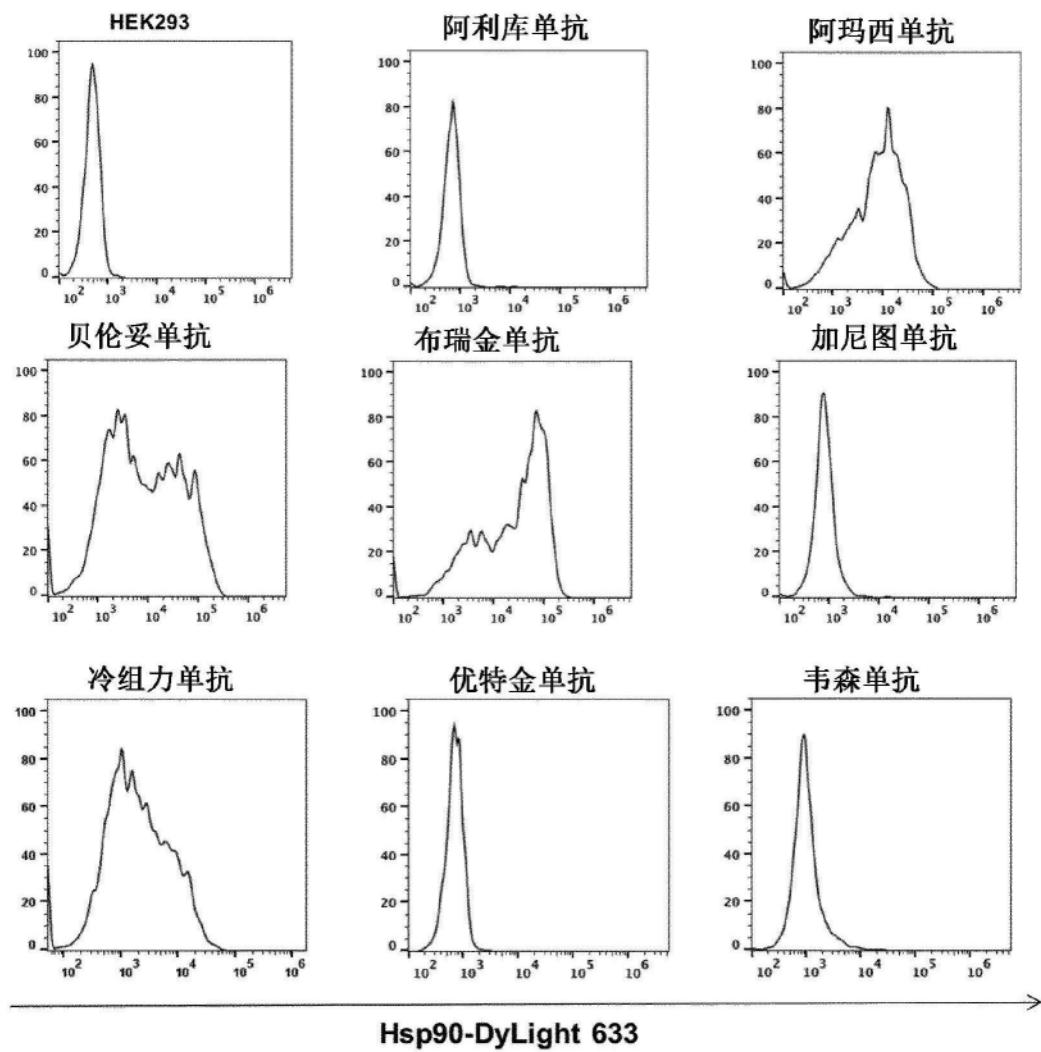


图36e



图|36f

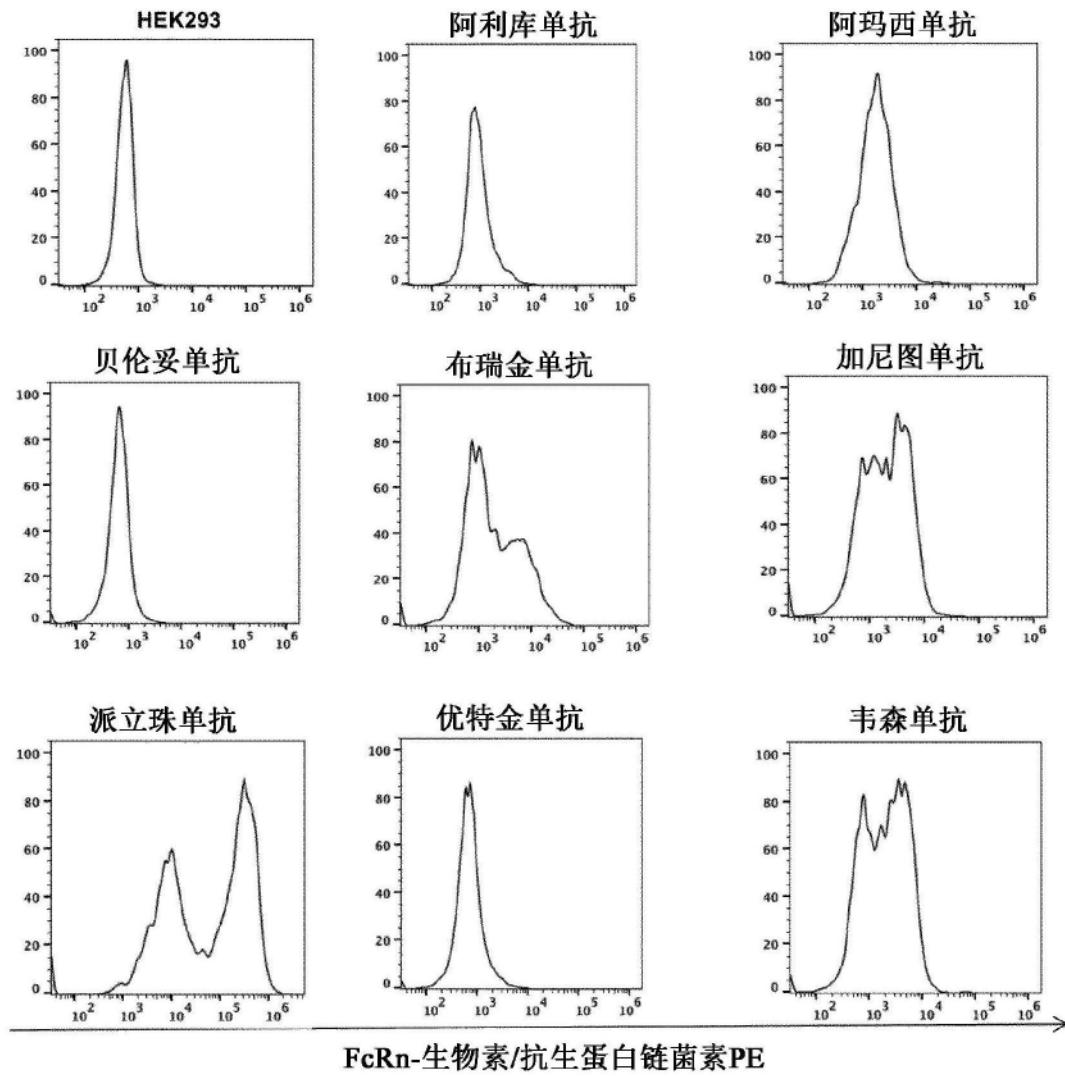


图36g

序列: PINT17-Tet 范围: 1516至7793

冬 37

图37(续)

图37(续)

2470            2480            2490            2500            2510            2520

CCA TGT CCT GCC CCT GAG CTG CTG GGC GGA CCC TCC GTG TTC CTG TTC CCT CCT AAG CCT  
 P C P A P E L L G G P S V F L F P P K P>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j j j j j >  
 P C P A P E L L G G P>  
 k k k k 铰链 k k k k k >

2530            2540            2550            2560            2570            2580

AAG GAC ACC CTG ATG ATC TCC CGG ACC CCT GAA GTG ACC TGC GTG GTG GTG GAC GTG TCC  
 K D T L M I S R T P E V T C V V V D V S>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j j j j >

2590            2600            2610            2620            2630            2640

CAC GAA GAT CCT GAA GTG AAG TTC AAT TGG TAC GTG GAC GGC GTG GAG GTG CAC AAC GCC  
 H E D P E V K F N W Y V D G V E V H N A>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j j j >

2650            2660            2670            2680            2690            2700

AAG ACC AAG CCT CGG GAG GAA CAG TAC AAC TCC ACC TAC CGG GTG GTG TCT GTG CTG ACC  
 K T K P R E E Q Y N S T Y R V V S V L T>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j j >

2710            2720            2730            2740            2750            2760

GTG CTG CAC CAG GAC TGG CTG AAC GGC AAA GAA TAC AAG TGC AAG GTG TCC AAC AAG GCC  
 V L H Q D W L N G K E Y K C K V S N K A>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j j >

2770            2780            2790            2800            2810            2820

CTG CCT GCC CCT ATC GAA AAG ACC ATC TCC AAG GCT AAG GGC CAG CCA CGG GAA CCT CAG  
 L P A P I E K T I S K A K G Q P R E P Q>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j >

2830            2840            2850            2860            2870            2880

GTC TAC ACA CTG CCT CCT AGC CGG GAC GAG CTG ACC AAG AAC CAG GTG TCC CTG ACC TGT  
 V Y T L P P S R D E L T K N Q V S L T C>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j >

2890            2900            2910            2920            2930            2940

CTG GTG AAG GGC TTC TAC CCT TCC GAT ATC GCC GTG GAG TGG GAG TCT AAC GGC CAG CCT  
 L V K G F Y P S D I A V E W E S N G Q P>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j >

2950            2960            2970            2980            2990            3000

GAG AAC AAC TAC AAG ACC ACC CCT CCT GTG CTG GAC TCC GAC GGC TCC TTC TTC CTG TAC  
 E N N Y K T T P P V L D S G S F F L Y>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j >

3010            3020            3030            3040            3050            3060

TCC AAG CTG ACC GTG GAC AAG TCC CGG TGG CAG CAG GGC AAC GTG TTC TCC TGC TCC GTG  
 S K L T V D K S R W Q Q G N V F S C S V>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j >

3070            3080            3090            3100            3110            3120

ATG CAC GAG GCC CTG CAC AAC CAC TAC ACC CAG AAG TCC CTG TCC CTG TCT CCT GGC AAG  
 M H E A L H N H Y T Q K S L S L S P G K>  
 j j j j j j j j j j IGG1 CH1-3 j j j j j j j j >

3130            3140            3150            3160            3170            3180

GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG AAT GCT GTG GGC CAG GAC ACG CAG GAG GTC  
 E Q K L I S E E D L> <N A V G Q D T Q E V>  
 1 1 1 MYC 抗原 1 1 1 >< m m m PDGFR TM m m m >

决定基

3190            3200            3210            3220            3230            3240

ATC GTG GTG CCA CAC TCC TTG CCC TTT AAG GTG GTG GTG ATC TCA GCC ATC CTG GCC CTG  
 I V V P H S L P F K V V V I S A I L A L>  
 m m m m m m m m m m PDGFR TM m m m m m m m >

图37(续)

>前聚腺苷酸

3310 3320 3330 | 3340 3350 3360  
TAG TAA AAG CTT GTC ACT TGG AAA GTA ATA GTT TTT CCT GCA CGG GTA GTA ATC AGC CTC  
\* \* >  
m >

3370	3380	3390	3400	3410	3420
GAC TGT GCC TTC TAG TTG CCA GCC ATC TGT TGT TTG CCC CTC CCC CGT GCC TTC CTT GAC	CCT GGA AGG TGC CAC TCC CAC TGT CCT TTC CTA ATA AAA TGA GGA AAT TGC ATC GCA TTG	TCT GAG TAG GTG TCA TTC TAT TCT GGG GGG TGG GGT GGG GCA GGA CAG CAA GGG GGA GGA			

<BGH 聚腺苷酸

3550 3560 3570 3580 3590 3600  
TTG GGA AGA CAA TAG CAG GCA TGC TGG GGA TGG CCC GGG CAT GAT AAC TTC GTA TAA TGT

>人类\EF1a\启动子

3610	3620	3630	3640	3650	3660
ATG CTA TAC GAA GTT ATG TAG AGT AAT TCA TAC AAA AGG ACT CGC CCC TGC CTT GGG GAA					
TCC CAG GGA CCG TCG TTA AAC TCC CAC TAA CGT AGA ACC CAG AGA TCG CTG CGT TCC CGC					
CCC CTC ACC CGC CCG CTC TCG TCA TCA CTG AGG TGG AGA AGA GCA TGC GTG AGG CTC CGG					
TGC CCG TCA GTG GGC AGA GCG CAC ATC GCC CAC AGT CCC CGA GAA GTT GGG GGG AGG GGT					
CGG CAA TTG AAC CGG TGC CTA GAG AAG GTG GCG CGG GGT AAA CTG GGA AAG TGA TGT CGT					
GTA CTG GCT CGG CCT TTT TCC CGA GGG TGG GGG AGA ACC GTA TAT AAG TGC AGT AGT CGC					
CGT GAA CGT TCT TTT TCG CAA CGG GTT TGC CGC CAG AAC ACA GGT AAG TGC CGT GTG TGG					
TTC CCG CGG GCC TGG CCT CTT TAC GGG TTA TGA CCC TTG CGT GCC TTG AAT TAC TTC CAC					
GCC CCT GGC TGC AGT ACG TGA TTC TTG ATC CCG AGC TTC GGG TTG GAA GTG GGT GGG AGA					
GTT CGA GGC CTT GCG CTT AAG GAG CCC CTT CGC CTC GTG CTT GAG TTG AGG CCT GGC TTG					
GGC GCT GGG GCC GCC GCG TGC GAA TCT GGT GGC ACC TTC GCG CCT GTC TCG CTG CTT TCG					
ATA AGT CTC TAG CCA TTT AAA ATT TTT GAT GAC CTG CTG CGA CGC TTT TTT TCT GGC AAG					
ATA GTC TTG TAA ATG CGG GCC ATG ATC TGC ACA CTG GTA TTT CGG TTT TTG GGG CCG CGG					
GCG GCG ACG GGG CCC GTG CGT CCC AGC GCA CAT GTT CGG CGA GGC GGG GCC TGC GAG CGC					
GGC CAC CGA GAA TCG GAC GGG GGT AGT CTC AAG CTG GCC CTG CTC TGG TGC CTG CGC					
TCG CGC CGC CGT GTA TCG CCC CGC CCT GGG CGG CAA GGC TGG CCC GGT CGG CAC CAG TTG					
CGT GAG CGG AAA GAT GGC CGC TTC CGG GCC CTG CTG CAG GGA GCT CAA AAT GGA GGA CGC					
GGC GCT CGG GAG AGC GGG CGG GTG AGT CAC CCA CAC AAA GGA AAA GGG CCT TTC CGT CCT					
CAG CCG TCG CTT CAT GTG ACT CCA CGG AGT ACC GGG CGC CGT CCA GGC ACC TCG ATT AGT					
TCA CGA GCT TTT GGA GTA CGT CGT CTT TAG GTT GGG GGG AGG GGT TTT ATG CGA TGG AGT					
TTC CCC ACA CTG AGT GGG TGG AGA CTG AAG TTA CGC CAG CTT GGC ACT TGA TGT ATT TCT					
CCT TGG AAT TTG CCC TTT TTG AGT TTG GAT CTT GGT TCA TTC TCA AGC CTC AGA CAG TGG					
TTC AAA GTT TTT TTC TCC CAT TTC AGG TGT CGT GAG AGC TCG TTT AGT GAA CGG TCA GAT					
CGC CTG GAG ACG CCA TCC ACG CTG TTT TGA CCT CCA TAG AAG ACA CGG GGA CGG ATC CAG					

5050 5060 5070 5080 5090 5100  
 CCT CCG CGG CCC CGA ATT CAC CAT GTC TAG ACT GGA CAA GAG CAA AGT CAT AAA CTC TGC  
 M S R L D K S K V I N S A>  
 n n n n n TET-开\3G n n n n n >

5170            5180            5190            5200            5210            5220  
 GCT GGG AGT TGA GCA GCC TAC CCT GTA CTG GCA CGT GAA GAA CAA GCG GGC CCT GCT CGA  
 L G V E Q P T L Y W H V K N K R A L L D>  
 n n n n n n TET-开\3G n n n n n n n n n n n n n n >

图37(续)

5230 5240 5250 5260 5270 5280  
 TGC CCT GCC AAT CGA GAT GCT GGA CAG GCA TCA TAC CCA CTC CTG CCC CCT GGA AGG CGA  
 A L P I E M L D R H H T H S C P L E G E>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5290 5300 5310 5320 5330 5340  
 GTC ATG GCA AGA CTT TCT GCG GAA CAA CGC CAA GTC ATA CCG CTG TGC TCT CCT CTC ACA  
 S W Q D F L R N N A K S Y R C A L L S H>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5350 5360 5370 5380 5390 5400  
 TCG CGA CGG GGC TAA AGT GCA TCT CCG CAC CCG CCC AAC AGA GAA ACA GTA CGA AAC CCT  
 R D G A K V H L G T R P T E K Q Y E T L>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5410 5420 5430 5440 5450 5460  
 GGA AAA TCA GCT CGC GTT CCT GTG TCA GCA AGG CTT CTC CCT GGA GAA CGC ACT GTA CGC  
 E N Q L A F L C Q Q G F S L E N A L Y A>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5470 5480 5490 5500 5510 5520  
 TCT GTC CGC CGT GGG CCA CTT TAC ACT GGG CTG CGT ATT GGA GGA ACA GGA GCA TCA AGT  
 L S A V G H F T L G C V L E E Q E H Q V>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5530 5540 5550 5560 5570 5580  
 AGC AAA AGA GGA AAG AGA GAC ACC TAC CAC CGA TTC TAT GCC CCC ACT TCT GAA ACA AGC  
 A K E E R E T P T T D S M P P L L K Q A>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5590 5600 5610 5620 5630 5640  
 AAT TGA GCT GTT CGA CCG GCA GGG AGC CGA ACC TGC CTT CCT TTT CGG CCT GGA ACT AAT  
 I E L F D R Q G A E P A F L F G L E L I>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5650 5660 5670 5680 5690 5700  
 CAT ATG TGG CCT GGA GAA ACA GCT AAA GTG CGA AAG CGG CGG GCC GAC CGA CGC CCT TGA  
 I C G L E K Q L K C E S G G P T D A L D>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5710 5720 5730 5740 5750 5760  
 CGA TTT TGA CTT AGA CAT GCT CCC AGC CGA TGC CCT TGA CGA CTT TGA CCT TGA TAT GCT  
 D F D L D M L P A D A L D D F D L D M L>  
 n n n n n n n n TET-开\3G n n n n n n n n >  
  
 5770 5780 5790 5800 5810 5820  
 GCC TGC TGA CGC TCT TGA CGA TTT TGA CCT TGA CAT GCT CCC CGG GTA ACT AAG TAA GGA  
 P A D A L D D F D L D M L P G \*>  
 n n n n n n TET-开\3G n n n n n n n n >  
  
 >SV40\聚腺苷酸信号  
 |  
 5830 | 5840 5850 5860 5870 5880  
 TCC AGA CAT GAT AAG ATA CAT TGA TGA GTT TGG ACA AAC CAC AAC TAG AAT GCA GTG AAA  
 AAA ATG CTT TAT TTG TGA AAT TTG TGA TGC TAT TGC TTT ATT TGT AAC CAT TAT AAG CTG  
 CAA TAA ACA AGT TAA CAA CAA CAA TTG CAT TCA TTT TAT GTT TCA GGT TCA GGG GGA GGT  
 GTG GGA GGT TTT TTA AAG CAA GTA AAA CCT CTA CAA ATG TGG TAT GGC TGA TTA TGA TCC  
 TGC AAG CCT CGT CGT CCT CGG CGG ACC ACG CTA TCT GTG CAA GGT CCC CGG CCC CGG ACG  
 CGC GCT CCA TGA GCA GAG CGC CGG CCG AGG CGA AGA CTC GGG CGG CGC CCT GCC CGT  
 CCC ACC AGG TCA ACA GGC GGT AAC CGG CCT CTT CAT CGG GAA TGC GCG CGA CCT TCA GCA  
  
 BstZ171  
 6250 6260 6270 16280  
 TCG CCG GCA TGT CCC CCT GGC GGA CGG GAA GTA TGT ATA C

图37(续)

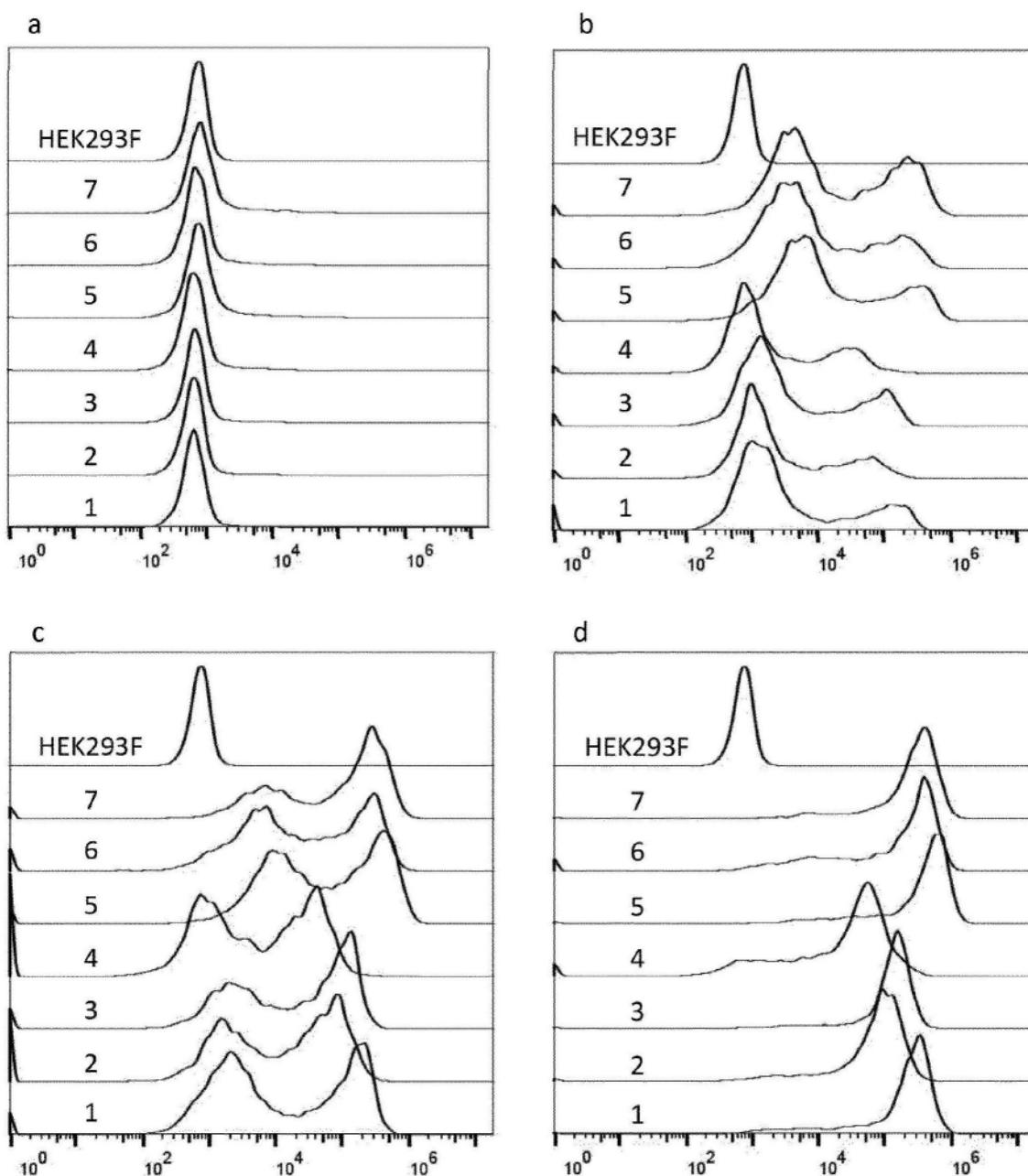


图38

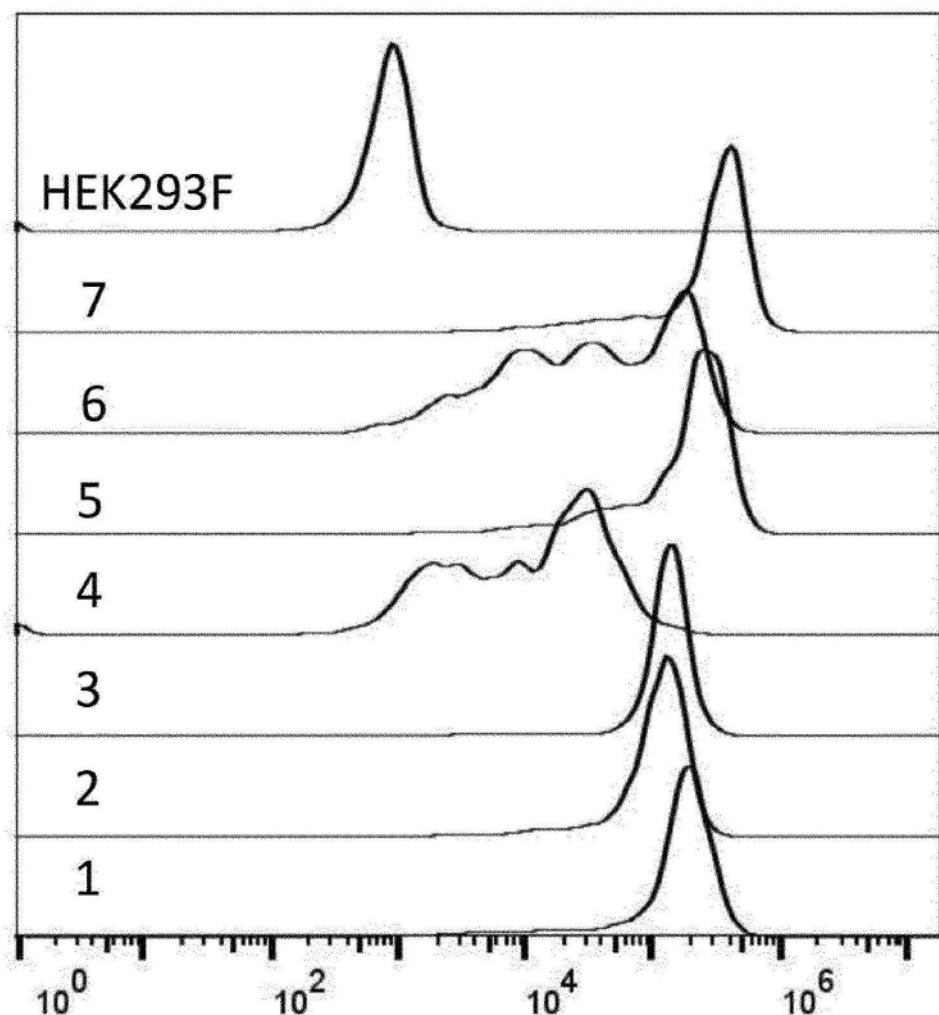


图39

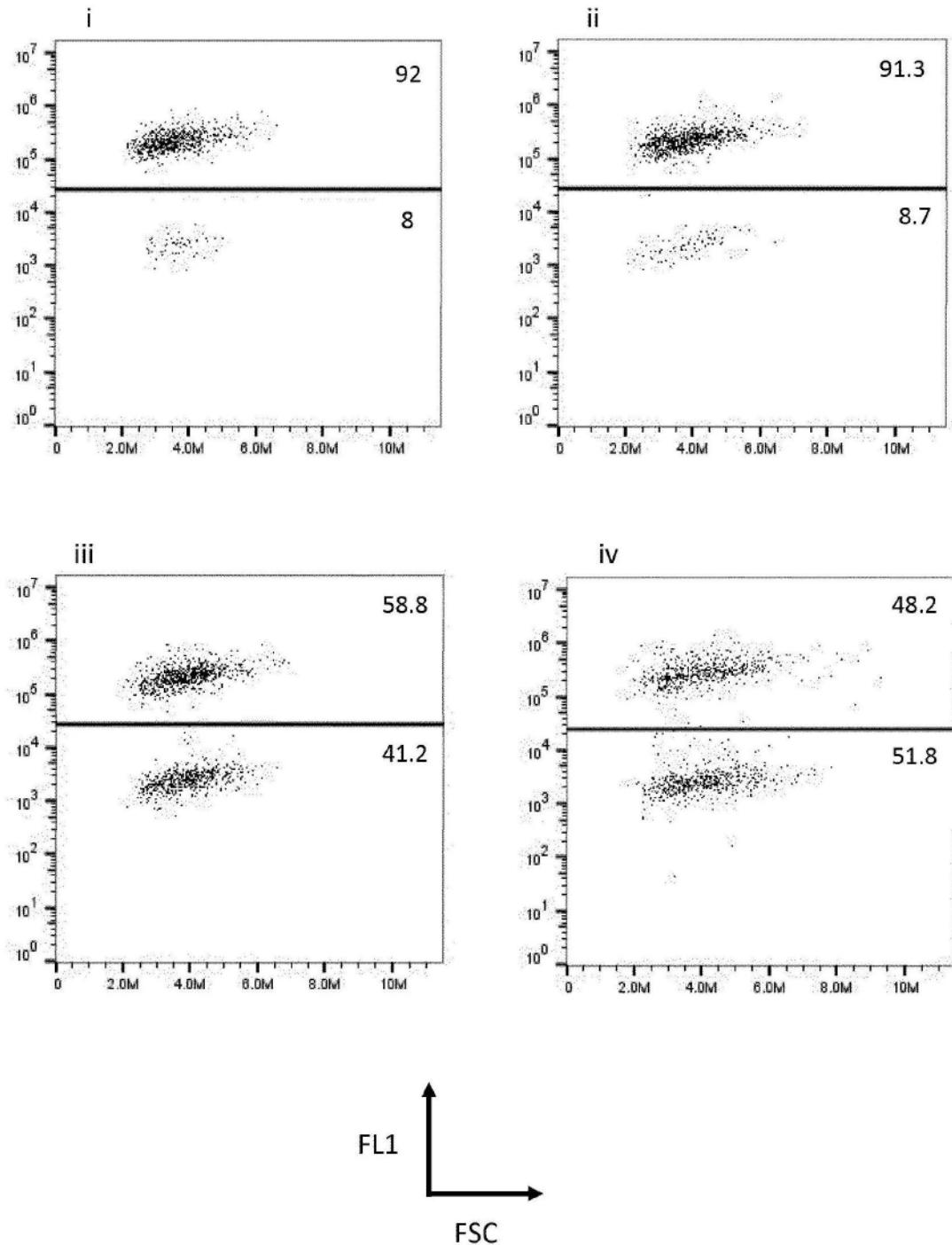


图40a

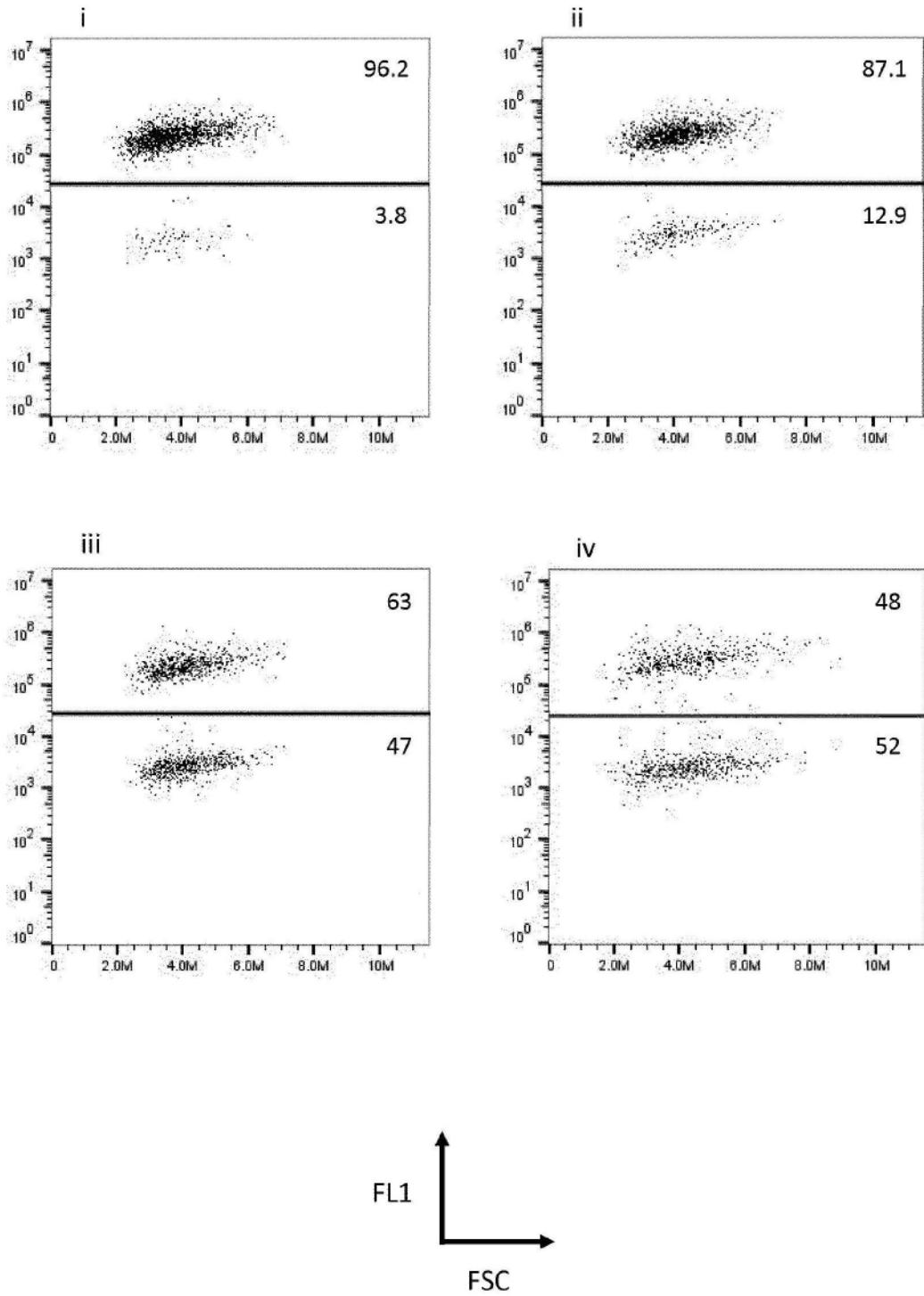


图40b

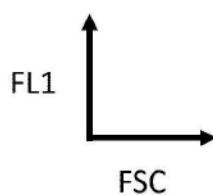
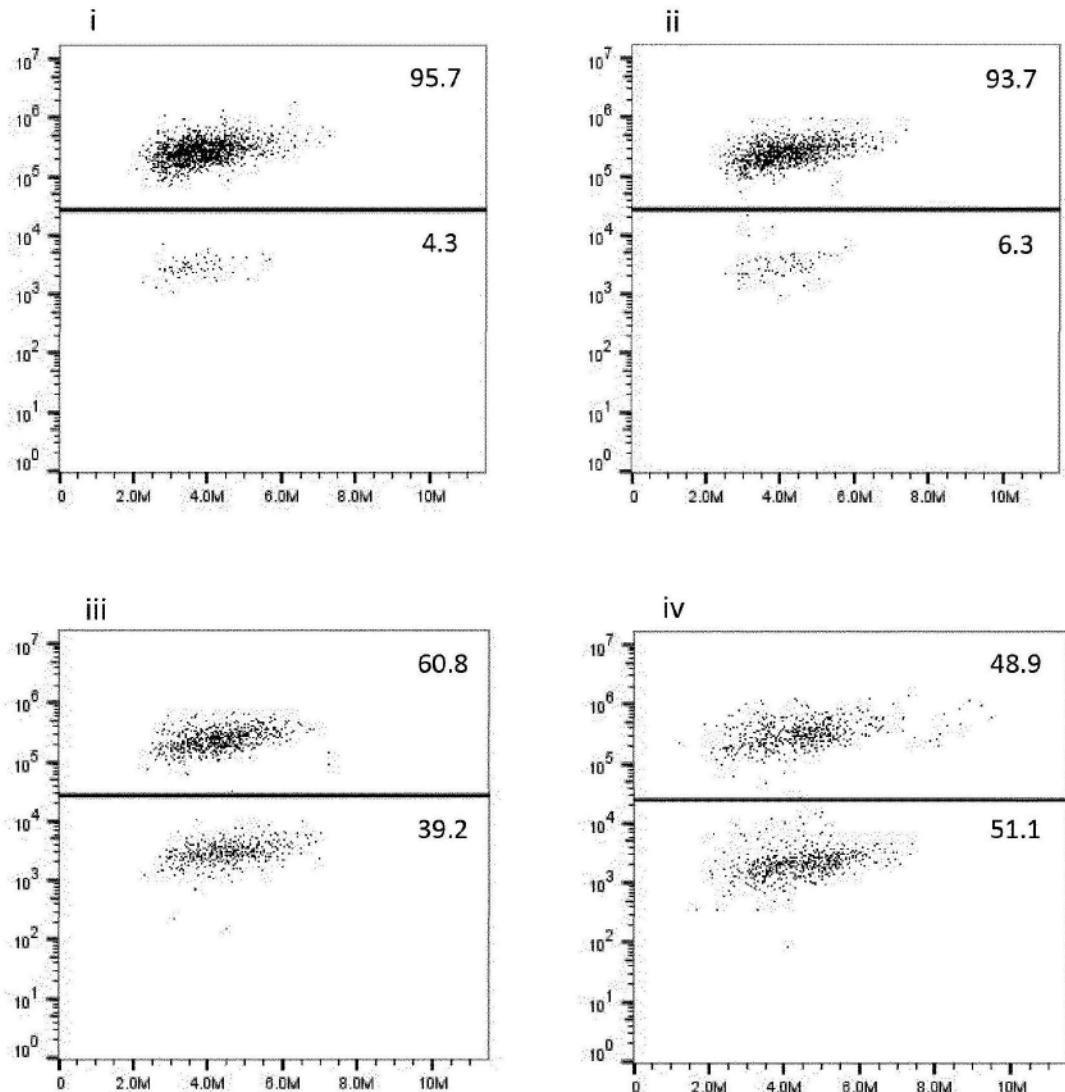
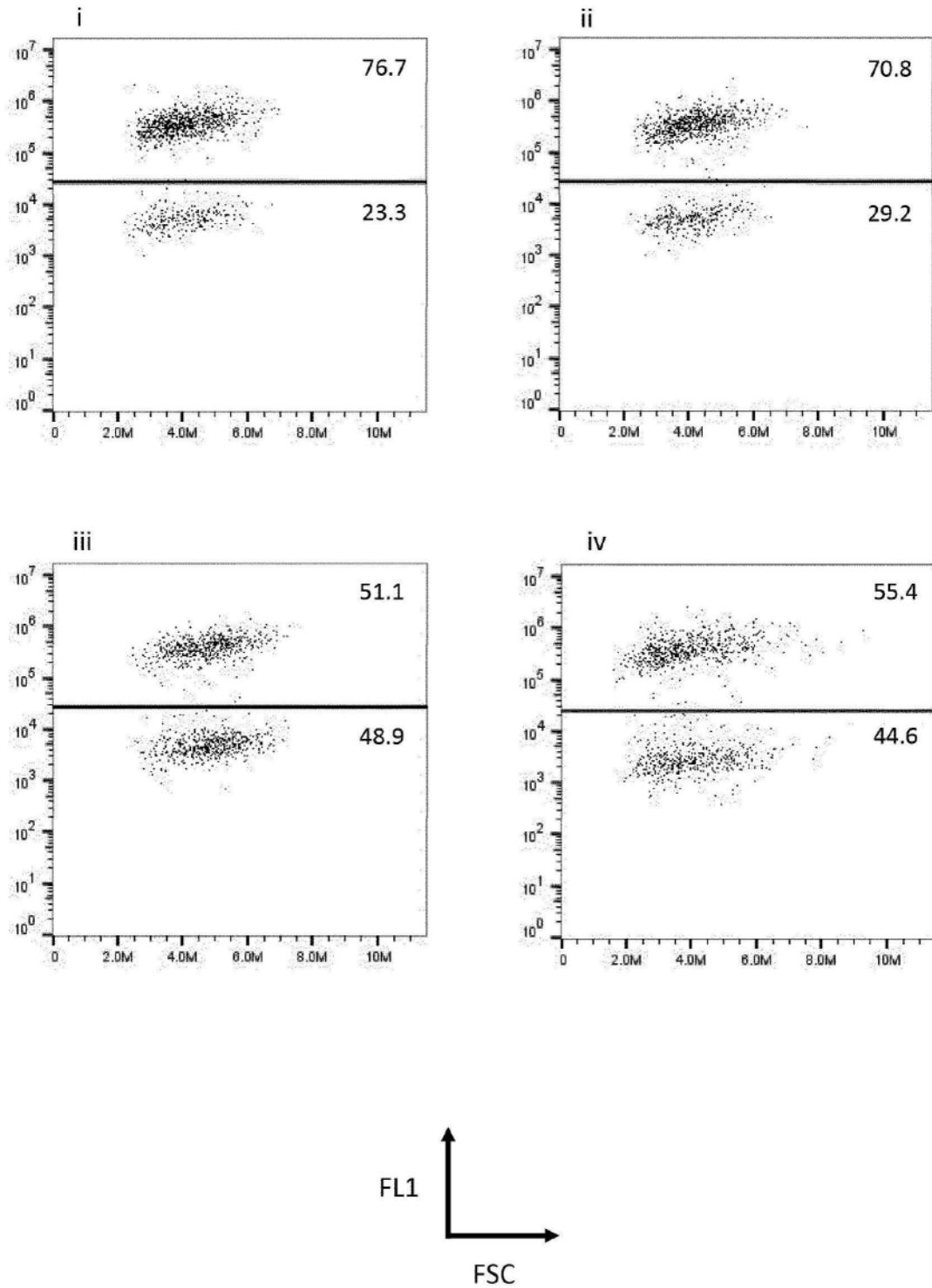


图40c



图|40d

**a** 优特金单抗和布瑞金单抗的重叠图

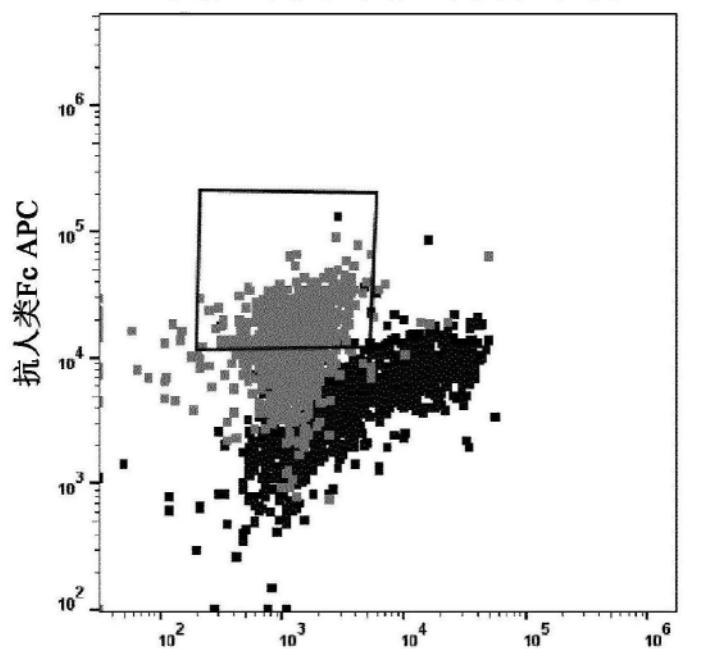


图41

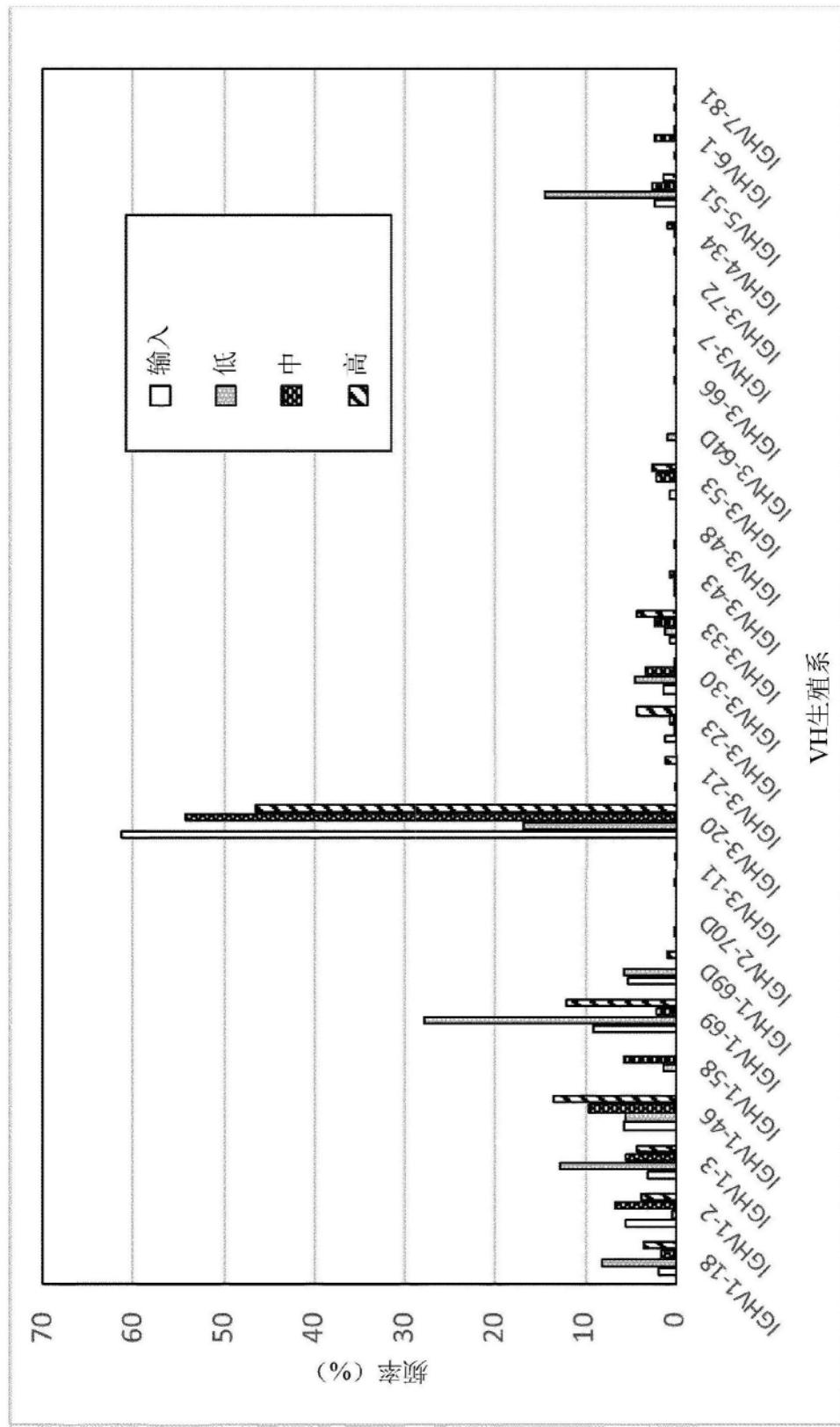
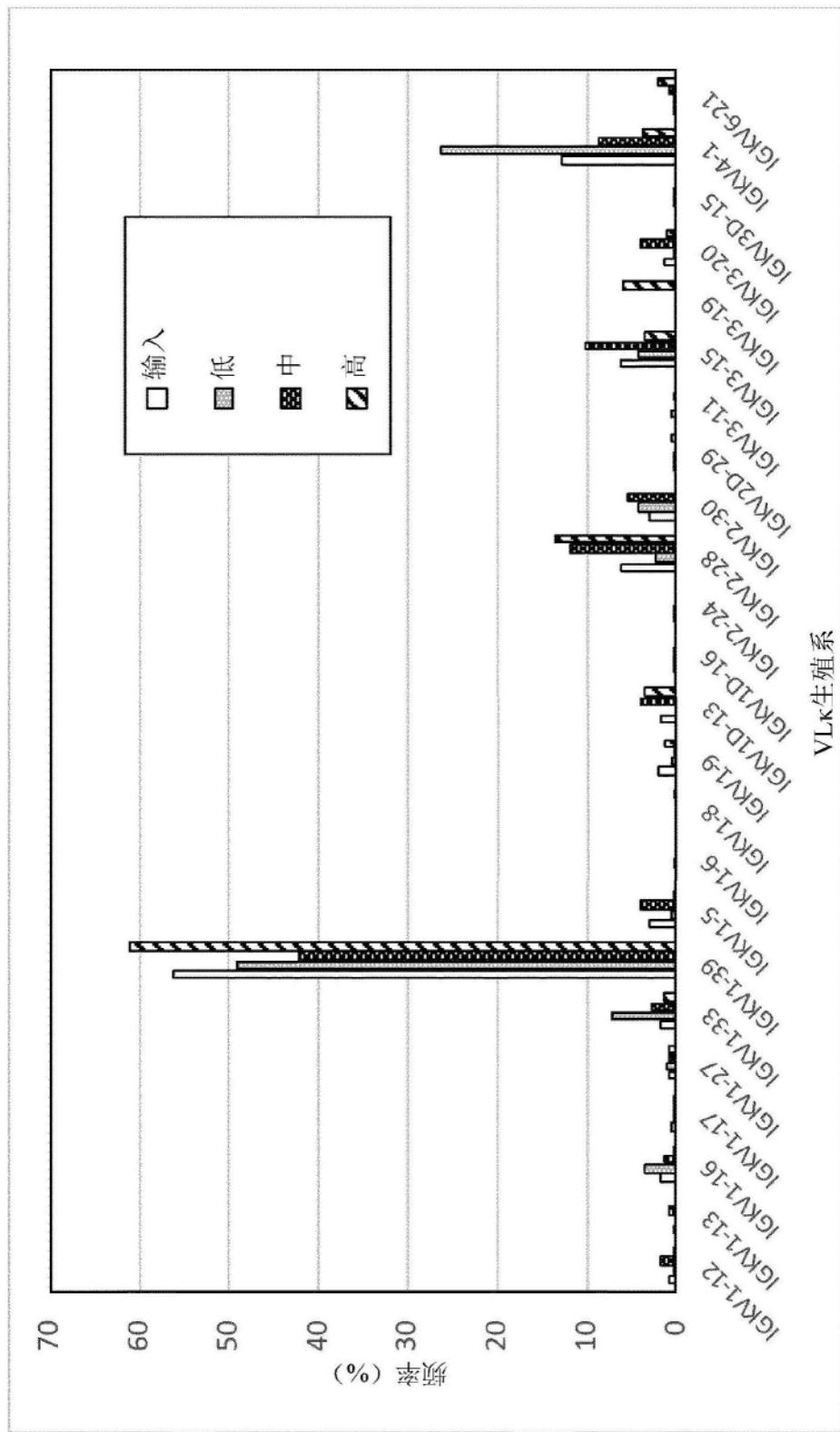


图42



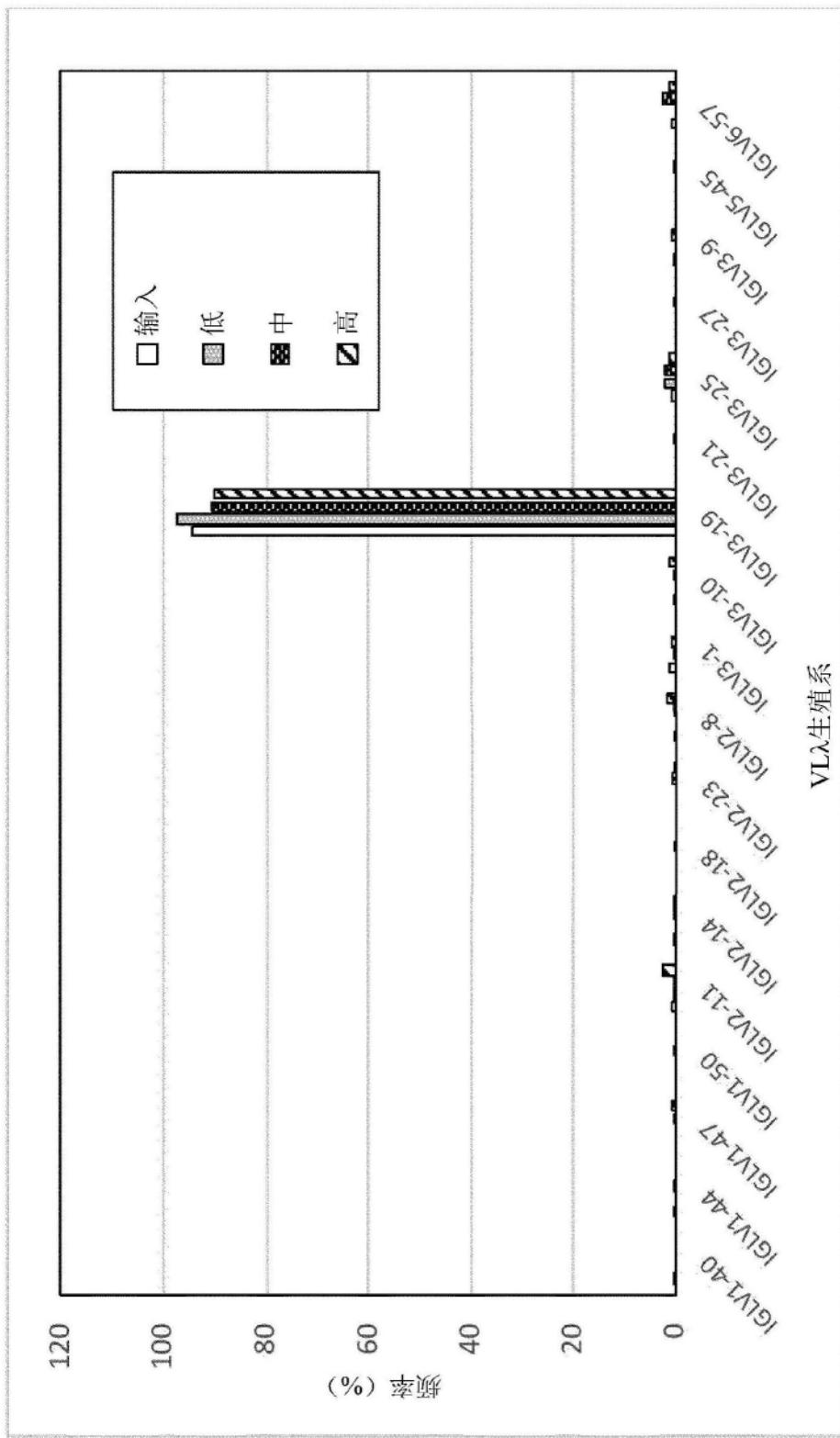


图44

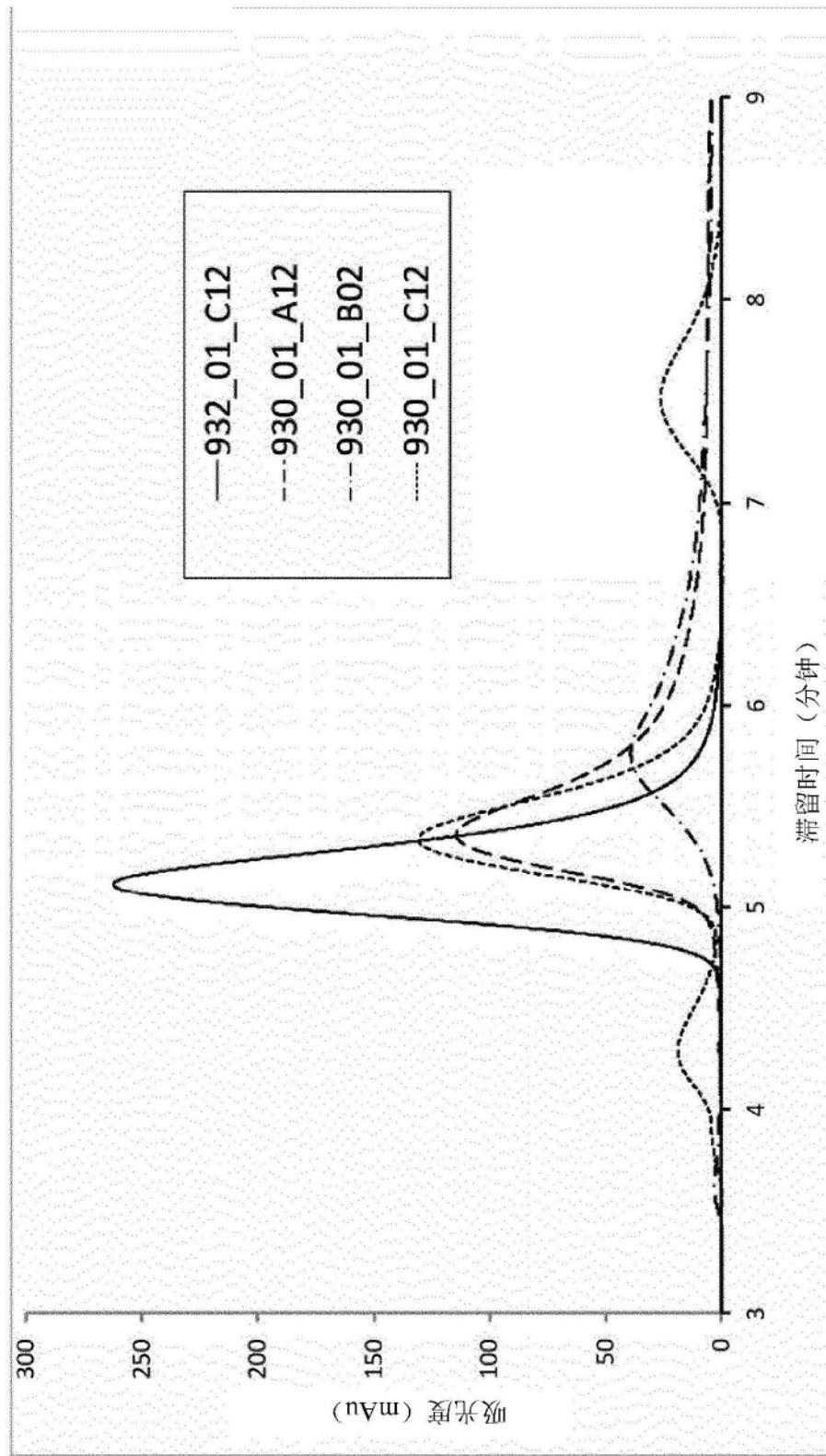


图45

	CRISPR2				
	1410	1420	1430	1440	1450
人类	CAGGGCCGGT	TAATGTGGCT	CTGGTTCTGG	GTACTTTAT	CTGTCCCCCTC
CHO	CGGGATC--C	TTCTG-----	--GATTGGGG	ATGCTTTAT	CT---CCCCGT
	* * * *	* **	* ***	* *****	* ****
CRISPR1					
CRISPR2					
	1460	1470	1480	1490	1500
人类	CACCCCACAG	-TGGGGCCAC	TAGGGACAGG	ATTGGTGAC	AGAAAAGCCCC
CHO	CATCCAAAAG	CTGGCATTGT	CAGGGACAAG	ATTAGTCAC	AGCAGAGCCCC
	* * * *	***	*****	* ***	* * * * *
CRISPR3					

图46

序列: pINT17-BSD-CHO-D1.3 范围: 1至960和6321至7360

>CHO-AAVS-左臂

```

>3199      >Ex-DraII
|           |
10          20          30          40          50          60          70          80
GCGATCGCGA TGGCTTACAT CCCGTGCCTT TCCAGGCTGG TGGCTGCCCT GGCTGTGCCG CAGGCTTCCA GGGCCCAGCT

         >3197      >Ex-NheI      <3200
|           |           |
90          100         110         120         130         140         150         160
CTGACTTGAC GCCCCCCCCC CAATCCCCCA CTCCCTCCTC TGACTCTAGC CAGGCCAGG CCCTTCAGTG TCACTTCTTT
TGGGGGGTCC ACCTTGTTC TCACTCCACT TCCTGTGACC CGTGTGCTCC GCTGTGGCCT CAGGAGGGTC CTTTGCCTCCT
CCAAGGGCACT GTACCTCTTT GTTCCCCCTGG AGAAAGGGCG CTCCACCGTC TTGTTGGCCC CATCCCATAA CCTTCCCTCCC
CTCCCCATGT CCTATTAAATT CAATTCTCGG CTCATCCCCCT TTGGTGATC TTGCCACCCA CCCCCACCCCC CCATAGTTT
CTCTCTCAACT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTGCTCTCTCT CTCTCTGTGT
CCACATTGAT TTGCTTCTCTT GGAACACTT AAGCATCTGG CCTTCCAAGC CCTATTCTCTC TCTCCTCTGG TTTTGTTTTT
GTTGTTGTTT TTGTTTTTTT CCCCTGTGTA GCCCAGGGCA ACCTCAAATA CAGGTCTGTG GTCTCTGACCC CCCTCATTTG

```

>3195 >Ex-NcoI <3198

```

650          660         670         680         690         700         710         720
TGGGATTACA GGCATGCCCA CTCCACCTAC CACCTCATGG ACTATATTTG GGCAGCATCC TGCTTTAAAAA ACTATGCCCT
GTCCTCTTAGA TTTCACCCCT GTCACCTCT AGAACCTTACCC CCACCAGGAC CCTGGGGCAG TCAGACACTG TGATTATCTT

```

>AAVS-TALE-左

```

810          820         830         840         850         860         870         880
TTTCCAGTGA TCTGTGCCCT GAGAGTGGAG ACCATCTTGT ATTGCCGGGA TCCTCTGGGA TTCGGGATGC TTTTATCTCC

```

<3196

>AAVS-左端	>分支位点	>剪接受体
890	900	930
900	910	940
910	920	950
CCGTCATCCA	AAAGCTGGCA	CATGGCAAAA CCTCTGAGCC
M	A	K
		P
b	SYN内含子	L
b	>	S>
杀稻瘟菌素		

>3201

>CHO-AAVS-右臂

```

>loxP
| 
6330        6340        6350        6360        6370        6380        6390        6400
GGGCATGATA ACTTCGTATA ATGATATGCTA TACGAAGTTA TGATACGGC GCGCCCTGTC GGGACAAGAT TAGTCACAGC
AGAGCCCCCA TCCCCCTCTC TCTTCCTCTC GGCCACAGTG TTGGTTCACT TTCACTCTCTT CATGGCAACT TCCATCTCTC
GGGTGTCGT CCTCTTGCTA GGATGCTCGT GGAGGTGGGG AATAAAGGTT CAGTGTTCAGT CTCATCCAGG CAGAGGAAAA
ACAGAACCGC TCTGTCTCTGG GCTGGGCTGT CTTTTAGACA GACTGCTTGC CTAGACTCCA GGAAGCCTCA CGTTCCAACC
CCAATACCCG AGAAATCAGG TGCTTGGAAA GTAGAGGCAG GAGGGTTGGA ATTCACTGTC ACACAGTGAC
TTGAGGCCA TCTTGATAGC CCACATGAGA CCCAGTGCA AAATCTTGAC CAAACAAAAA AAGTCCGTGC TGACTTTTG
TGTCAGTCTG CTTCCTGTGT CTTGATCTT GAGTGTCTT TCTCGGTGCG CCTTACTTTT TTATACCTGTA GAATGGGACC
AGTGGTCACC CCTGTCCCCA GAGGGACTCT ATGGTGAGAG GCGTCCTGTG GAAAATTC CTTATGGGGC CGGCCTGAGC
TGTGTGGAC ATCGTCCTCT ATATATACCA GGGTCTGTG TTCCCGAGGC CACACTCAGG GTGTCAGCGT CTTTGACCGAG

```

>3203 >Ex-BclI <3202

```

7050        7060        7070        7080        7090        7100        7110        7120
TCCAGGAGAT CAGCTCAGGC TGCTCACATT ATCACCAAAG TCCCTGGCCC TCCCAAGAGT TCCCTCCCAA CTGCATCCCC
TTCTCTGCA TCTGCTGGAG GTCCCTAGAG CATCCTCTCT GACAGGAACC TGTGACTCTA GGCGCTGACGT GTCAGCCTC

```

>Sbf1

```

7210        7220        7230        7240        7250        7260        7270        7280
CAGGTGGAC TCTCCTCCCG CCATGCAGAC ACCTTGGTA CCTGACTCTT CAGGCCTTTC CAGGAGCTG CAGGCGATCT

```

>3' β 球蛋白隔离子

```

7290        7300        7310        7320        7330        7340        7350        7360
CTCGATCTCT CGATTCGAT CAAGACATTC CTTTAATGGT CTTTCTGGA CACCACTAGG GGTCAGAACT AGTTCATCAA

```

图47

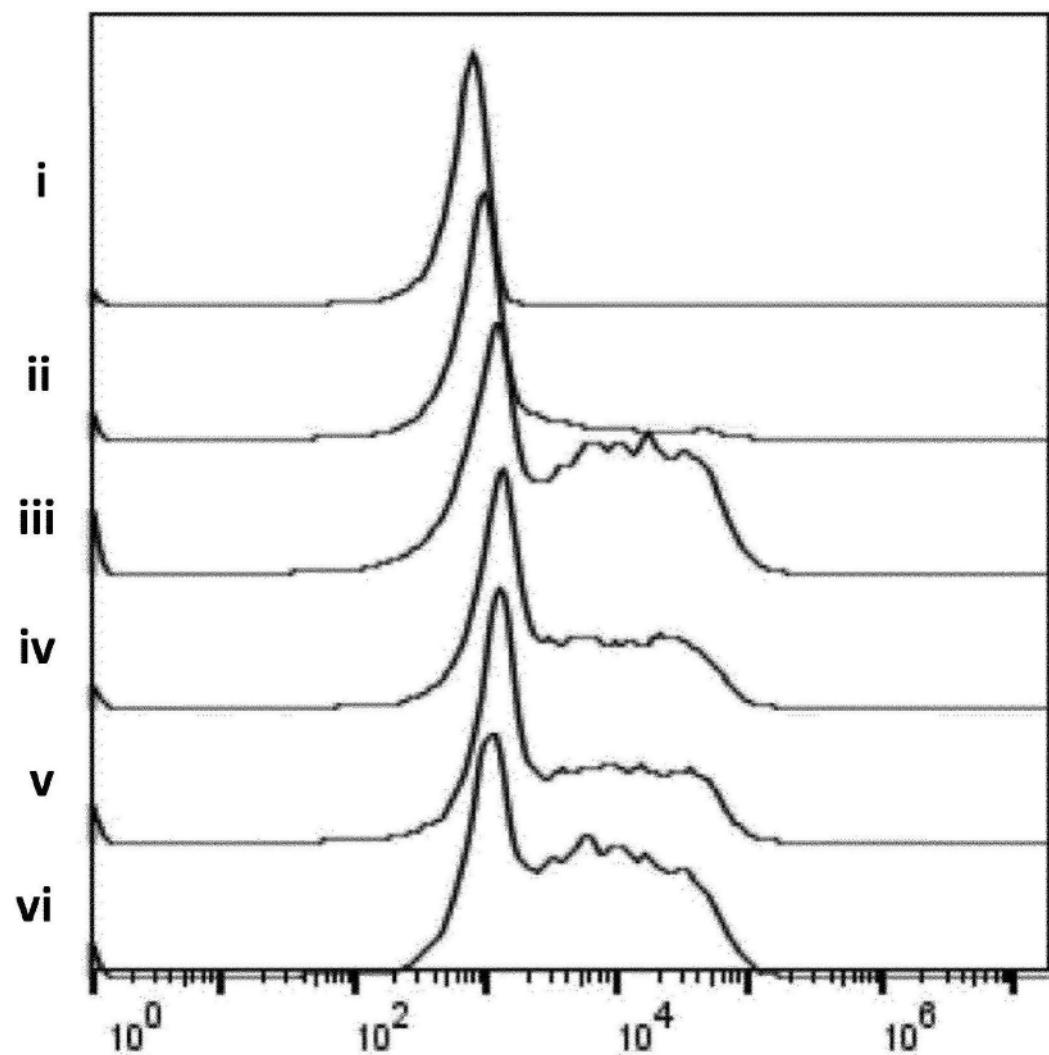


图48

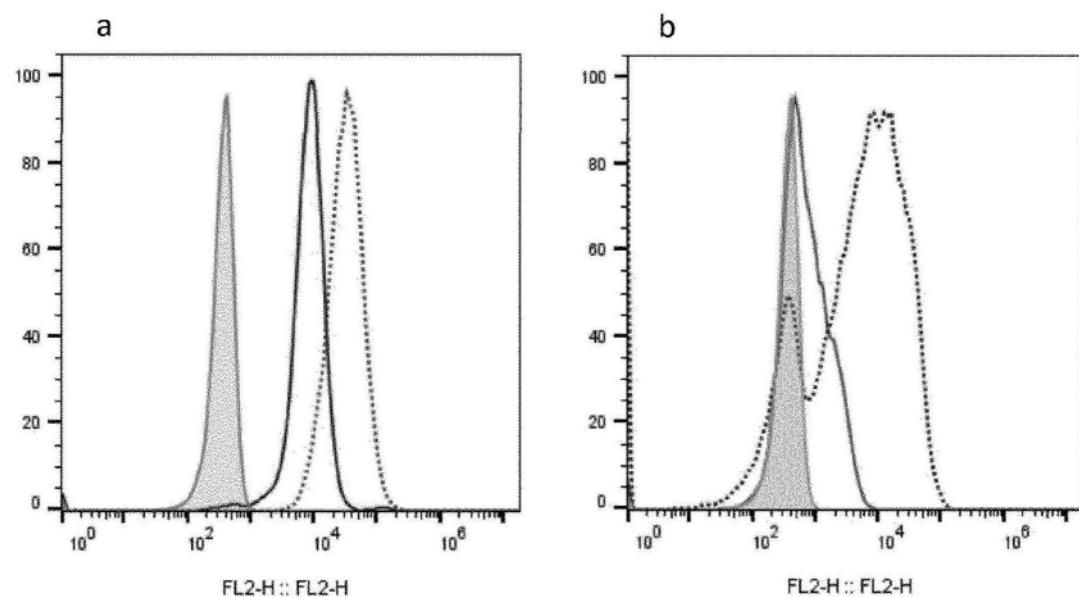


图49

i

E V Q L V Q S G A E V  
 GCCTTTCTCTCCACA GGCGCCATGGCC GAA GTG CAG CTG GTG CAG TCT GGA GCT GAG GTG  
NcoI

CDR1  
 K R P G A S V T V S C K A S G Y T F T N  
 AAG AGG CCT GGG GCC TCA GTG ACA GTT TCC TGC AAG GCT TCC GGT TAC ACC TTT ACC AAC

H G I S W V R Q A P G Q G L E W M G W N  
 CAC GGT ATC AGC TGG GTG CGA CAG GCC CCT GGA CAA GGC CTT GAG TGG ATG GGA TGG AAC

CDR2  
 S P Y N G N T N Y A Q R F Q G R V T M T  
 AGC CCT TAC AAT GGA AAC ACA AAC TAT GCA CAG AGG TTC CAG GGC AGA GTC ACC ATG ACC

T D T S T N T A Y M E L R T L T S D D T  
 ACA GAC ACA TCC ACG AAC ACA GCC TAC ATG GAG CTG AGG ACC CTG ACA TCT GAC GAC ACG

CDR3  
 A M Y Y C A R D R D Y Y D A G S Y W G Q  
 GCC ATG TAT TAC TGT GCC AGA GAT AGG GAT TAC TAT GAT GCG GGG AGC TAC TGG GGC CAG

G T L V T V S S a s t k g p s v  
 GGA ACC CTG GTC ACC GTC TCG AGT GCC AGC ACC AAG GGC CCC AGC GTG  
XhoI \_\_\_\_\_ CH1 \_\_\_\_\_ >

ii

A S S Y E L T Q P P S V S V S P G Q  
NheI  
 GCT AGC TCC TAT GAG CTG ACT CAG CCA CCC TCG GTG TCA GTA TCC CCA GGA CAG

CDR1  
 T A R I T C S G D S L P K Q Y A Y W Y Q  
 ACG GCC AGG ATC ACC TGC TCT GGA GAT TCA TTG CCA AAG CAA TAT GCT TAT TGG TAC CAG

CDR2  
 Q K P G Q A P V L V I Y K D S E R P S G  
 CAA AAG CCA GGC CAG GCC CCT GTA TTA GTG ATA TAT AAA GAC AGC GAG AGG CCC TCA GGG

CDR3  
 I P E R F S G S G T T V T L T I S G  
 ATC CCT GAG CGA TTC TCT GGC TCC GGC TCA GGG ACA ACA GTC ACA TTG ACC ATC AGT GGA

V Q A E D E A D Y Y C Q S A D S D N A F  
 GTC CAG GCA GAA GAC GAG GCT GAC TAT TAT TGT CAA TCA GCA GAC AGT GAT AAT GCT TTT

V F G R G T K L T V L G Q P A A A  
GTC TTT GGA AGA GGG ACC AAG CTG ACC GTC CTA GGT CAG CCC GCG GCC GCT

图|50a

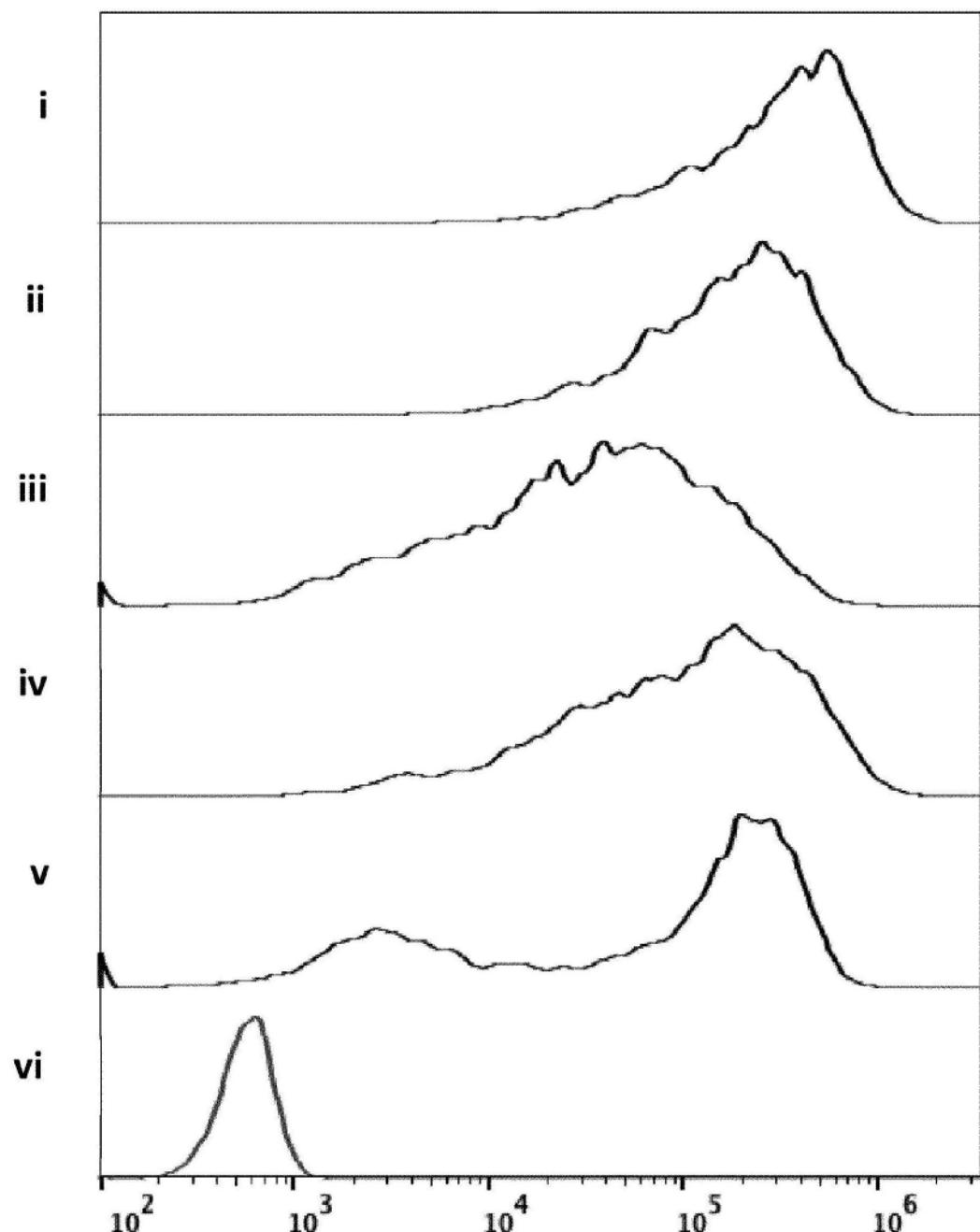


图50b

序列: pINT17-Bi-CMV-艾密次单抗: 1514至8904

BglII >BGH 聚腺苷酸

	10	20	30	40	50	60
GAG ATC TTC CCC AGC ATG CCT GCT ATT GTC TTC CCA ATC CTC CCC CTT GCT GTC CTG CCC						
CAC CCC ACC CCC CAG AAT AGA ATG ACA CCT ACT CAG ACA ATG CGA TGC AAT TTC CTC ATT						
TTA TTA GGA AAG GAC AGT GGG AGT GGC ACC TTC CAG GGT CAA GGA AGG CAC GGG GGA GGG						
<b>&lt;BGH_聚腺苷酸</b>						
	190	200	210	220	230	240
GCA AAC AAC AGA TGG CTG GCA ACT AGA AGG CAC AGT CGA GGC TCT AGA TTA ACA TTC						
<* C E						
< b b b						
250	260	270	280	290	300	
GCC CCG GTT GAA GCT CTT GGT CAC AGG GCT AGA AAG GCC CTG GTG GGT CAC TTC GCA GGC						
<G R N F S K T V P S S L G Q H T V E C A						
< b b b b b b b CL b b b b b b b b b b b b b b b b b						
310	320	330	340	350	360	
GTA CAC TTT GTG CTT CTC GTA GTC GGC CTT GCT CAG GGT CAG TGT GCT GCT CAG GCT GTA						
<Y V K H K E Y D A K S L T L T S S L S Y						
< b b b b b b b CL b b b b b b b b b b b b b b b						
370	380	390	400	410	420	
GGT AGA GTC CTT GCT GTC CTG CTC GGT CAC GCT CTC TTG GCT ATT GCC GCT CTG GAG GGC						
<T S D K S D Q E T V S E Q S N G S Q L A						
< b b b b b b b CL b b b b b b b b b b b b b						
430	440	450	460	470	480	
GTT GTC CAC CTT CCA CTG CAC CTT GGC TTC TCT GGG GTA GAA GTT GTT CAG CAG GCA CAC						
<N D V K W Q V K A E R P Y F N N L L C V						
< b b b b b b b CL b b b b b b b b b b b b b						
490	500	510	520	530	540	
GAC AGA GGC TGT GCC AGA CTT CAG CTG CTC GTC GCT AGG TGG AAA GAT GAA CAC GCT AGG						
<V S A T G S K L Q E D S P P F I F V S P						
< b b b b b b b CL b b b b b b b b b b b b						
550	560	570	580	590	600	
GGC GGC CGC TGT CCG CTT GAT TTC CAC CTT GGT GCC TCC GCC AAA TGT CAG TGG AGG GTC						
<A A A T R K I E V K T G G G F T L P P D						
< b < c c c c c c c c VL c c c c c c c c c						
610	620	630	640	650	660	
GCT GTA CTG CTG GCA GTA GTC GGT GGC GAT ATC CTC AGG CTG GAG GCT GGA TAT TGT CAG						
<S Y Q Q C Y Y T A I D E P Q L S S I T L						
< c c c c c c c c c c VL c c c c c c c c c						
670	680	690	700	710	720	
GGT GAA GTC GGT GCC GTA TCT GCT GCC GCT GAA TCT ATC GGG CAC GCC GCT TTC TTT TCT						
<T F D T G Y R S G S F R D P V G S E K R						
< c c c c c c c c c c VL c c c c c c c c c						
730	740	750	760	770	780	
GCT GGC CTG ATA GAT CAG CAG CTC AGG AGC CTG TCC AGG CTT CTG CTG ATA CCA GGC CAG						
<S A Q Y I L L E P A Q G P K Q Q Y W A L						
< c c c c c c c c c c VL c c c c c c c c c						
790	800	810	820	830	840	
CTG TCT CTC GAT GTT CCG GCT GGC CTT GCA TGT GAT GGT CAC TCT GTC TCC CAC GCT GGC						
<Q R E I N R S A K C T I T V R D G V S A						
< c c c c c c c c c c VL c c c c c c c c c						

图51

>RG                    <剪接受体

850                860                870                | 880                890                900

AGA CAG GCT GCT AGG GCT CTG TGT CAT CTG GAT GTC TCC TCT GCT AGC TGA AAA TAT ACA  
 <S L S S P S Q T M Q I D G R  
 <—c—c—c—c—c—VL—c—c—c—c—c—c—

910                920                930                940                950                960

GCA AAC ATC AGT ACA ACA TAA ATA TCT GTG TAT GAA AAT CAC CTT TAA TCT TGC TAG ACA  
 TGA AGA AAG AAT ATG CAA TAC ATT TTT AAA ATT AGG ATT TTA AAA TCA AGC CAA AAA TCA  
 CCT ATT GCA GAG TCC CCA ATG AAA GAA ATT ACA GAT TGA AAG ATT ATC TCC GCC TAG GTT

<内含子

1090                | 1100                1110                1120                1130                1140

TGT GGA AAT ATT CTC ACC TGG AAC CCA GAG CAG CAG AAA CCC AAT GAG TTG TGA TGG CAA  
 <P V W L L F G I L Q S P L  
 <—g—g—g—g—g—前导序列—g—g—g—g—g—

<TATA盒

1150                1160 | 1170                1180                1190                | 1200

CAT GGT GGC CAC TCC GGA TCT GAT CTG ACG GTT CAC TAA ACC AGC TCT GCT TAT ATA GAC  
 <M

<来自M60321的最小CMV启动子 >CMV启动子

1210                1220                1230                | 1240                1250                1260

CTC CCA CCG TAC ACG CCT ACC GCC CAT TTG GGA TCT AGT AAT CAA TTA CGG GGT CAT TAG  
 TTC ATA GCC CAT ATA TGG AGT TCC GCG TTA CAT AAC TTA CGG TAA ATG GCC CGC CTG GCT  
 GAC CGC CCA ACG ACC CCC GCC CAT TGA CGT CAA TAA TGA CGT ATG TTC CCA TAG TAA CGC  
 CAA TAG GGA CTT TCC ATT GAC GTC AAT GGG TGG AGT ATT TAC GGT AAA CTG CCC ACT TGG  
 CAG TAC ATC AAG TGT ATC ATA TGC CAA GTA CGC CCC CTA TTG ACG TCA ATG ACG GTA AAT  
 GGC CGC CCT GGC ATT ATG CCC AGT ACA TGA CCT TAT GGG ACT TTC CTA CTT GGC AGT ACA  
 TCT ACG TAT TAG TCA TCG CTA TTA CCA TGC TGA GGT TTT GGC AGT ACA TCA ATG GGC  
 GTG GAT AGC GGT TTG ACT CAC GGG GAT TTC CAA GTC TCC ACC CCA TTG ACG TCA ATG GGA  
 GTT TGT TTT GGC ACC AAA ATC AAC GGG ACT TTC CAA AAT GTC GTA ACA ACT CCG CCC CAT

>TATA盒

1750                1760                1770                1780                1790                1800

TGA CGC AAA TGG GCG GTA GGC GTG TAC GGT GGG AGG TCT ATA TAA GCA GAG CTG GTT TAG

>Cla1

1810                1820 | 1830                1840                1850                1860

TGA ACC GTC AGA TCA GAT CCA TCG ATT GGC CAC CAT GAG TTG GAG CTG TAT CAT CCT CTT  
 M S W S C I I L F  
 —h—h—h—前导序列—h—h—h—>

>内含子

1870                1880 | 1890                1900                1910                1920

CTT GGT AGC AAC AGC TAC AGG TAA GGG GTT AAC AGT AGC AGG CTT GAG GTC TGG ACA TAT  
 L V A T A T>  
 —h—h—前导序列—h—h—>

>Nco1

1930                1940                1950                1960                | 1970                1980

ATA TGG GTG ACA ATG ACA TCC ACT TTG CCT TTC TCT CCA CAG GCG CCA TGG CCC AGG TGC  
 Q V>  
 —i—i—>

图51(续)

图51(续)

2770 2780 2790 2800 2810 2820  
 CCT GCG TGG TGG TGG ACG TTT CAC AAG AGG ACC CCG AGG TGC AGT TTA ATT GGT ACG TGG  
 T C V V V D V S Q E D P E V Q F N W Y V>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 2830 2840 2850 2860 2870 2880  
 ACG GCG TGG AAG TGC ACA ACG CCA AGA CCA AGC CTA GAG AGG AAC AGT ACA ACA GCA CCT  
 D G V E V H N A K T K P R E E Q Y N S T>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 2890 2900 2910 2920 2930 2940  
 ACA GAG TGG TGT CCG TGC TGA CAG TGC ACC AGG ATT GGC TGA ACG GCA AAG AGT ACA  
 Y R V V S V L T V L H Q D W L N G K E Y>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 2950 2960 2970 2980 2990 3000  
 AGT GCA AGG TGT CCA ACA AGG GCC TGC CAA GCA GCA TCG AGA AAA CCA TCA GCA AGG CCA  
 K C K V S N K G L P S S I E K T I S K A>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 3010 3020 3030 3040 3050 3060  
 AGG GCC AGC CTA GGG AAC CCC AGG TTT ACA CAC TGC CTC CAA GCC AGA AAG AGA TGA CCA  
 K G Q P R E P Q V Y T L P P S Q K E M T>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 3070 3080 3090 3100 3110 3120  
 AGA ACC AGG TGT CCC TGA CCT GCC TCG TGA AGG GCT TCT ACC CTT CCG ATA TCG CCG TGG  
 K N Q V S L T C L V K G F Y P S D I A V>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 3130 3140 3150 3160 3170 3180  
 AAT GGG AGA GCA ATG GCC AGC CAG AGA ACA ACT ACA AGA CCA CAC CTC CTG TGC TGG ACA  
 E W E S N G Q P E N N Y K T T P P V L D>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 3190 3200 3210 3220 3230 3240  
 GCG ACG GCT CAT TCT TCC TGT ACA GCA AGC TGA CCG TGG ACA AGA GCA GAT GGC AAG AGG  
 S D G S F F L Y S K L T V D K S R W Q E>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>  
  
 3250 3260 3270 3280 3290 3300  
 GCA ACG TGT TCA GCT GCA GCG TGA TGC ACG AGG CCC TGC ACA ACA GAT ACA CCC AGA AGT  
 G N V F S C S V M H E A L H N R Y T Q K>  
 j j j j j CH1-3\_重链1-IGG2-杆 j j j j j j j>

>Myc-抗原决定基 >PDGFR\_TM  
 3310 3320 3330 3340 3350 3360  
 CCC TGT CTC TGA GCC CCG AAC AAA AAC TCA TCT CAG AAG AGG ATC TGA ATG CTG TGG GCC  
 S L S L S P>  
 CH1-3 重链1-JG i >

3370	3380	3390	3400	3410	3420
AGG ACA CGC AGG AGG TCA TCG TGG TGC CAC ACT CCT TGC CCT TTA AGG TGG TGG TGA TCT					
3430	3440	3450	3460	3470	3480
CAG CCA TCC TGG CCC TGG TGG TGC TCA CCA TCA TCT CCC TTA TCA TCC TCA TCA TGC TTT					

>前聚腺苷酸  
|  
3490            3500            3510            3520            3530            3540

GGC AGA AGA AGC CAC GTT AGT AAA AGC TTG TCA CTT GGA AAG TAA TAG TTT TTC CTG CAC  
GGG TAG TAA TCA GCC TCG ACT GTG CCT TCT AGT TGC CAG CCA TCT GTT GTT TGC CCC TCC

图51(续)

CCC GTG CCT TCC TTG ACC CTG GAA GGT GCC ACT CCC ACT GTC CTT TCC TAA TAA AAT GAG  
 GAA ATT GCA TCG CAT TGT CTG AGT AGG TGT CAT TCT ATT CTG GGG GGT GGG GTG GGG CAG  
 <BGH\_聚腺苷酸  
 |  
 3730            3740            3750            3760            3770            3780  
 GAC AGC AAG GGG GAG GAT TGG GAA GAC AAT AGC AGG CAT GCT GGG GAT GGC CCG CGC ATG

## &gt;人类\EF1a\启动子

|  
 3790            3800            3810            3820            3830            3840  
 ATA ACT TCG TAT AAT GTA TAT ACG AAG TTA TGT AGA GTA ATT CAT ACA AAA GGA CTC  
 GCC CCT GCC TTG GGG AAT CCC AGG GAC CGT CGT TAA ACT CCC ACT AAC GTA GAA CCC AGA  
 GAT CGC TGC GTT CCC GCC CCC TCA CCC GCC CGC TCT CGT CAT CAC TGA GGT GGA GAA GAG  
 CAT GCG TGA GGC TCC GGT GCC CGT CAG TGG GCA GAG CGC ACA TCG CCC ACA GTC CCC GAG  
 AAG TTG GGG GGA GGG GTC GGC AAT TGA ACC GGT GCC TAG AGA AGG TGG CGC GGG GTA AAC  
 TGG GAA AGT GAT GTC GTG TAC TGG CTC CGC CTT TTT CCC GAG GGT GGG GGA GAA CCG TAT  
 ATA AGT GCA GTA GTC GCC GTG AAC GTT CTT TTT CGC AAC GGG TTT GCC GCC AGA ACA CAG  
 GTA AGT GCC GTG TGT GGT TCC CGC GGG CCT GGC CTC TTT ACG GGT TAT GGC CCT TGC GTG  
 CCT TGA ATT ACT TCC ACG CCC CTG GCT GCA GTA CGT GAT TCT TGA TCC CGA GCT TCG GGT  
 TGG AAG TGG GTG GGA GAG TTC GAG GCC TTG CGC TTA AGG AGC CCC TTC GCC TCG TGC TTG  
 AGT TGA GGC CTG GCT TGG CGC CTG GGG CGG CGC CGT GCG AAT CTG GTG GCA CCT TCG CGC  
 CTG TCT CGC TGC TTT CGA TAA GTC TCT AGC CAT TTA AAA TTT TTG ATG ACC TGC TGC GAC  
 GCT TTT TTT CTG GCA AGA TAG TCT TGT AAA TGC GGG CCA TGA TCT GCA CAC TGG TAT TTC  
 GGT TTT TGG GGC CGG CGG CGA CGG CGC CCG TGC GTC CCA CGC CAC ATG TTC CGC GAG  
 GCG GGG CCT CGC AGC GCG GCC ACC GAG AAT CGG ACG GGG GTA GTC TCA AGC TGG CCG GCC  
 TGC TCT GGT GCC TGG CCT CGC GCC GTG TAT CGC CCC GCC CTG GGC GGC AAG GCT GGC  
 CCG GTC GGC ACC AGT TGC GTG AGC GGA AAG ATG GCC GCT TCC CGG CCC TGC TGC AGG GAG  
 CTC AAA ATG GAG GAC GCG GCG CTC GGG AGA GCG GGC GGG TGA GTC ACC CAC ACA AAG GAA  
 AAG GGC CTT TCC GTC CTC AGC CGT CGC TTC ATG TGA CTC CAC CGG GTA CCG GGC CGC GTC  
 CAG GCA CCT CGA TTA GTT CAC GAG CTT TTG GAG TAC GTC GTC TTT AGG TTG GGG GGA GGG  
 GTT TTA TGC GAT GGA GTT TCC CCA CAC TGA GTG GGT GGA GAC TGA AGT TAG GCC AGC TTG  
 GCA CTT GAT GTA ATT CTC CTT GGA ATT TGC CCT TTT TGA GTT TGG ATC TTG GTT CAT TCT  
 CAA GCC TCA GAC AGT GGT TCA AAG TTT TCT TCC ATT TCA GGT GTC GTG AGA GCT CGT  
 TTA GTG AAC CGT CAG ATC GCC TGG AGA CGC CAT CCA CGC TGT TTT GAC CTC CAT AGA AGA  
 CAC CGG GAC CGA TCC AGC CTC CGC GGC CCC GAA TTC AAC ATG GAC TGG ACC TGG AGG GTC

M D W T W R V>  
 \_\_\_\_1\_\_\_\_SIGPP-部分1\_1\_\_\_\_1\_\_\_\_>

## &gt;剪接供体

## &gt;内含子

|  
 5290            5300            5310            5320            5330            5340  
 TTC TGC TTG CTG GCT GTA GCT CCA GGT AAA GGG CCA ACT GGT TCC AGG GCT GAG GAA GGG  
 F C L L A V A P G>  
 \_\_\_\_1\_\_\_\_SIGPP-部分1\_1\_\_\_\_1\_\_\_\_1\_\_\_\_>

5350            5360            5370            5380            5390            5400  
 ATT TTT TCC AGT TTA GAG GAC TGT CAT TCT CTA CTG TGT CCT CTC CGC AGG TGC TCA CTC  
 5410            5420            5430            5440            5450            5460  
 CCA GGT TCA GCT GGT GCA GTC TGG CAG CGA GCT GAA AAA ACC TGG CGC CTC CGT GAA GGT  
 Q V Q L V Q S G S E L K K P G A S V K V>  
 \_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_>

5470            5480            5490            5500            5510            5520  
 GTC CTG CAA GGC TTC TGG CTA CAC CTT TAC CGA CAA CAA CAT GGA CTG GGT CCG ACA GGC  
 S C K A S G Y T F T D N N M D W V R Q A>  
 \_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_>

5530            5540            5550            5560            5570            5580  
 CCC TGG ACA AGG ACT TGA GTG GAT GGG CGA CAT CAA CAC CAG AAG CGG CGG CAG CAT CTA  
 P G Q G L E W M G D I N T R S G G S I Y>  
 \_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_n\_\_\_\_>

图51(续)

图51(续)

图51(续)

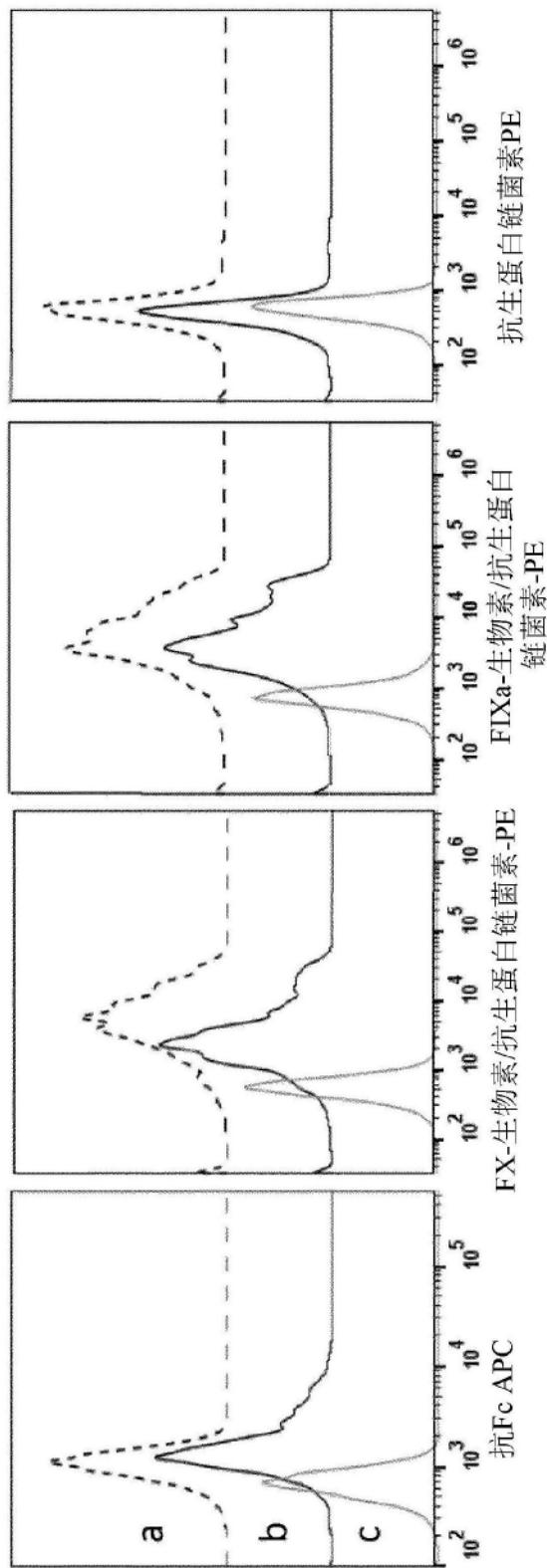


图52

艾密次单抗 E30Y E30Y_E55Y E30Y_E55Y_D93S E30Y_K54R_E55Y_D93S	10	20	— CDR1 —	40	50	— CDR2 —	60
	DIQMTQSPSS	LSASVGDRV	ITCKASRNIE	RQLAWYQQKP	GQAPELLIYQ	ASRKESGV	PD
	.....	.....	.....	Y .....	.....	.....	.....
	.....	.....	.....	Y .....	.....	.....	Y .....
	.....	.....	.....	Y .....	.....	.....	RY .....
艾密次单抗 E30Y E30Y_E55Y E30Y_E55Y_D93S E30Y_K54R_E55Y_D93S	70	80	— CDR3 —	100	110		
	RFSGSRYGTD	FTLTSSLQP	EDIATYYCQQ	YSDPPLTFGG	GTKVEIKRTA		
	.....	.....	.....	.....	.....	.....	.....
	.....	.....	.....	..S .....	.....	.....	.....
	.....	.....	.....	..S .....	.....	.....	.....

图53

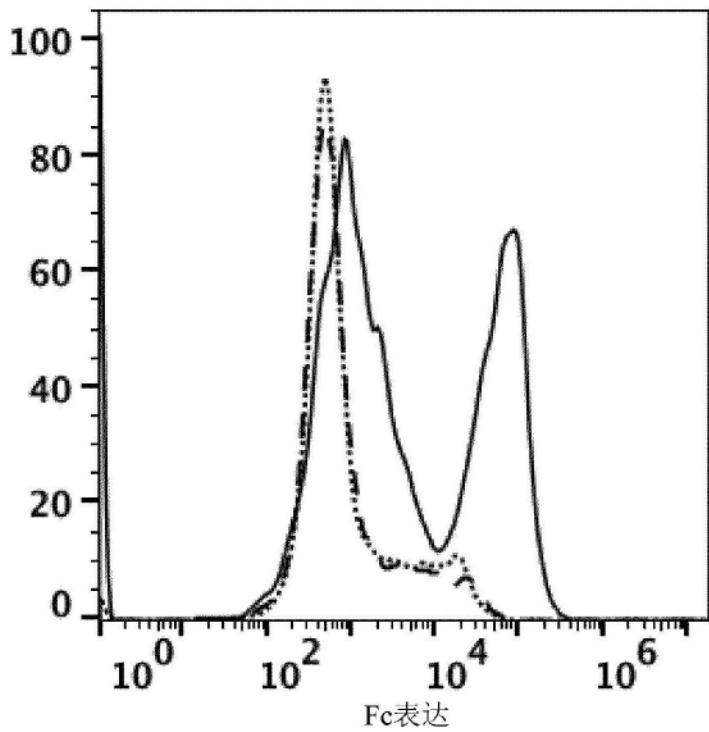
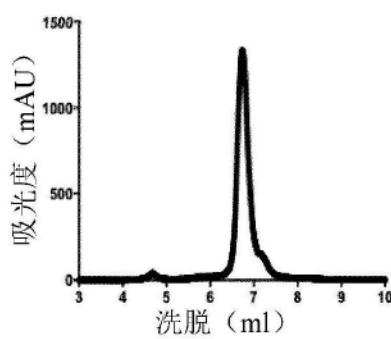
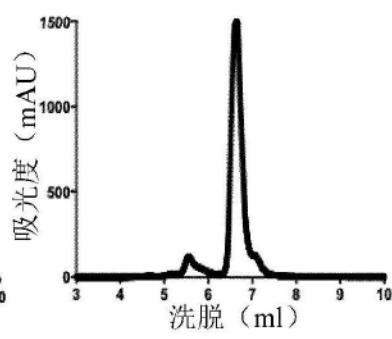
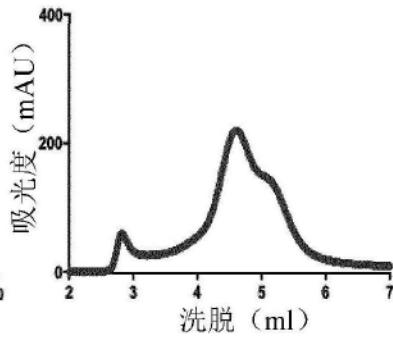
**a****b****c****d**

图54

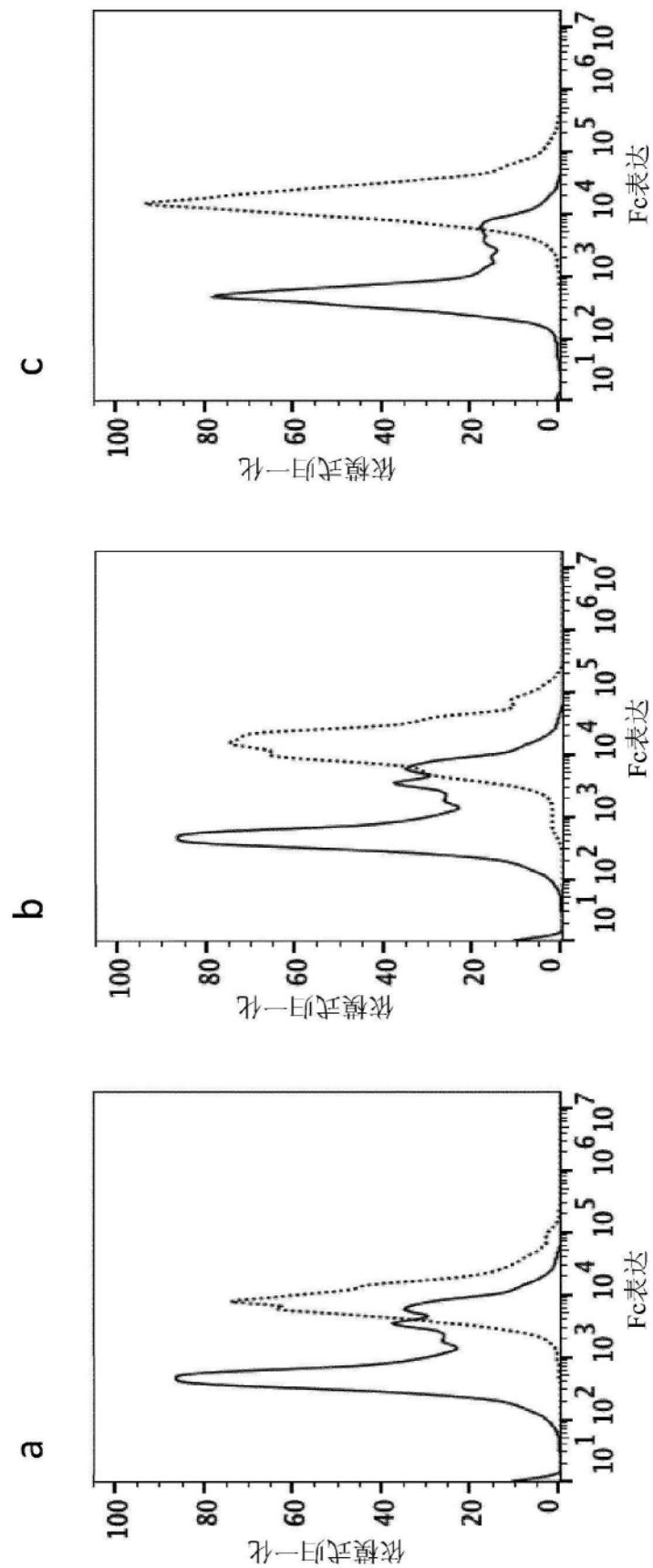


图55